General linear model and classical inference

SPM for M/EEG
May 2019

Martin Dietz
CFIN, Aarhus University, DK

AARHUS UNIVERSITY

Overview

Introduction

- Statistical parametric maps of MEG/EEG data
- ERP example

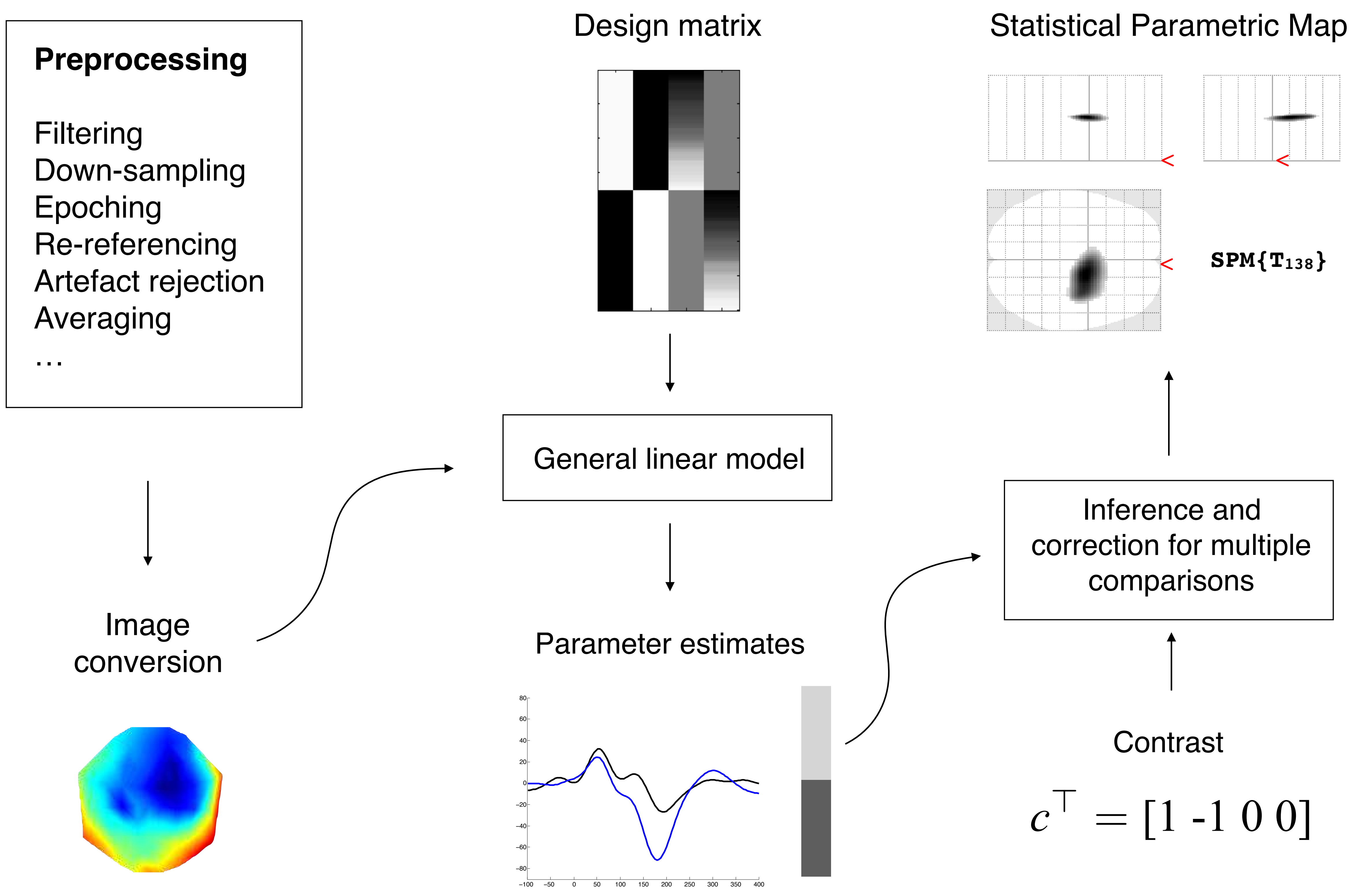
General linear model (GLM)

- Definition & design matrix
- Parameter estimation

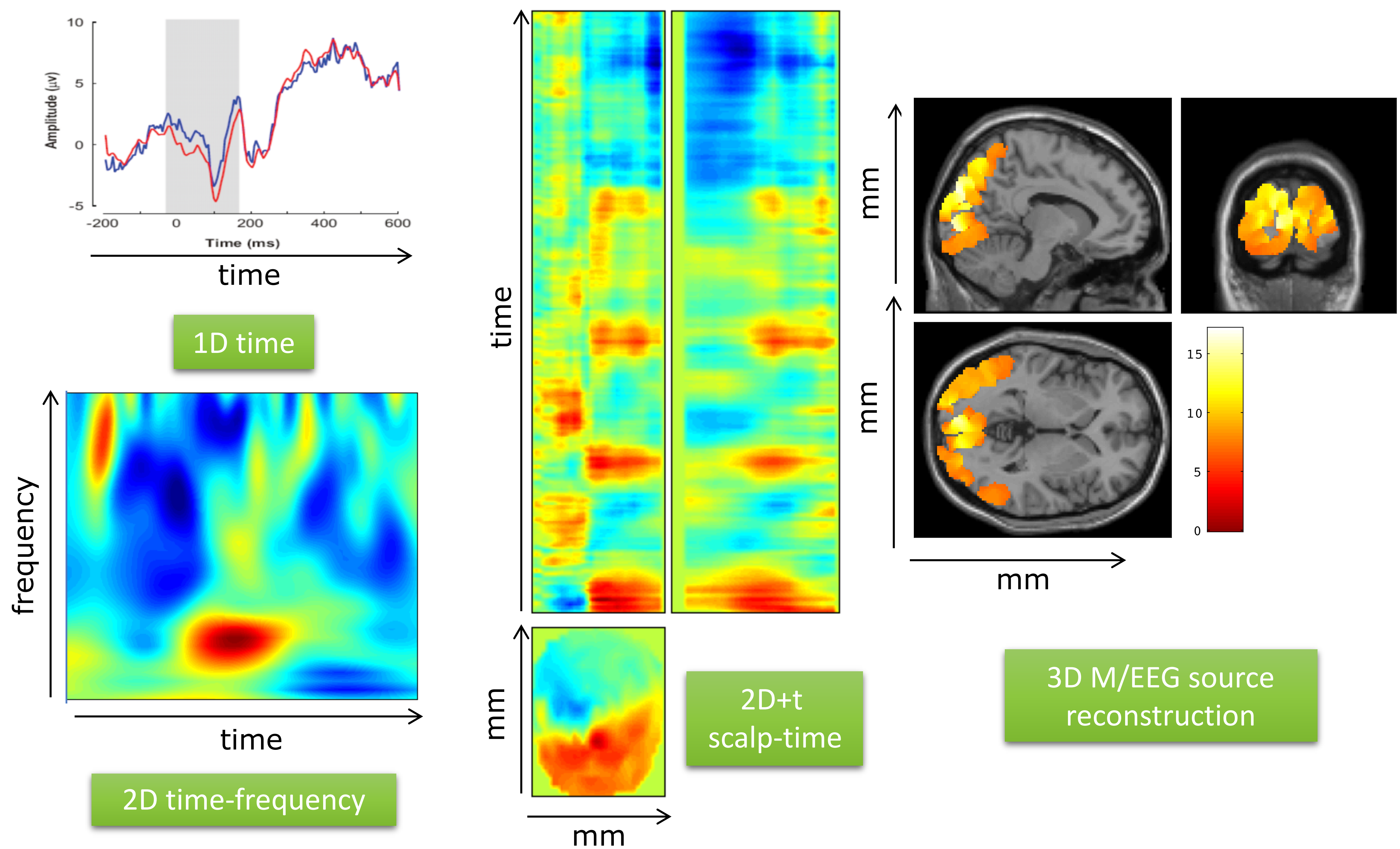
Classical inference

- Contrasts and inference
- Correlated regressors

Overview of SPM for M/EEG

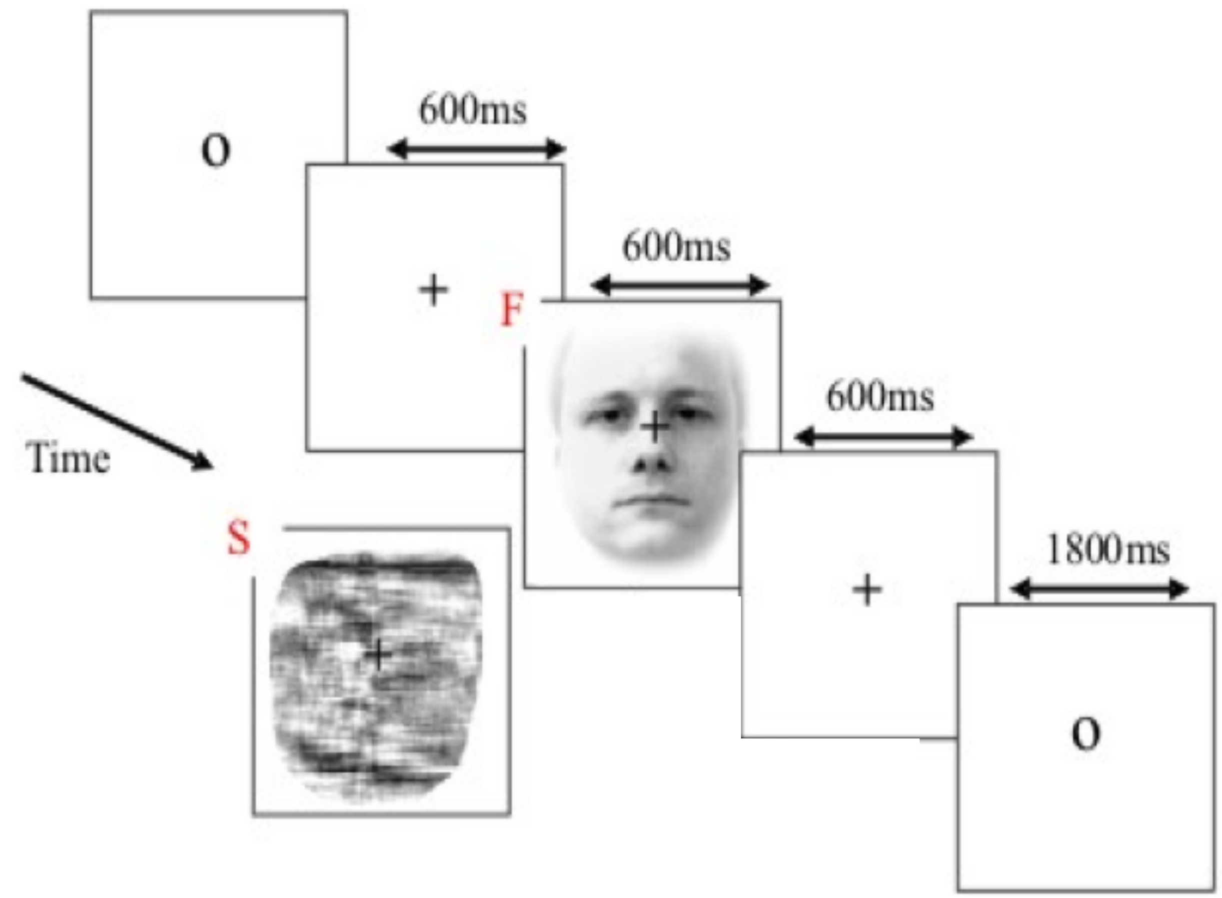


Statistical parametric maps



Event-related potential (ERP) example

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

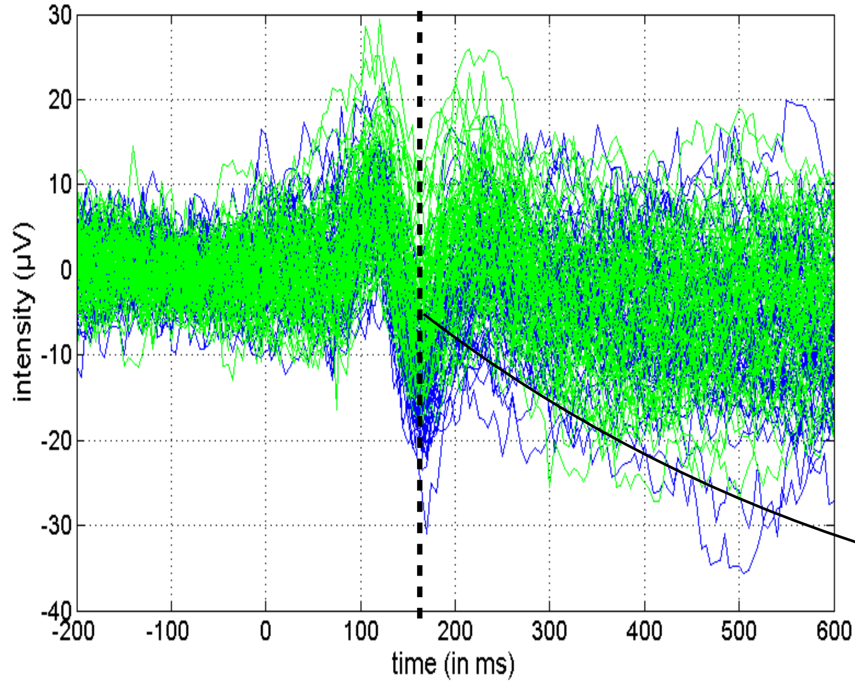


Question:

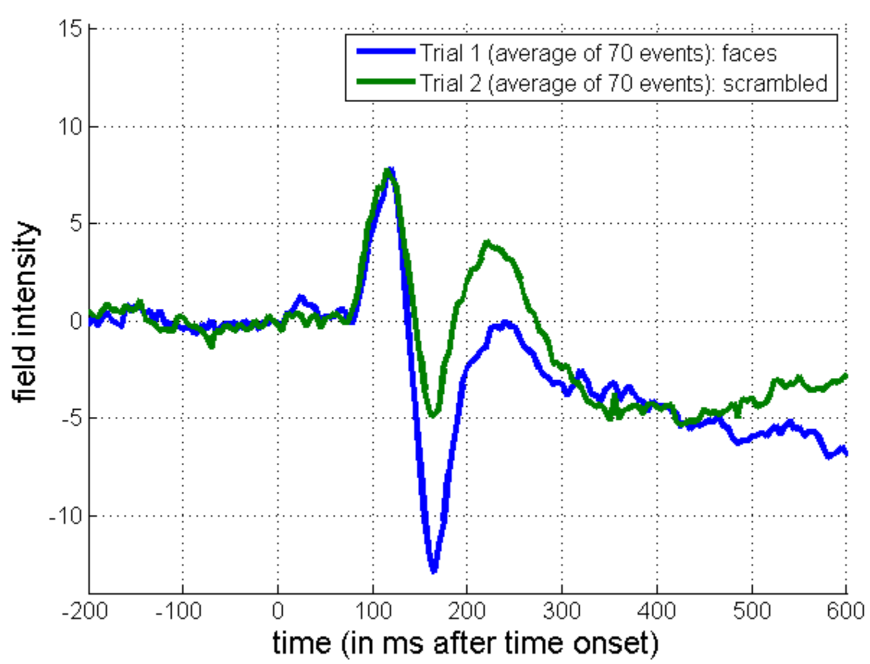
Is there a difference between the ERP of '*faces*' and '*scrambled faces*' ?

ERP example: ones channel

Trial-by-trial variability

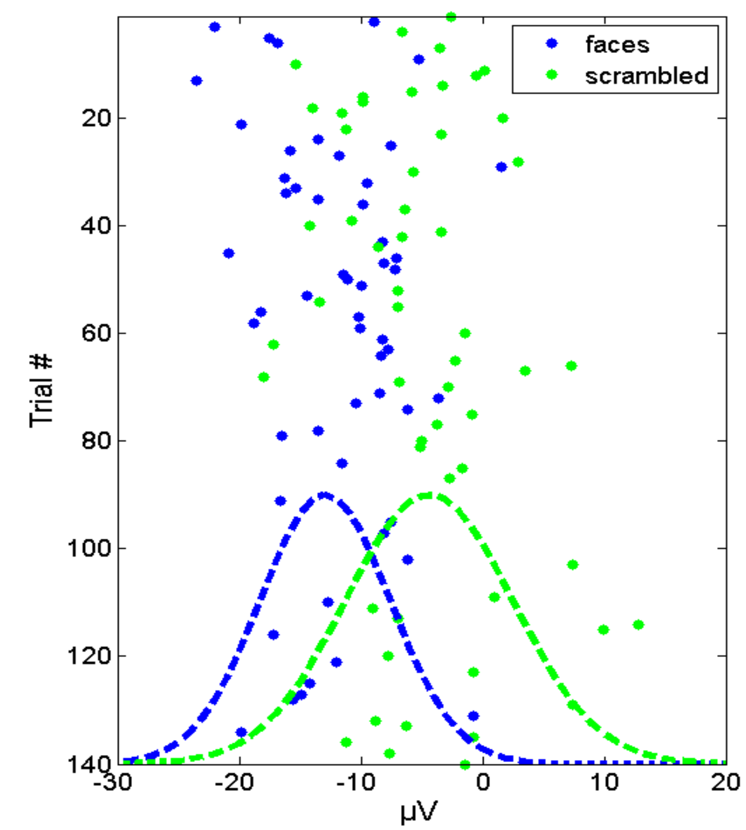


Average



Compare size of effect
to its standard error

(*Signal-to-noise ratio*)



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

Overview

Introduction

- Statistical parametric maps of MEG/EEG data
- ERP example

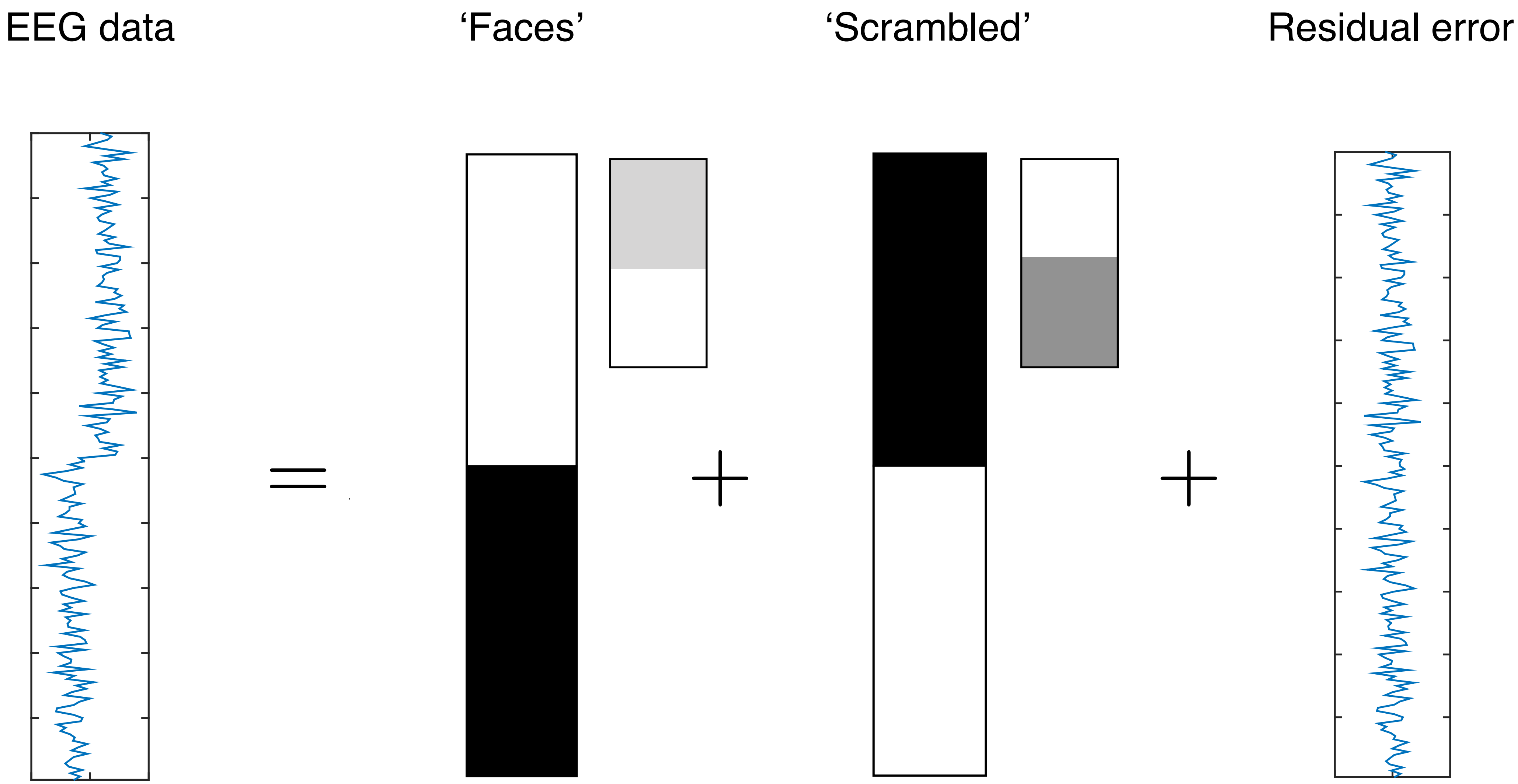
General linear model (GLM)

- Definition & design matrix
- Parameter estimation

Classical inference

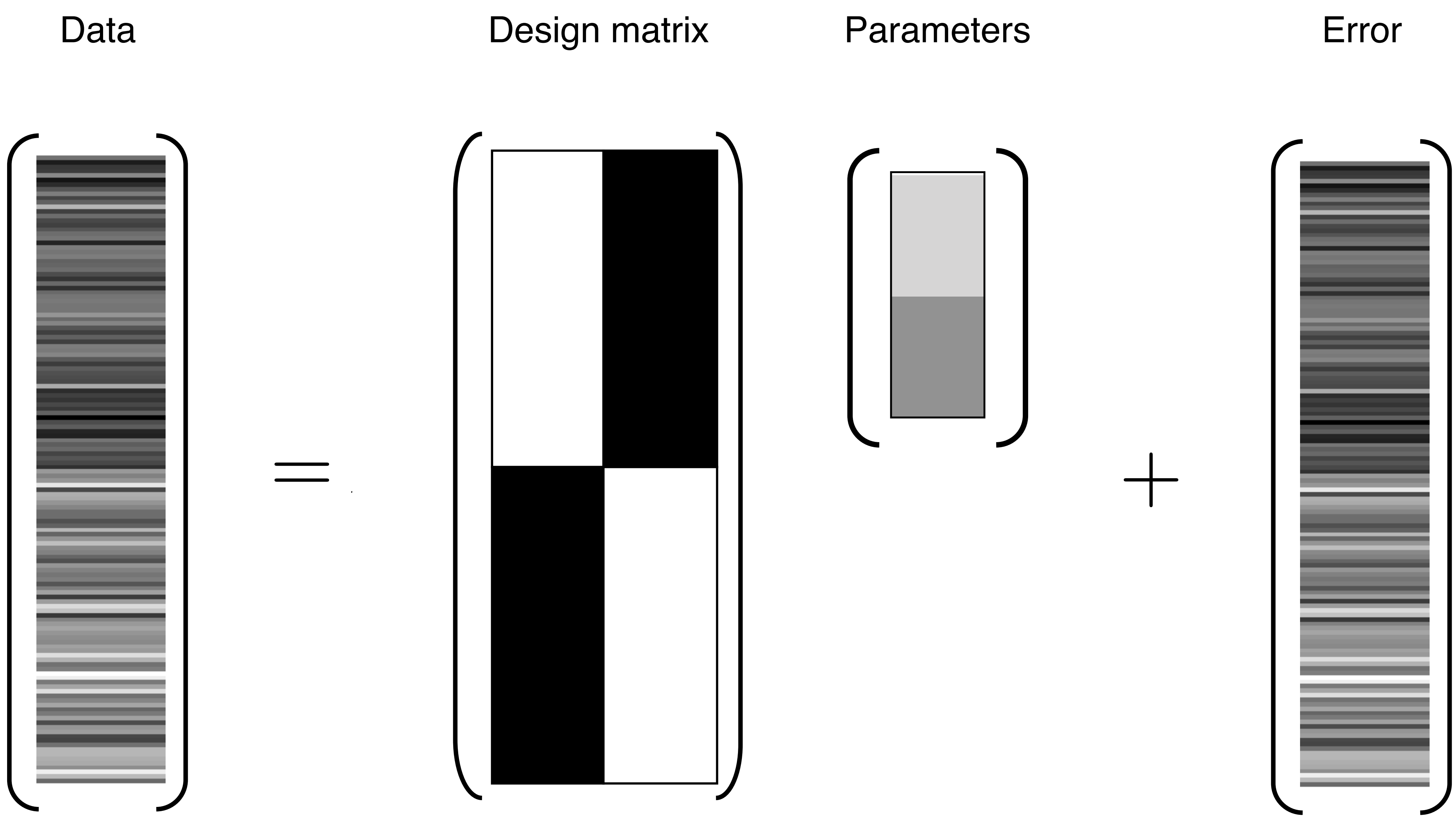
- Contrasts and inference
- Correlated regressors

Data modelling



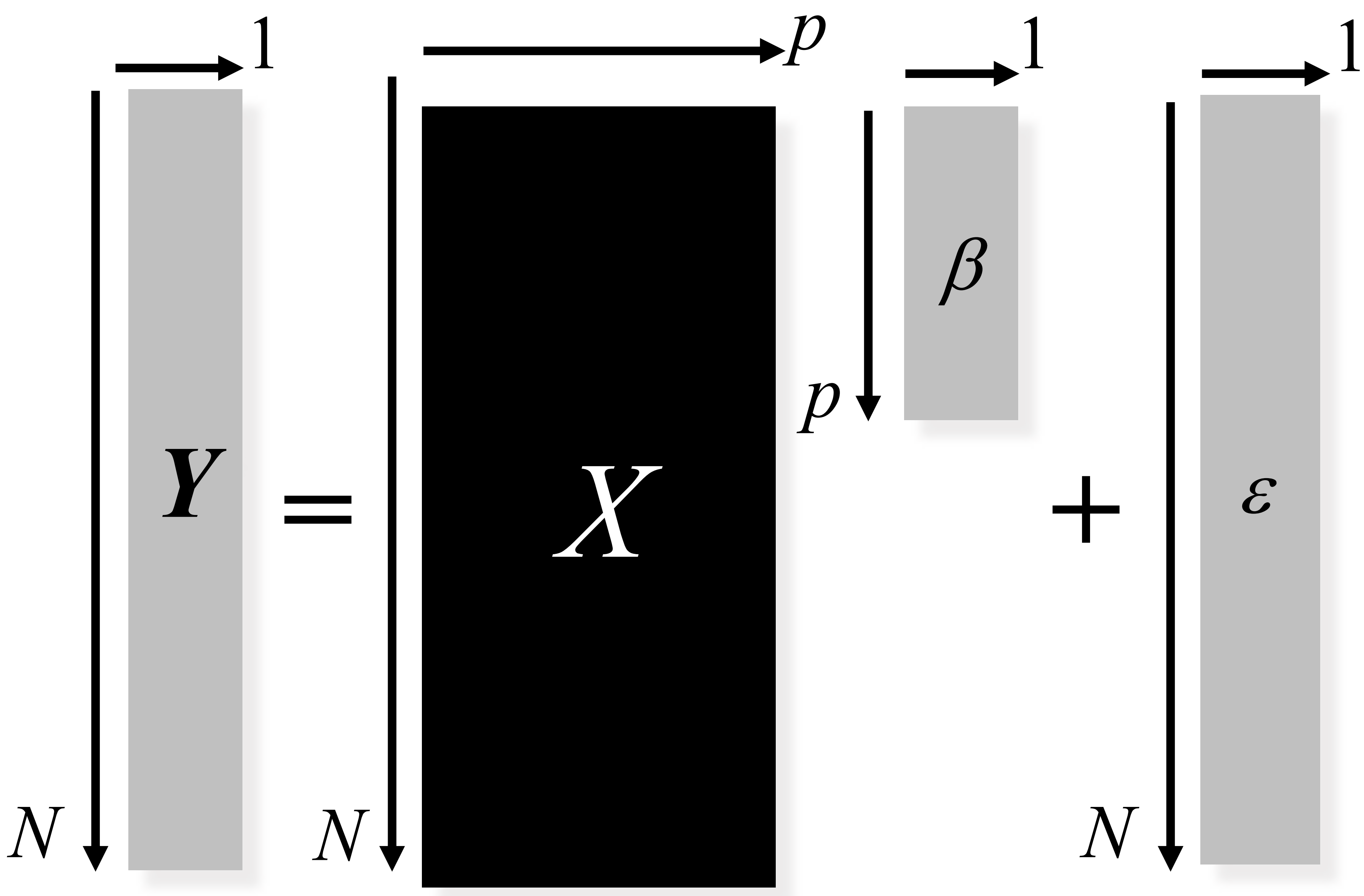
$$Y = x_1\beta_1 + x_2\beta_2 + \epsilon$$

General linear model



$$Y = X \cdot \beta + \epsilon$$

General linear model



GLM defined by

$$Y = X\beta + \epsilon$$

$$\epsilon \sim \mathcal{N}(0, \sigma^2 I)$$

The design matrix *X* embodies all knowledge about *experimentally controlled factors* and *confounds*

General linear model

Linear regression

$$y = x_0\beta_0 + x_1\beta_1 + ... + x_p\beta_p + \epsilon$$

Matrix form

$$Y = X\beta + \epsilon$$

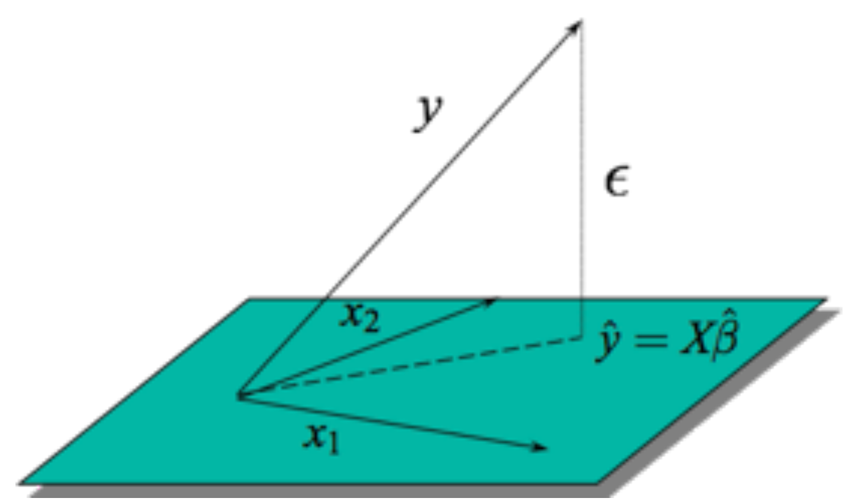
$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} x_{11} & \dots & x_{1p} & \dots & x_{1P} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{np} & \dots & x_{nP} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{N1} & \dots & x_{Np} & \dots & x_{NP} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \\ \vdots \\ \beta_P \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \\ \vdots \\ \epsilon_N \end{bmatrix}$$

General linear model

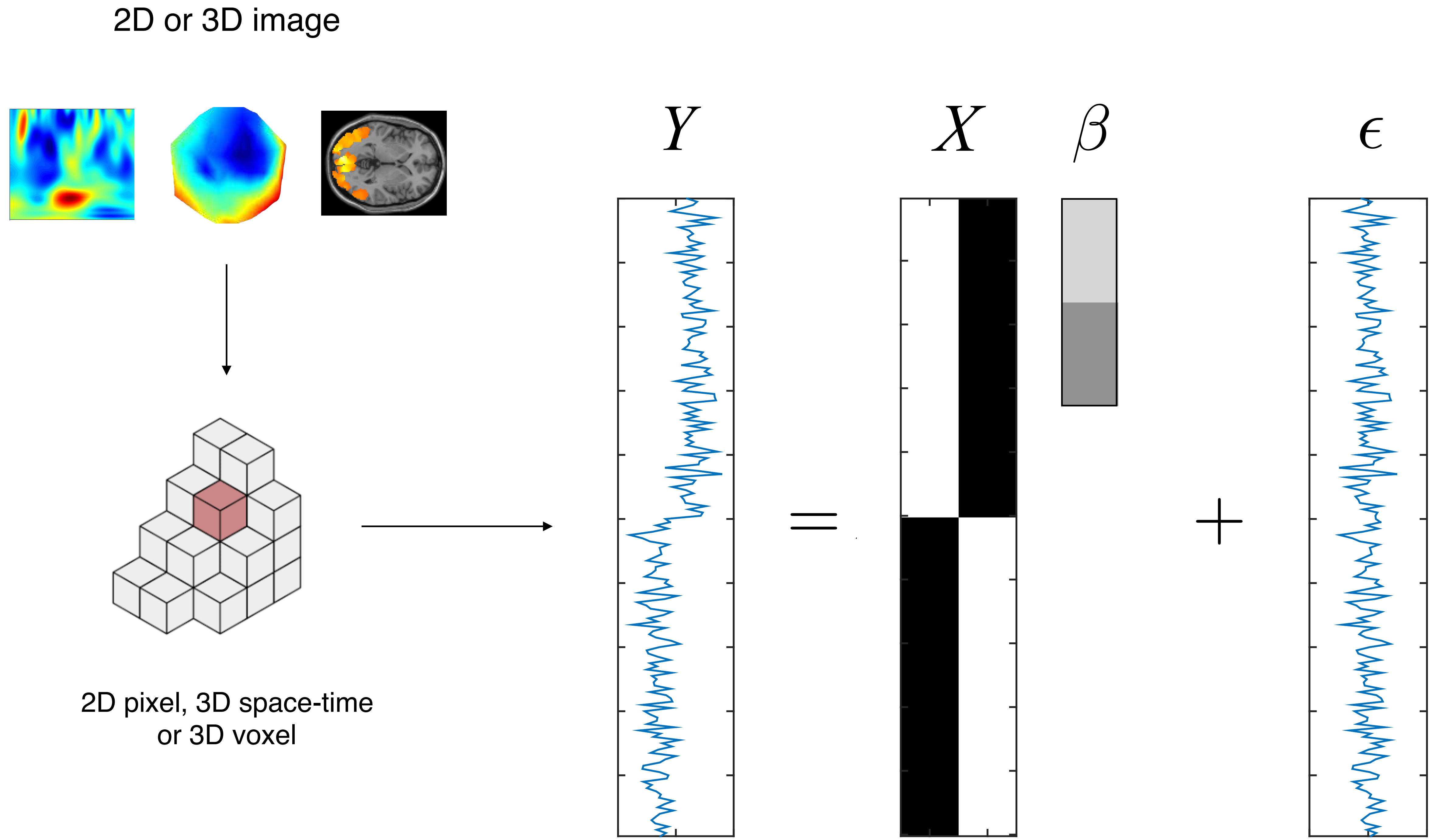
Special cases of the GLM

- *t*-test
- *F*-test
- Analysis of variance (ANOVA)
- Analysis of covariance (AnCova)
- multiple regression

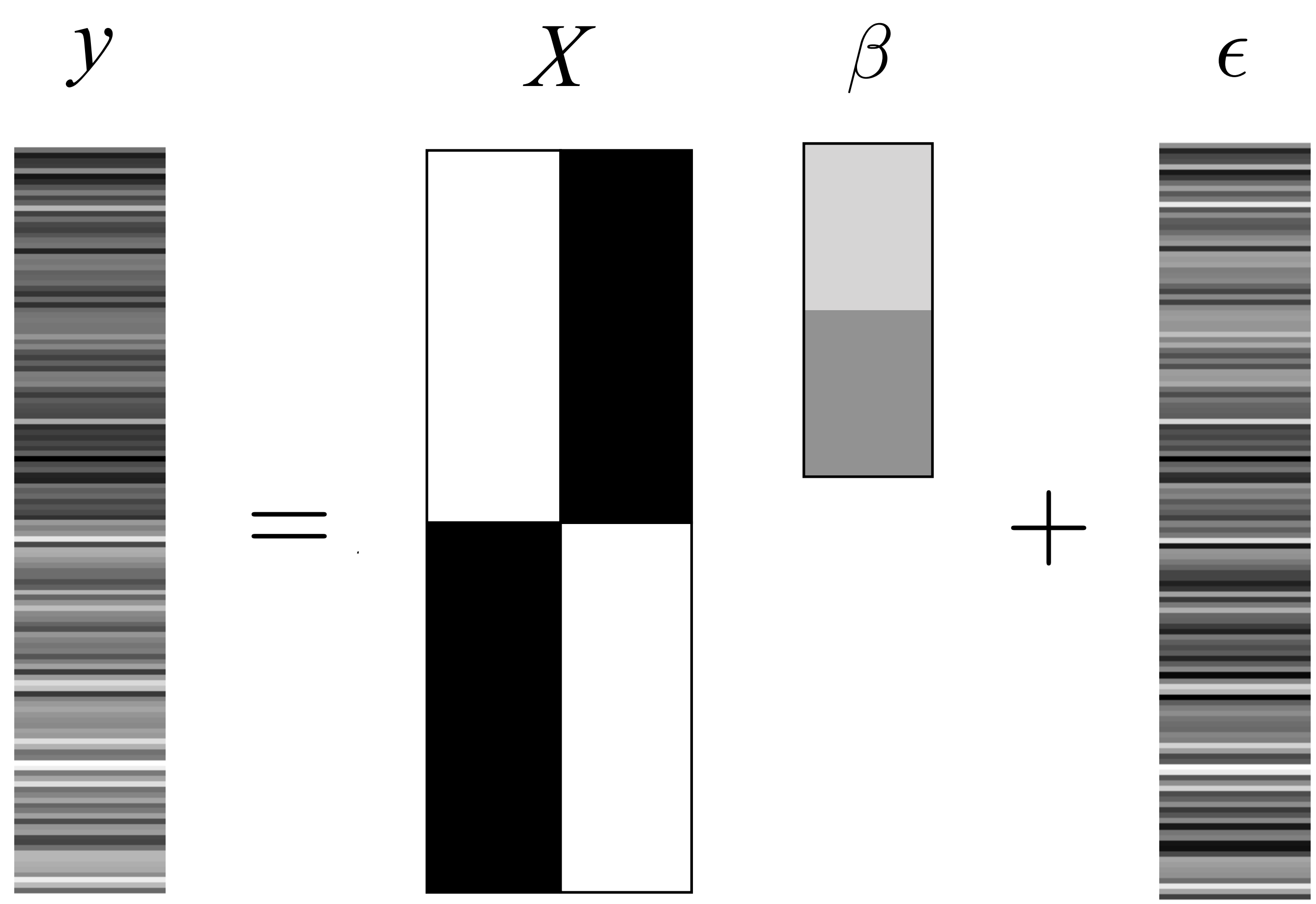
$$Y = X\beta + \epsilon$$



Voxel-wise general linear model



Parameter estimation



Iff residual error is i.i.d.

$$\epsilon = y - X\hat{\beta}$$

$$\epsilon \sim \mathcal{N}(0, \sigma^2 I)$$

↓

Ordinary least squares

minimise $\sum_i \hat{\epsilon}_i^2$ w.r.t $\hat{\beta}$ ⟶ $\hat{\beta} = (X^\top X)^{-1} X^\top y$

Parameter estimation: a geometric perspective

Ordinary least squares

$$\hat{\beta} = (X^\top X)^{-1} X^\top y$$

Residual forming matrix

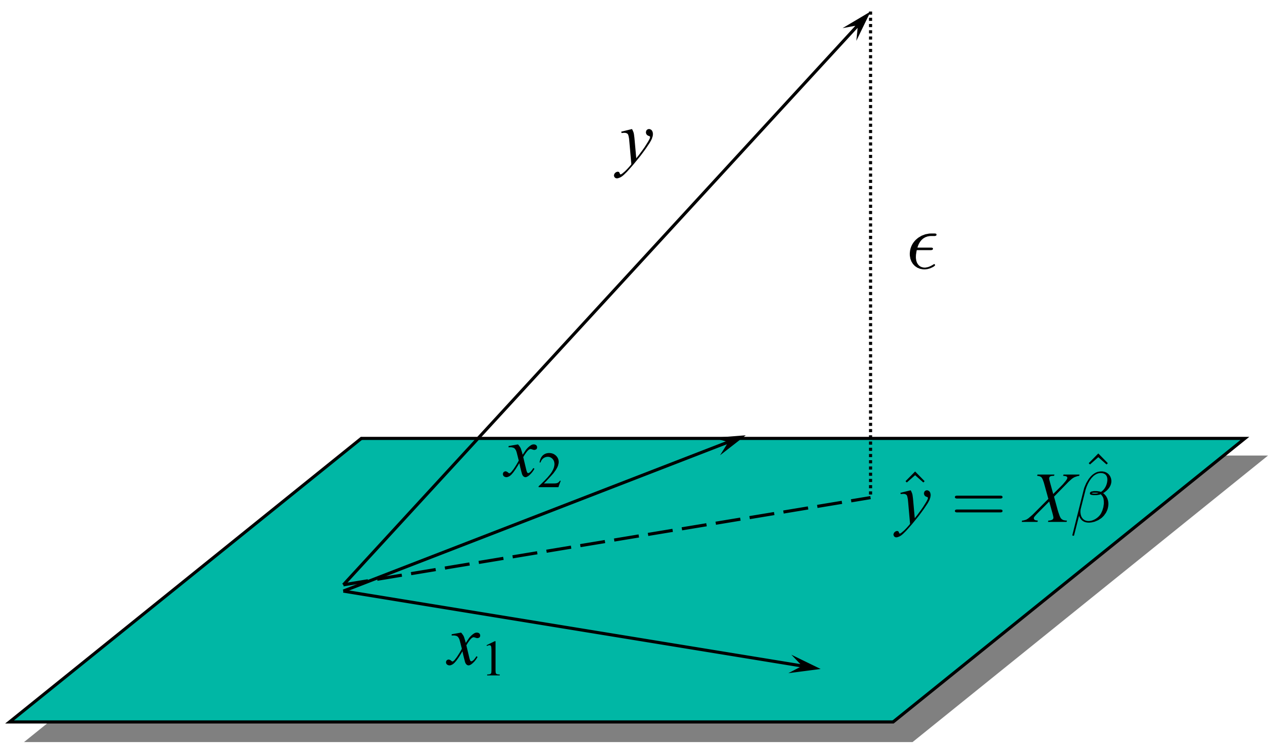
$$\epsilon = Ry$$

$$R = I - P$$

Projection matrix

$$\hat{y} = Py$$

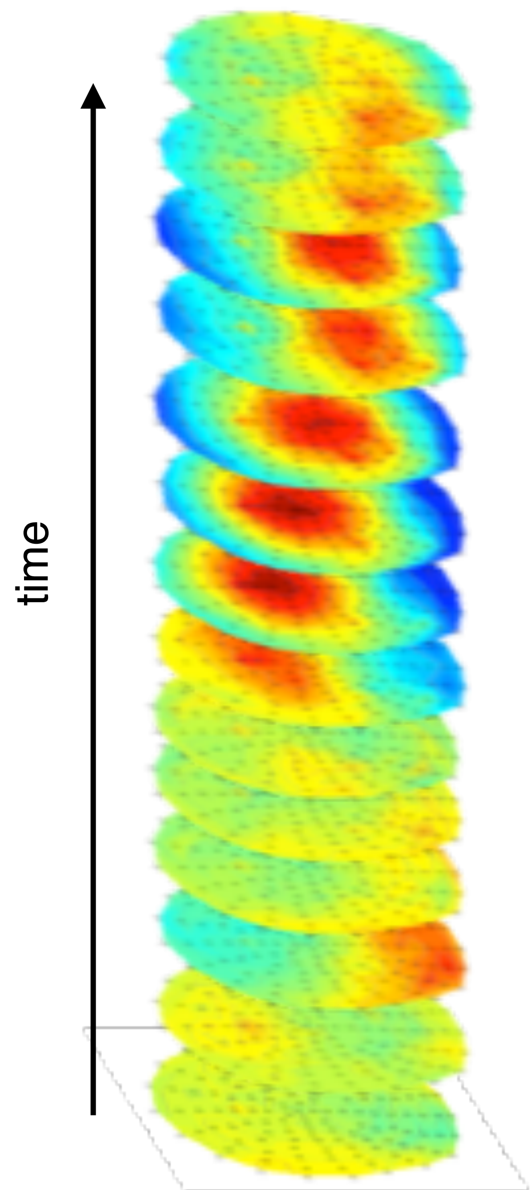
$$P = X(X^\top X)^{-1} X^\top$$



Design space of *X*

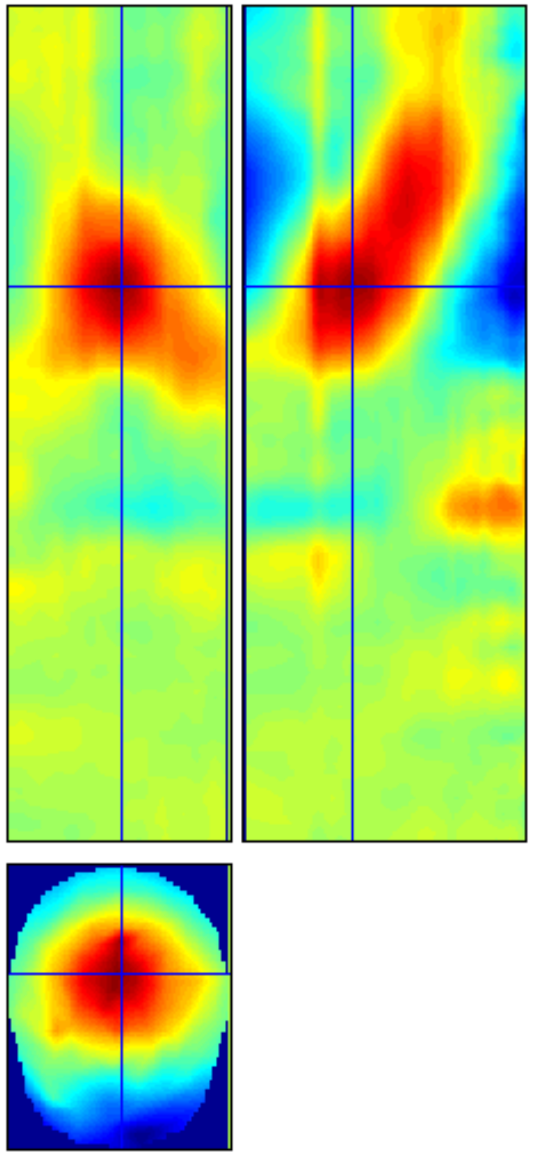
Channels to voxels transformation

M/EEG channels

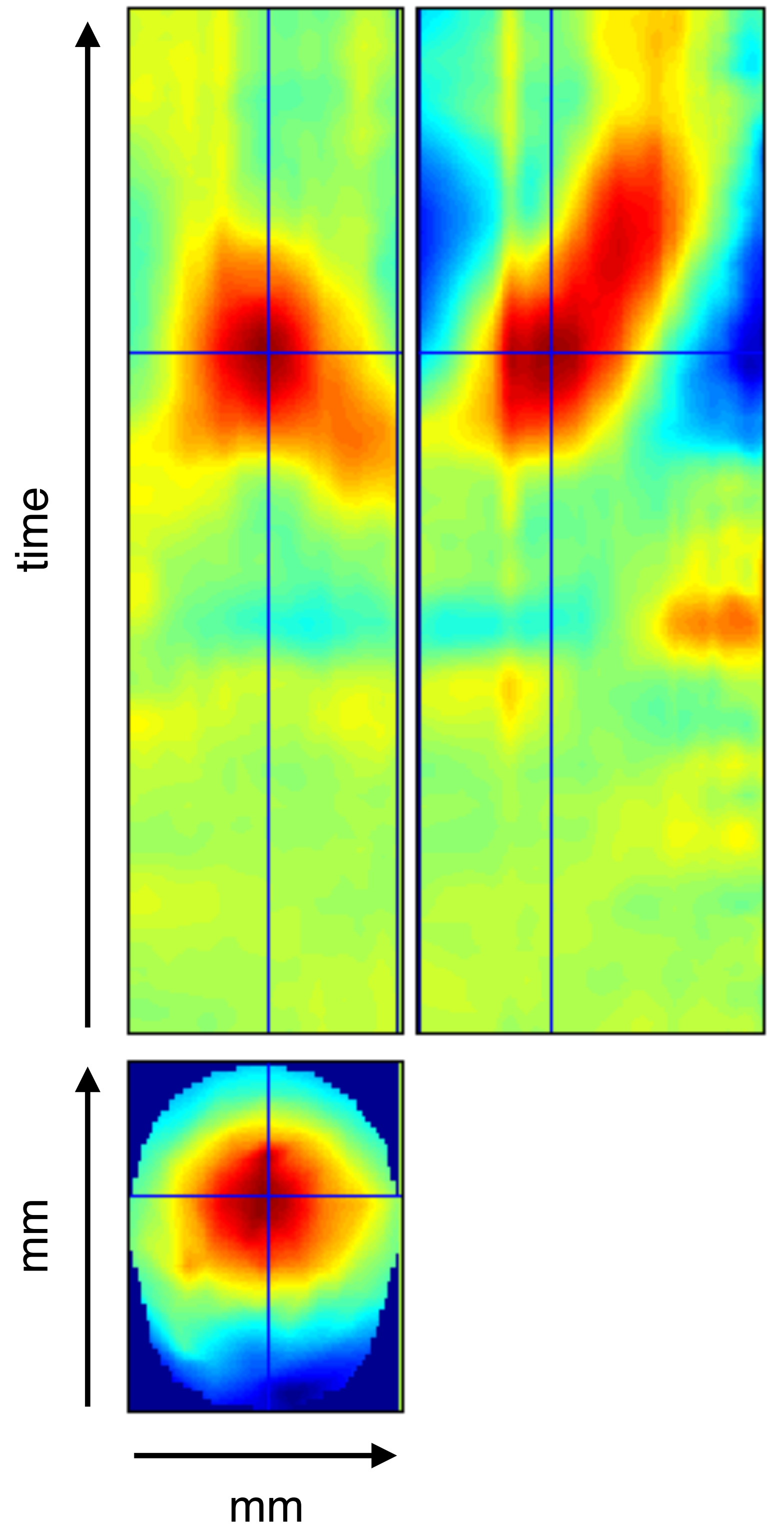


Spatio-temporal
interpolation

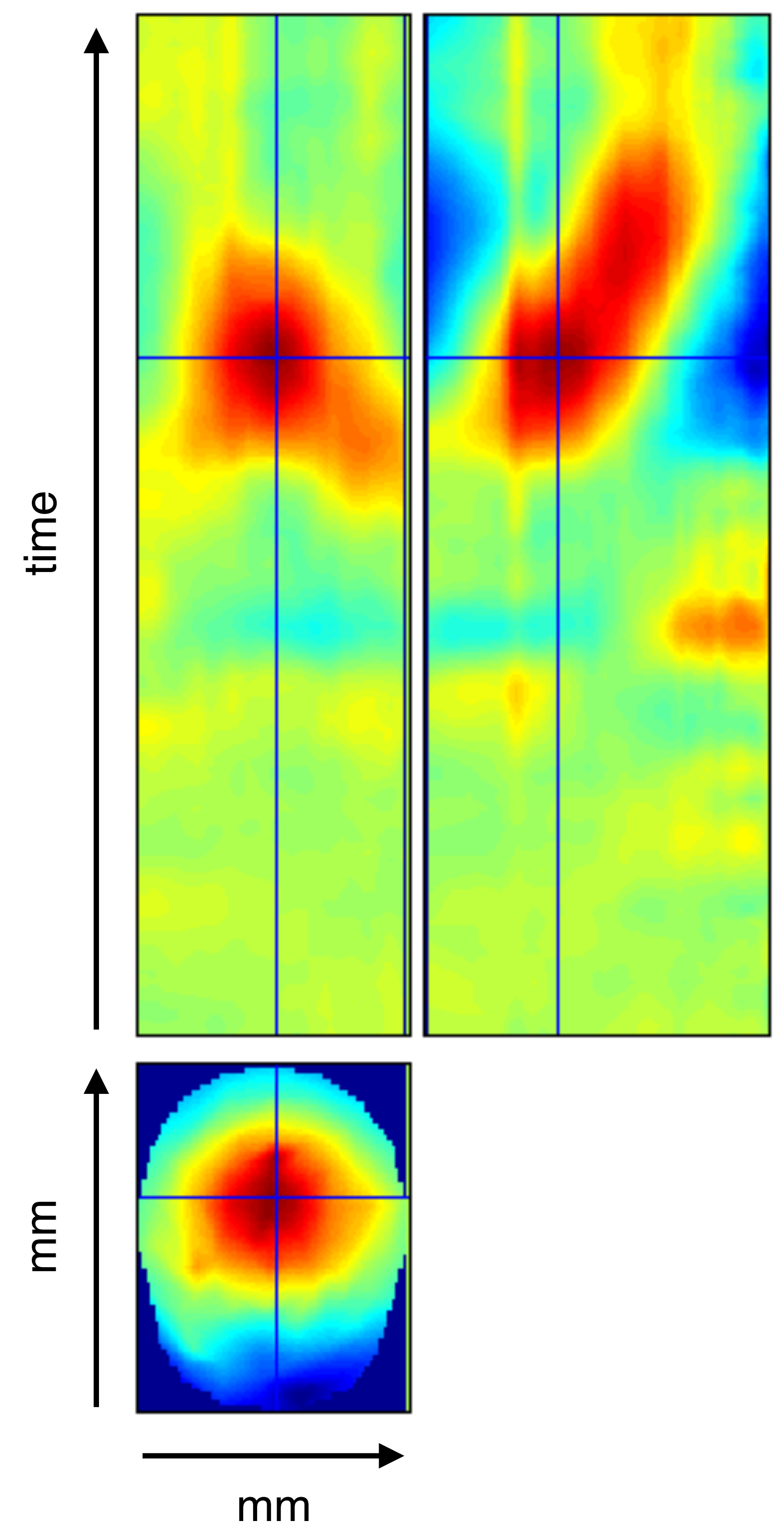
3D image (Nifti)



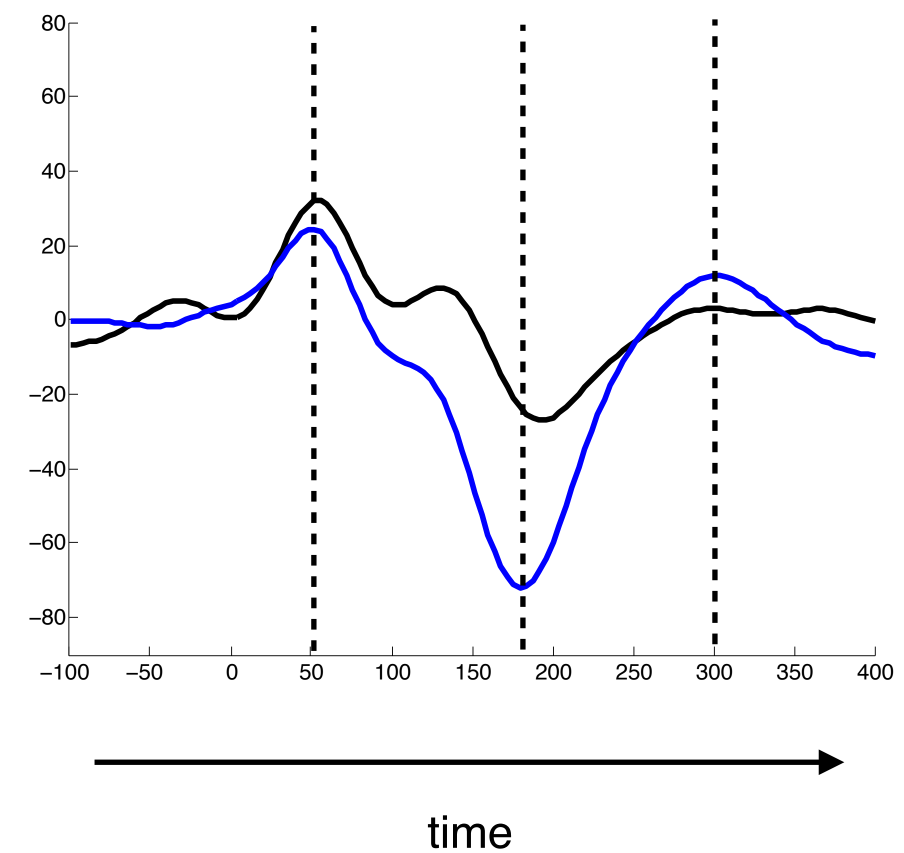
Mass-univariate statistical framework



Mass-univariate statistical framework



Avoid selection bias ("cherry picking")



Overview

Introduction

- Statistical parametric maps of MEG/EEG data
- ERP example

General linear model (GLM)

- Definition & design matrix
- Parameter estimation

Classical inference

- Contrasts and inference
- Correlated regressors

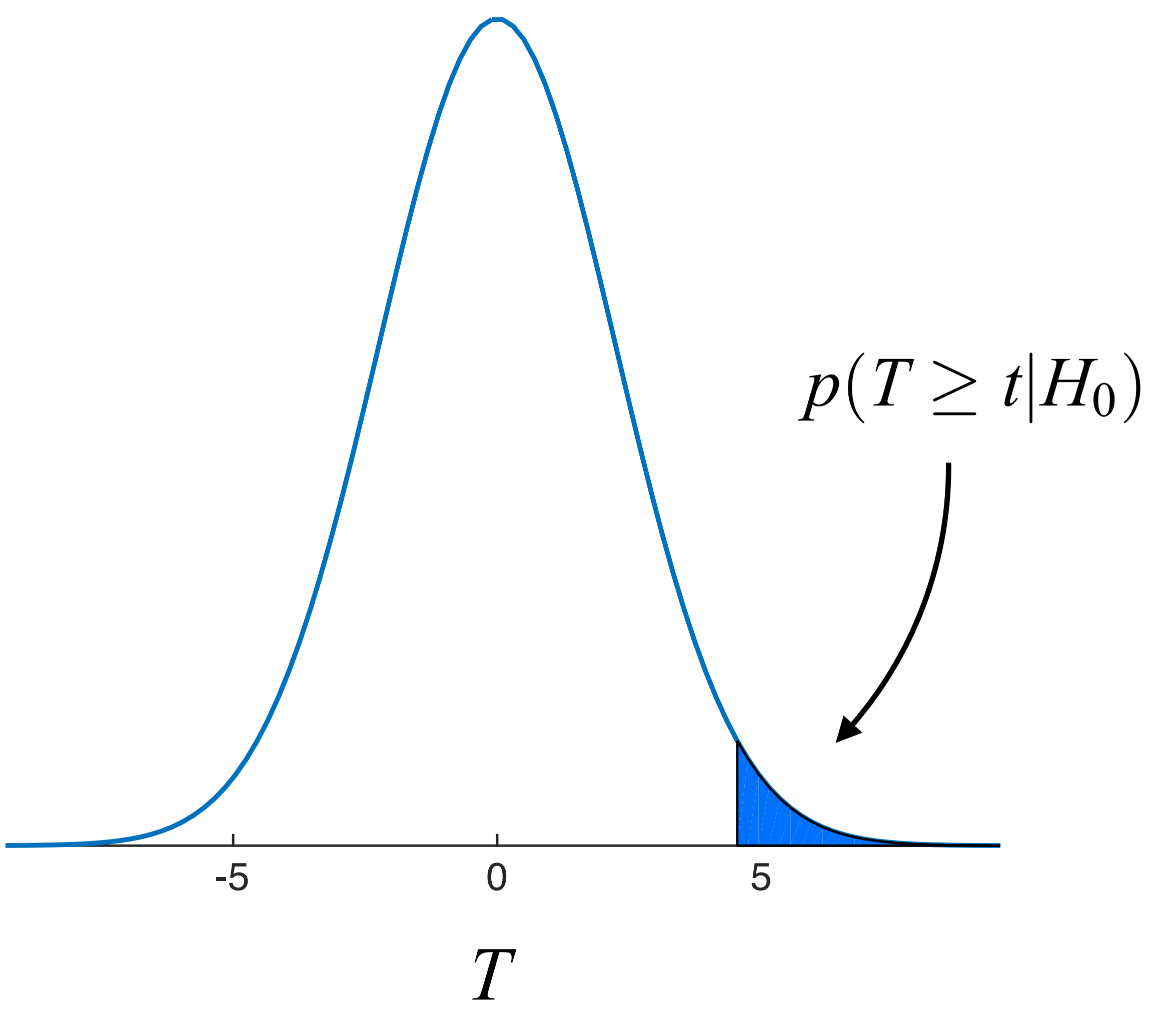
Hypothesis testing in classical inference

The null hypothesis *H_0*

No effect under null distribution

Alternative hypothesis *H_A*

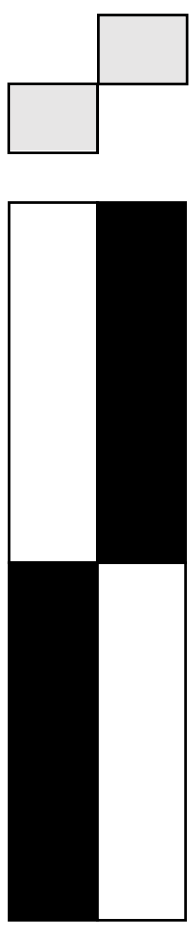
Unlikely or surprising effect
under null distributions



Contrast and t-test

Contrast vector

$$c^\top = [1 \; -1]$$



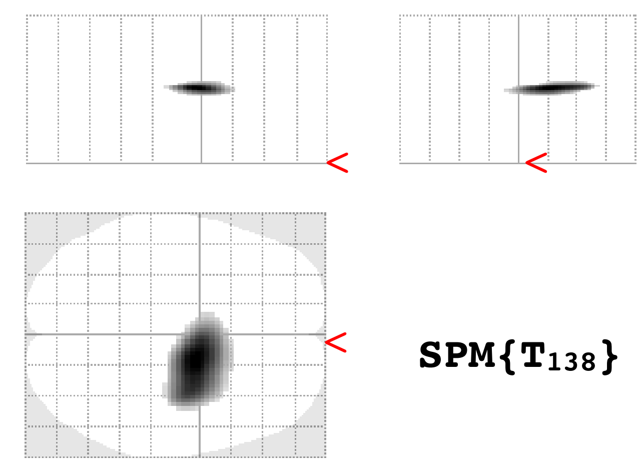
Linear combination of parameters

$$\hat{\beta}_1 > \hat{\beta}_2$$

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

$$c^\top \hat{\beta} > 0$$

SPM-*t* over time-space



t-contrast

$$t = \frac{c^\top \hat{\beta}}{\sqrt{\sigma^2 (c^\top (X^\top X)^{-1} c)}}$$

Classical inference

Constrast

$$c = [1 - 1]^\top$$

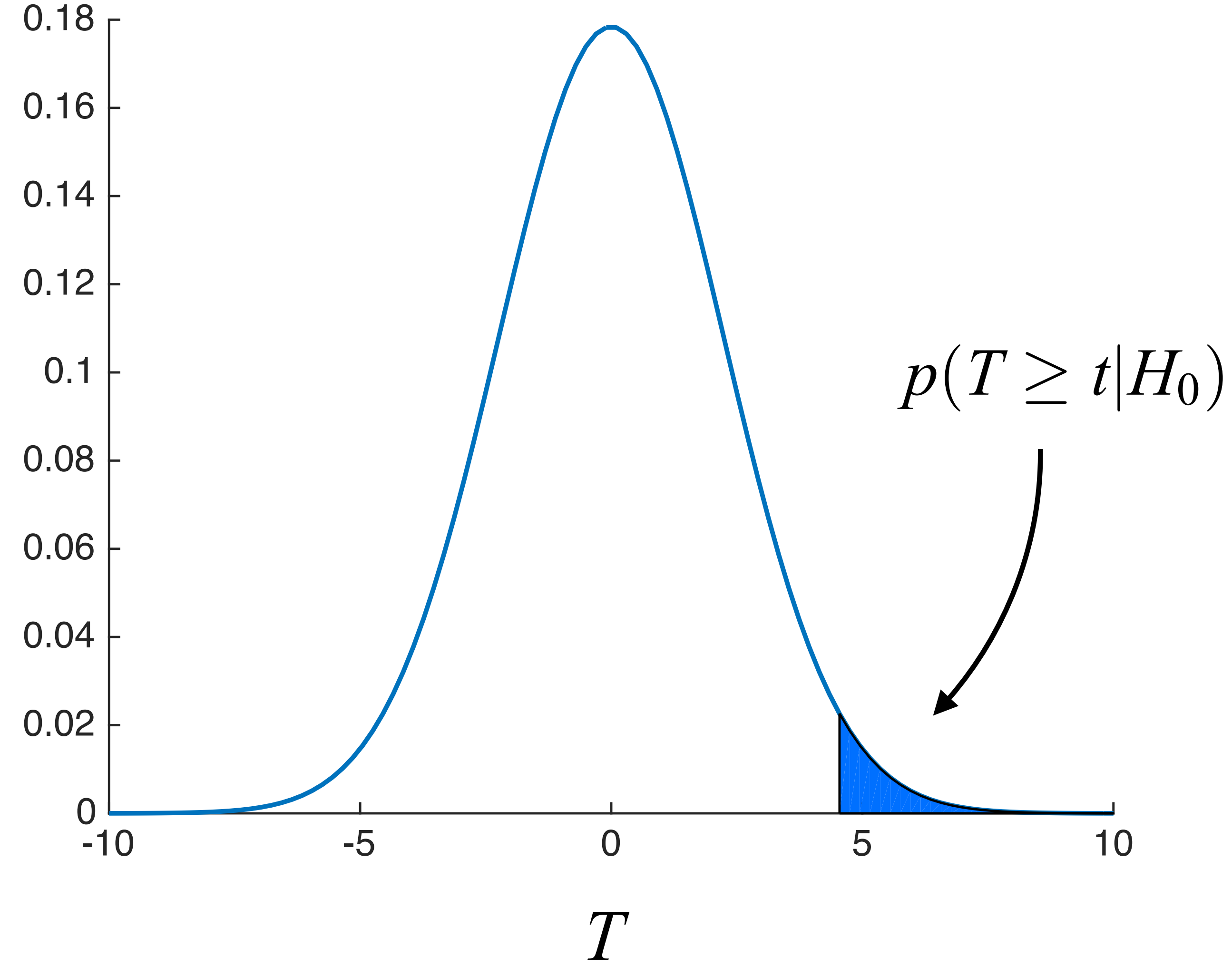
t statistic

$$t = \frac{c^\top \hat{\beta}}{\sqrt{\sigma^2 (c^\top (X^\top X)^{-1} c)}}$$

p value

$$p(T \geq t | H_0)$$

Analytic *T* null distribution



Classical inference

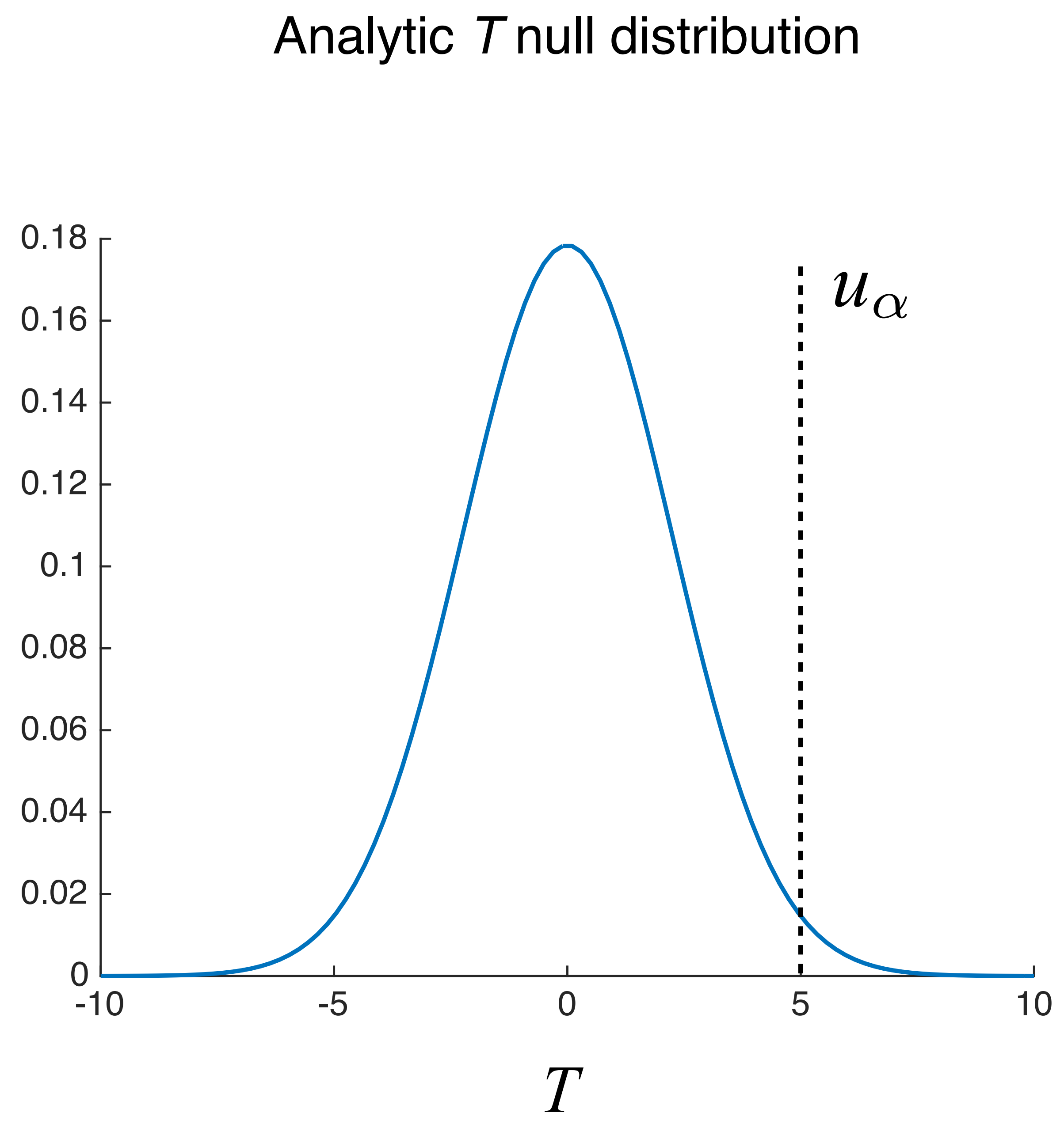
False-positive rate

$$\alpha = p(T \geq u_\alpha | H_0)$$

Hight threshold

$$u_\alpha = T \propto \alpha$$

$$\alpha = 0.05$$

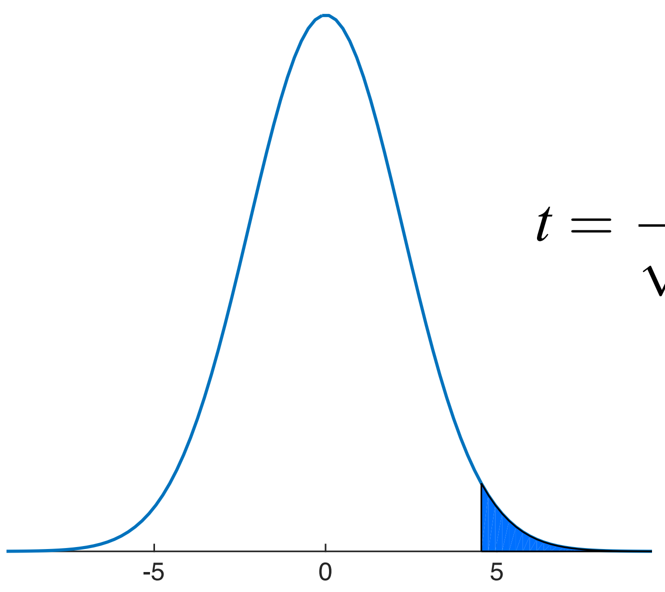


t-test summary

Contrast
Linear combination of parameters

***t*-test**
Signal-to-noise ratio measure

***t*-statistic**
No dependence on
scaling of regressors or contrast



$$t = \frac{\text{contrast}}{\sqrt{\text{variance}}}$$

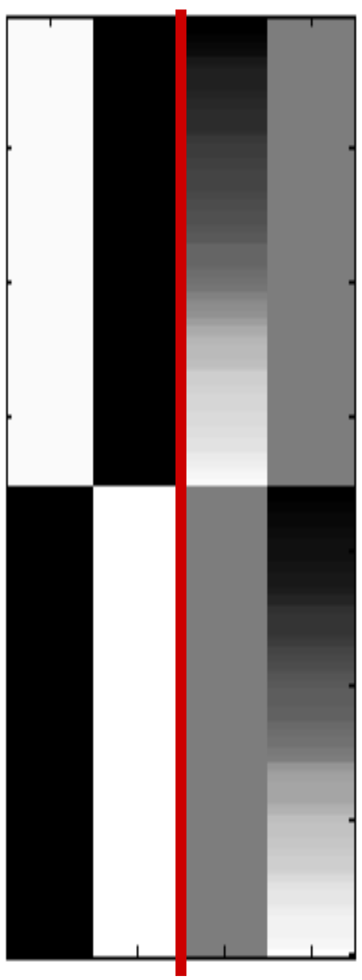
Unilateral test

H₀: $c^\top \hat{\beta} = 0$ **H_A:** $c^\top \hat{\beta} > 0$

F-test and extra sum-of-squares

Full model

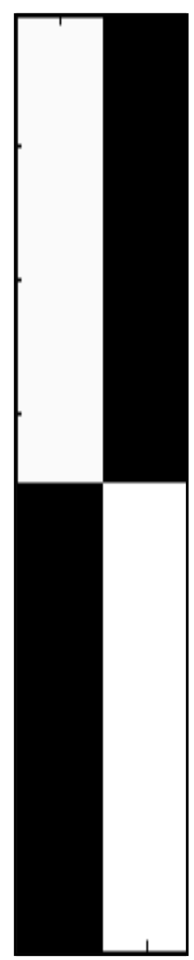
$X_0 \quad X_1$



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

Reduced model

X_0



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

***F*-statistic**

Ratio of explained and unexplained variance

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

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

$\nu_1 = \text{rank}(X) - \text{rank}(X_0)$

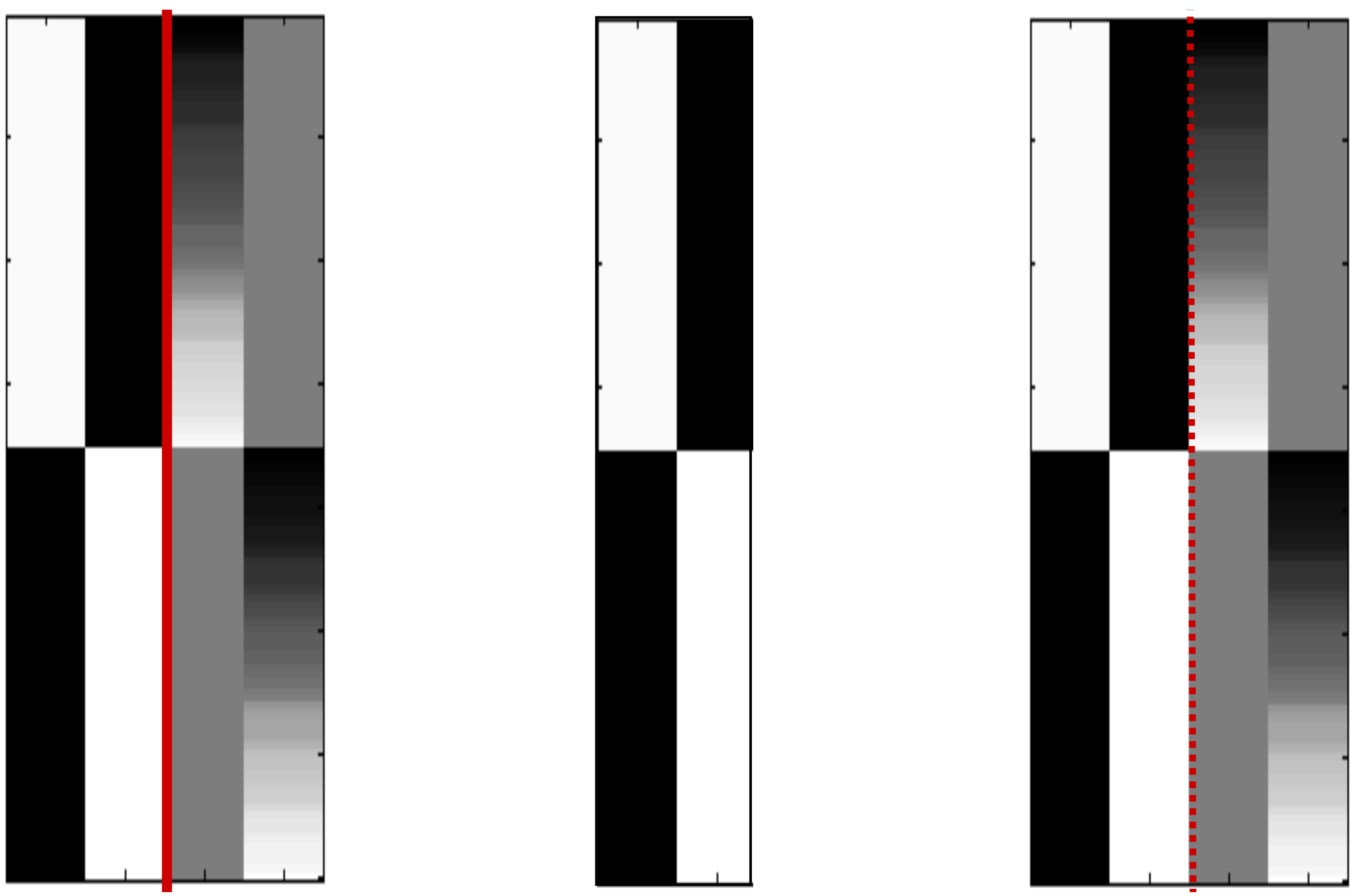
$\nu_2 = N - \text{rank}(X)$

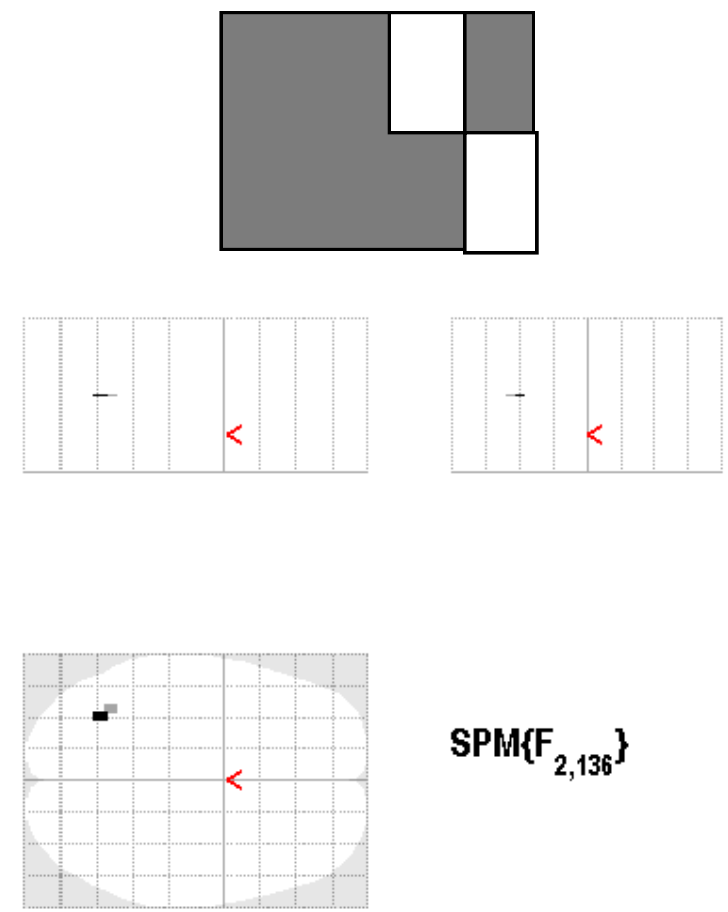
F-test and multi-dimensional contrasts

H₀: $\beta_3 = \beta_3 = 0$

H_A: $c^\top \hat{\beta} > 0$

$X_0 \quad X_1 \qquad\qquad X_0 \qquad c^\top = \begin{matrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}$





F-test summary

_F_-test

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

Additional variance explained by full model
w.r.t. a reduced (nested) model

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

Uni-dimensional contrast

Testing $b_1 - b_2$ is the same as testing $b_2 - b_1$
It is exactly the square of the *t*-test, testing for both positive and negative effects

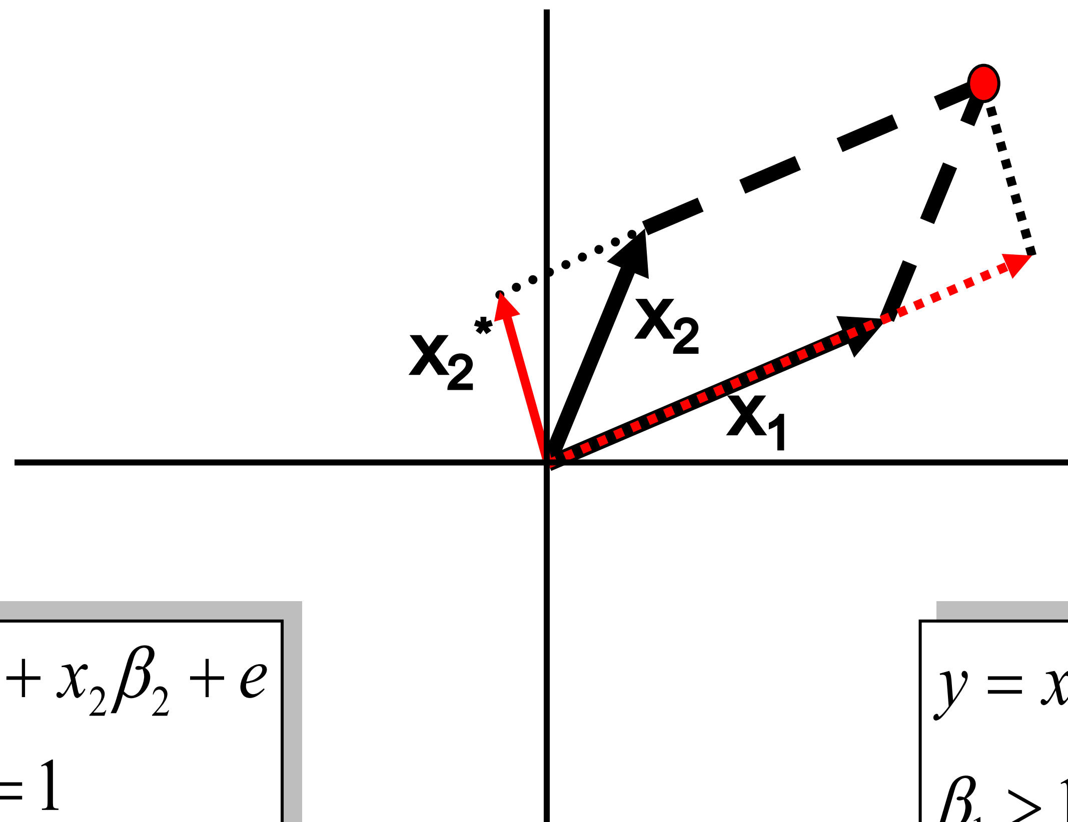
Multi-dimensional contrast

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

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$

$$H_A : \text{at least one } \beta_k \neq 0$$

Correlated and orthogonal regressors



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

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

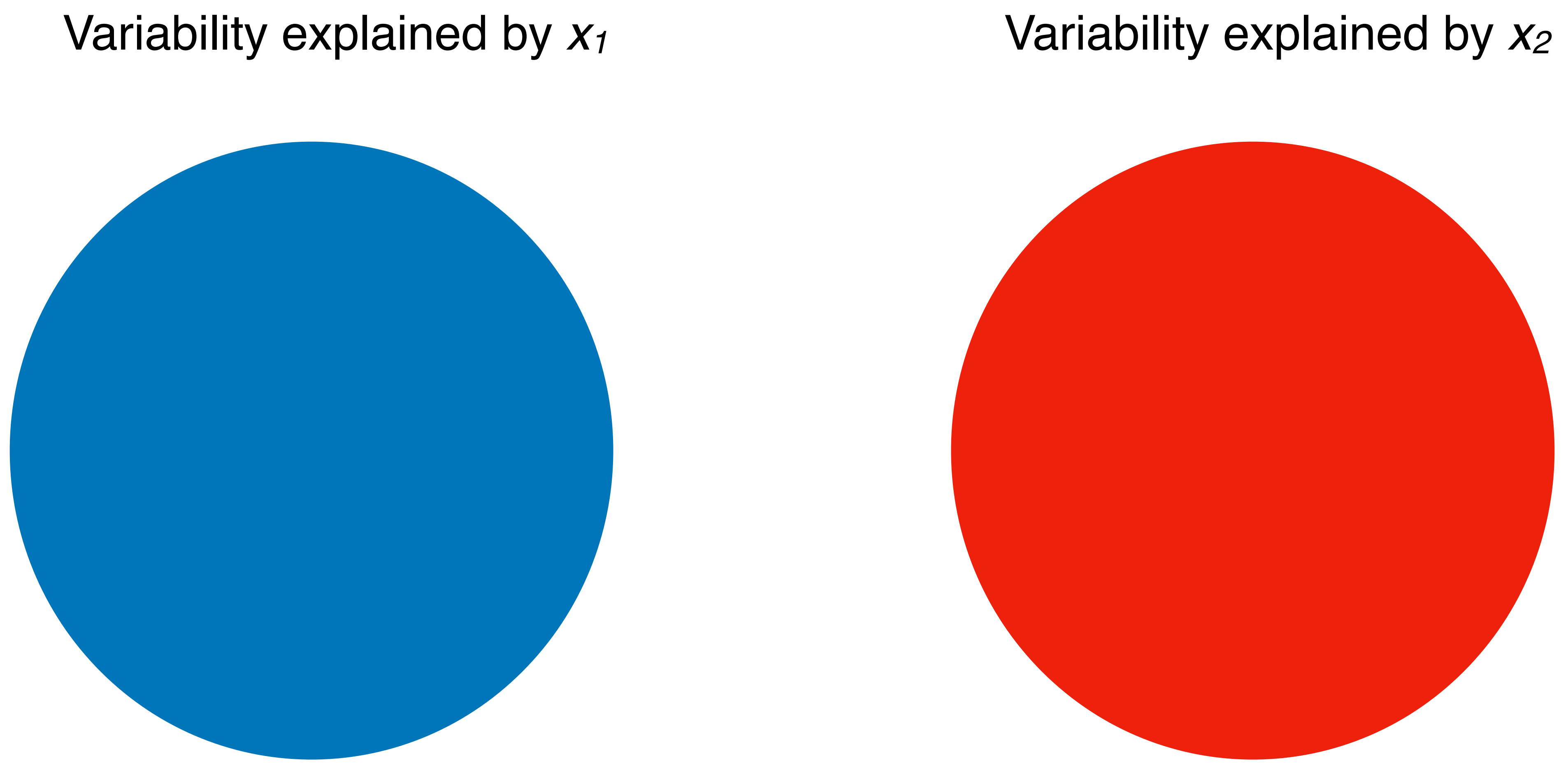
Correlated regressors

Explained variance shared between regressors

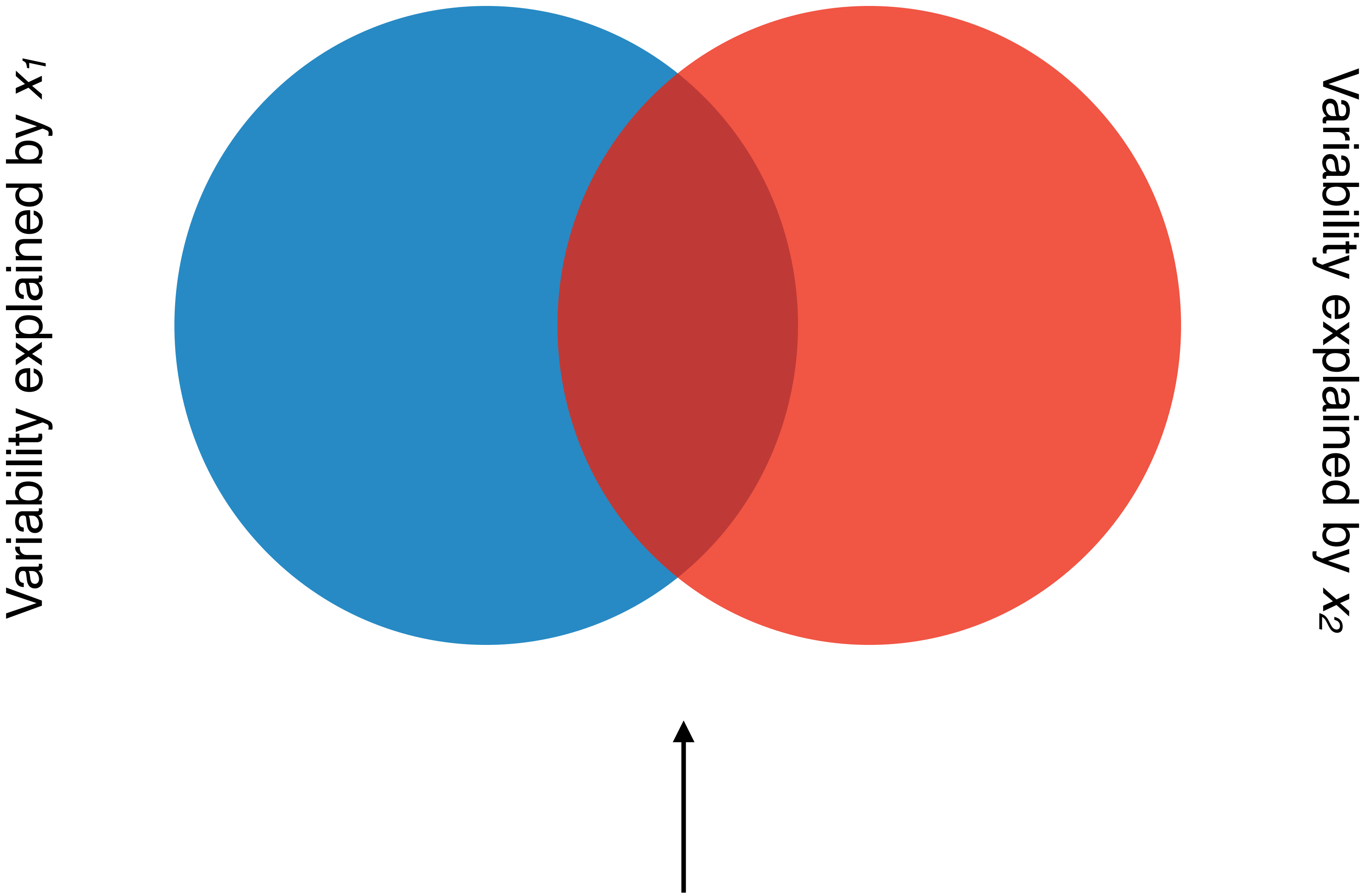
Orthogonalised regressors

Parameter estimate for x_1 changes, not for x_2

Orthogonal regressors

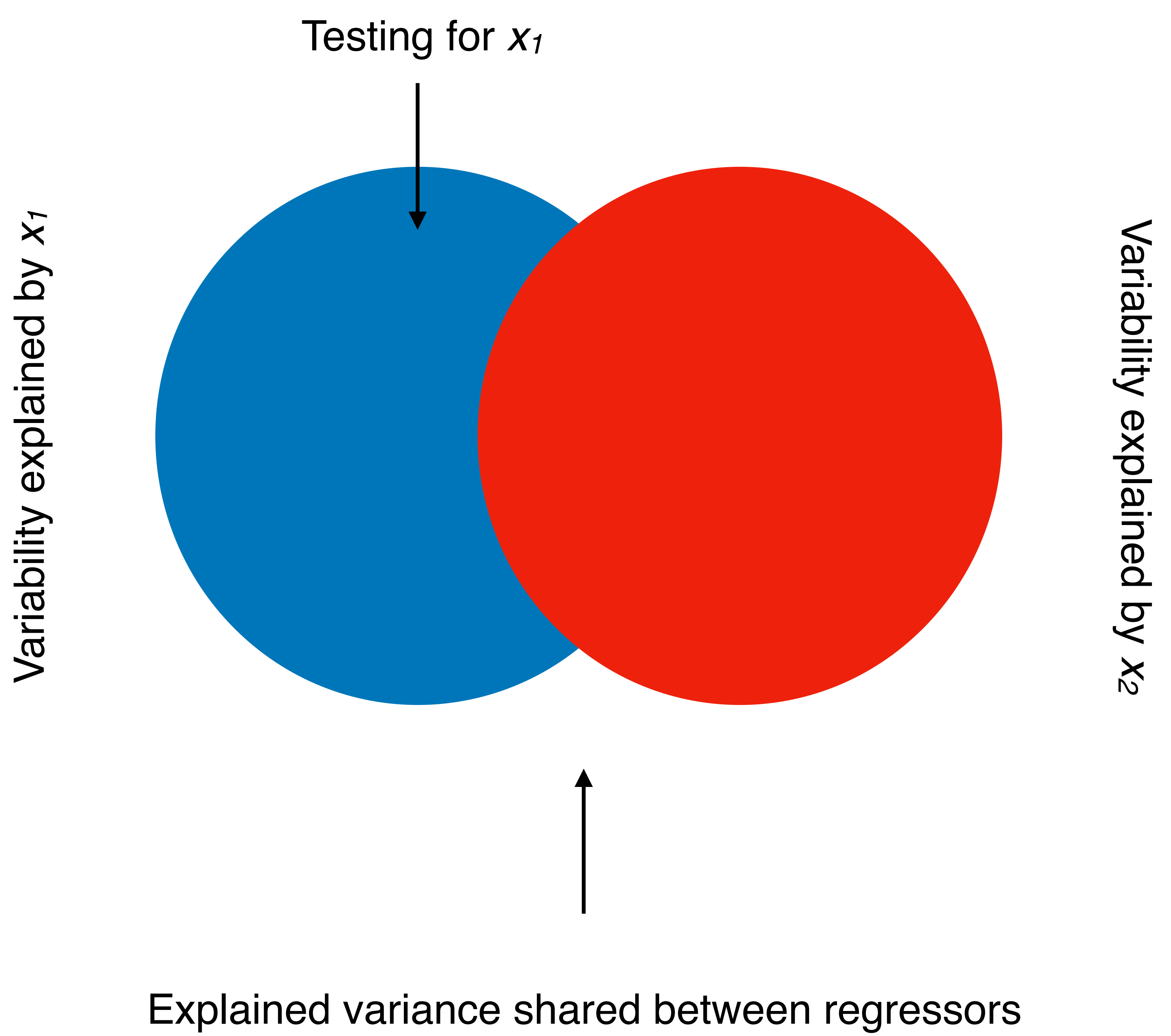


Correlated regressors

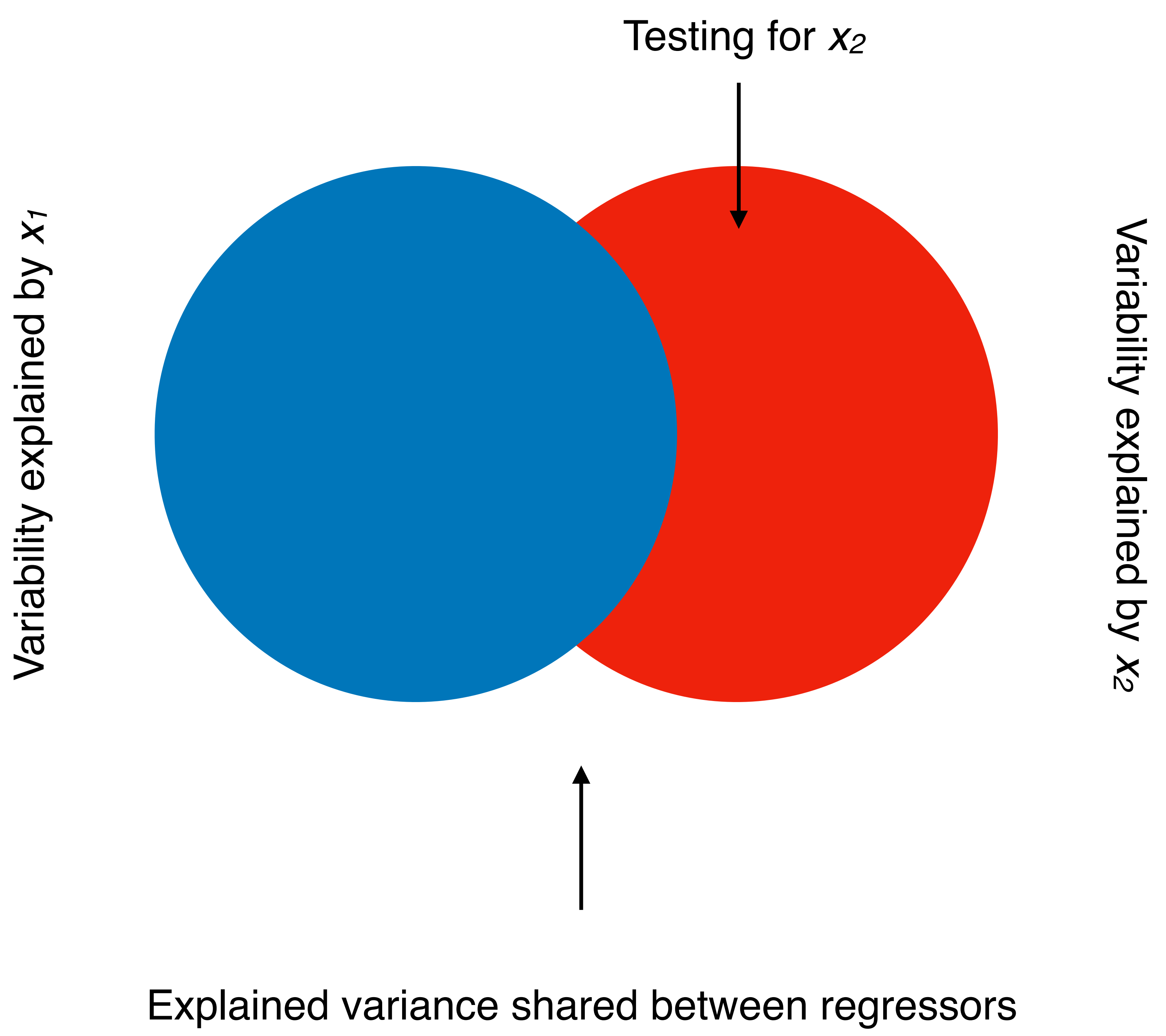


Explained variance shared between regressors

Correlated regressors



Correlated regressors

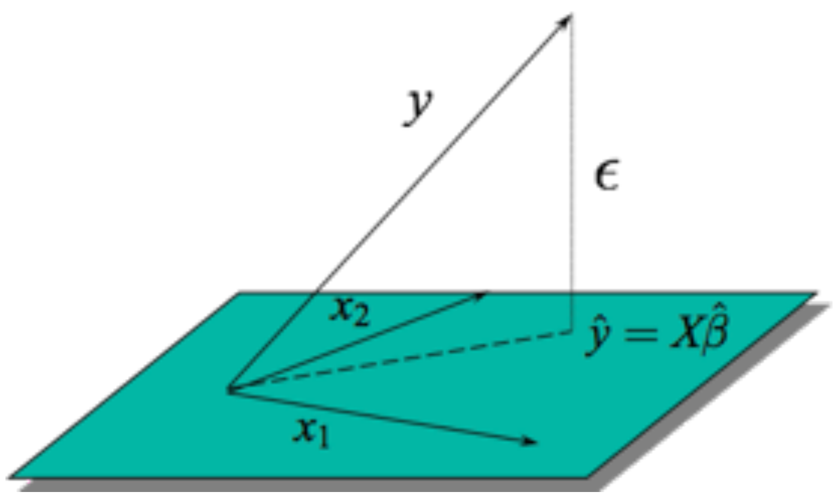


Summary

Special cases of the GLM

- *t*-test
- *F*-test
- Analysis of variance (ANOVA)
- Analysis of covariance (AnCova)
- multiple regression

$$Y = X\beta + \epsilon$$



Thank you!

Any questions ?

Thanks to Christophe, Guillaume, Will, Rik, Stefan and Karl!