wellcome centre human neuroimaging

AARHUS UNIVERSITY

# General linear model and classical inference

*SPM for M/EEG course*

Martin Dietz
CFIN, Aarhus University, DK

# Overview

**Introduction**

SPM for MEG/EEG data

ERP example

**General linear model (GLM)**

Definition and design matrix

Parameter estimation

**Classical inference**

Contrasts and inference (*t*-tests and *F*-tests)

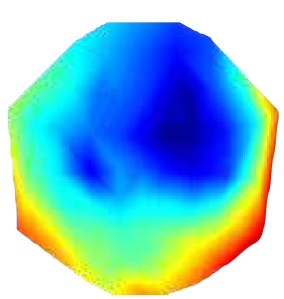
Correlated regressors and orthogonalisation

# Overview of SPM for M/EEG

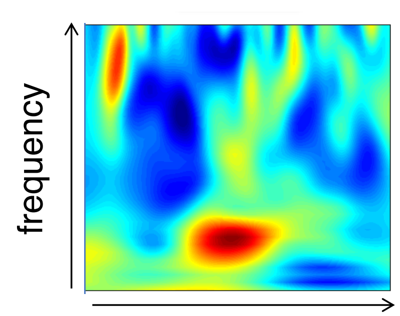
**Preprocessing**

Filtering
Epoching
Re-referencing
Artefact rejection
Averaging
…

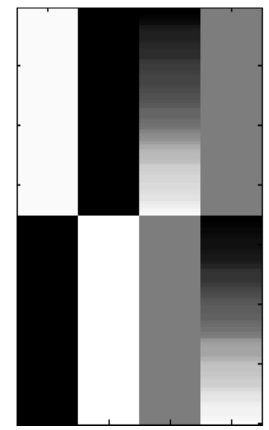
Image conversion



Scalp-time



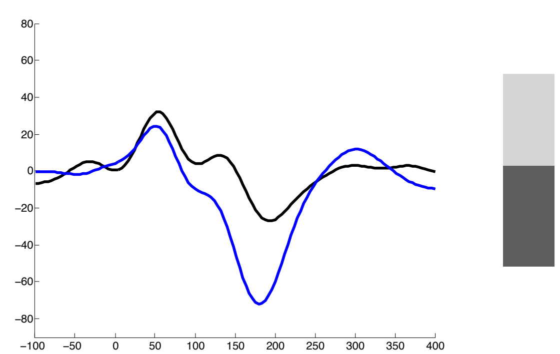

Design matrix $X$



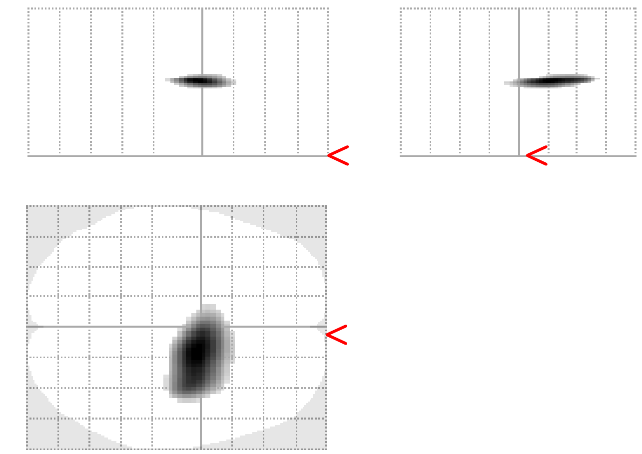
General linear model

$$y = X\beta + \epsilon$$

Parameter estimates $\beta$



Statistical parametric map (SPM)



Inference and correction for multiple comparisons

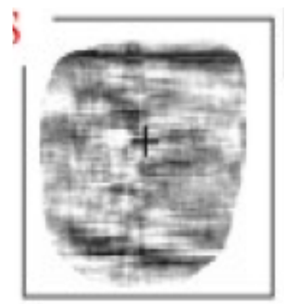
Contrast $c^\top \beta$

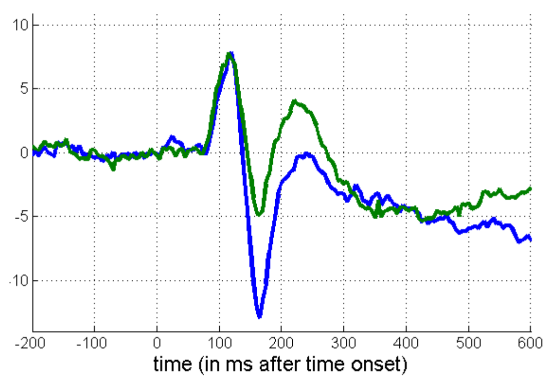
$$c = [1 \text{ -1 } 0 \ 0]^\top$$

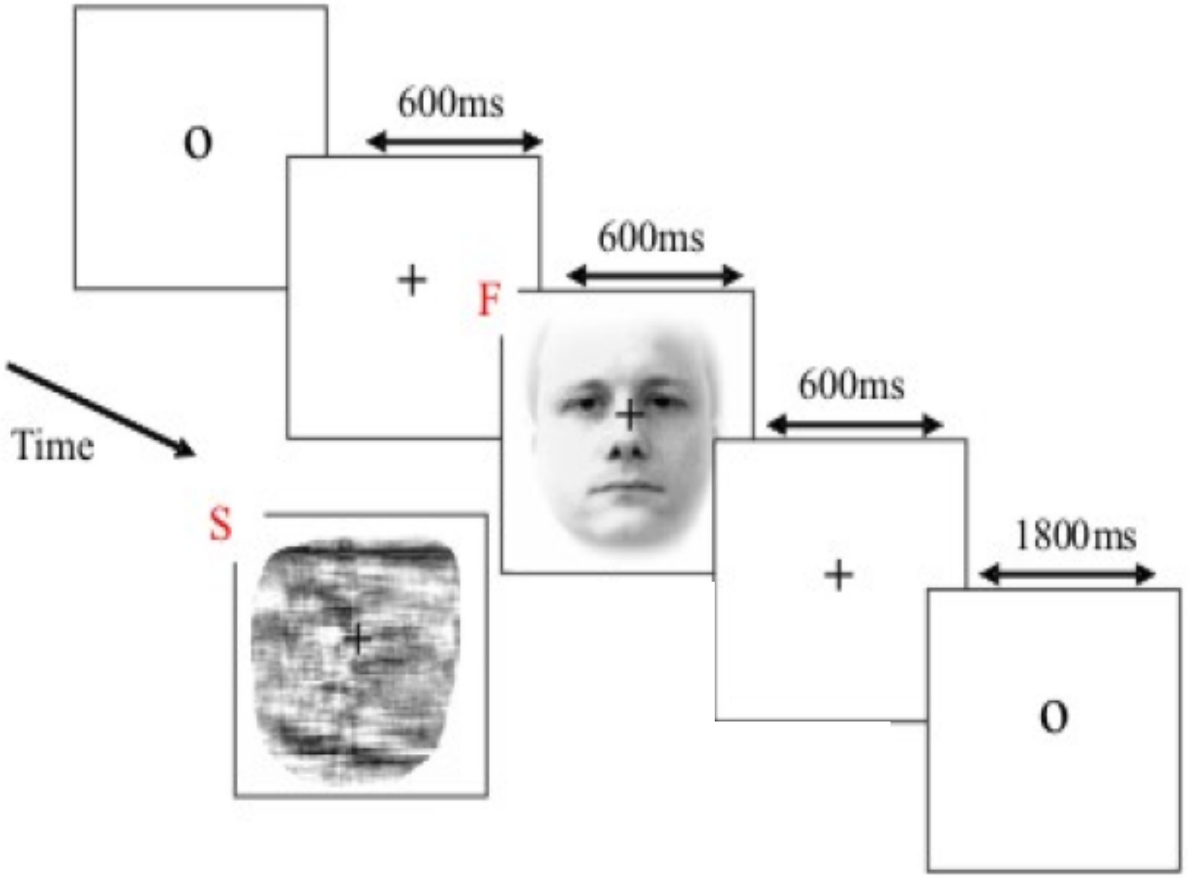
# Event-related potential (ERP) example

Visual stimuli

**Faces** **Scrambled**

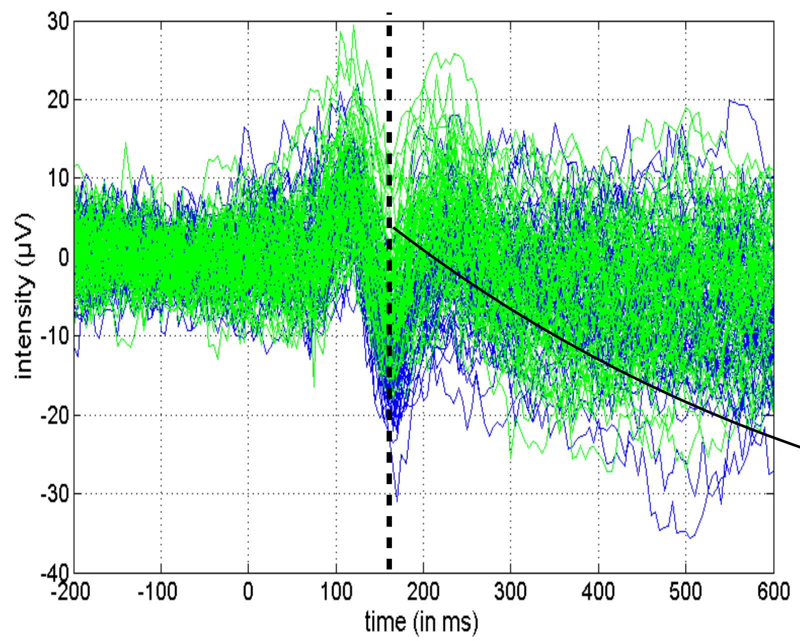


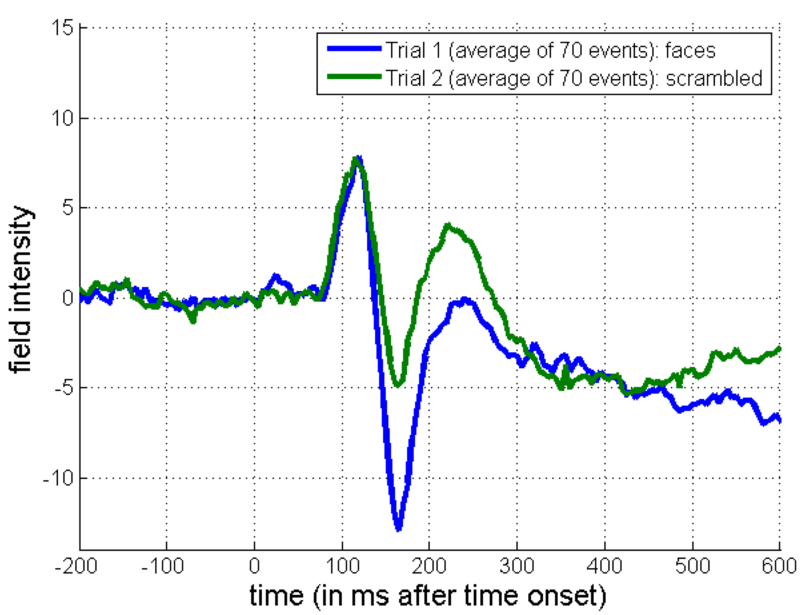

*Hypothesis*: difference in ERP to ***faces*** and ***scrambled faces*** ?

# ERP example: one channel and time point

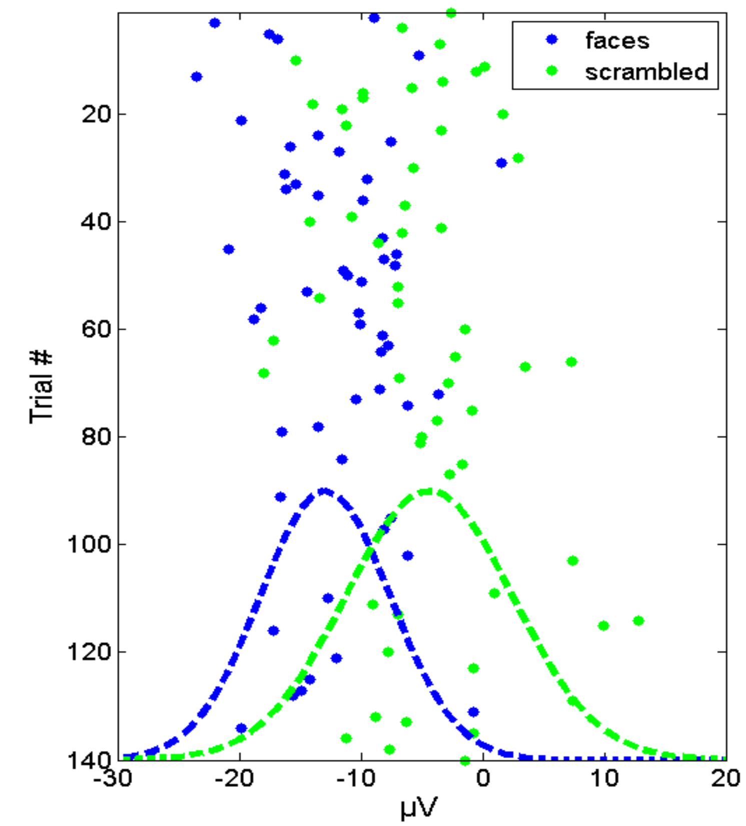
## Trial-by-trial variability




## Average




## Compare size of effect (average) to its standard error (variability)

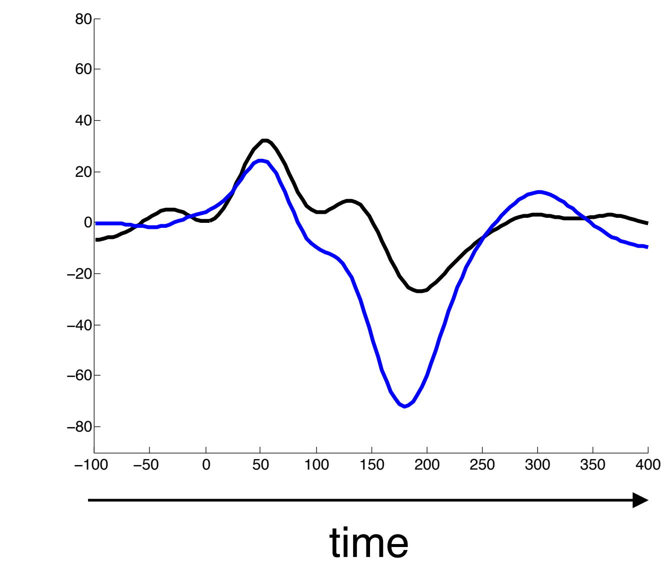



## *t* statistic (signal-to-noise ratio)

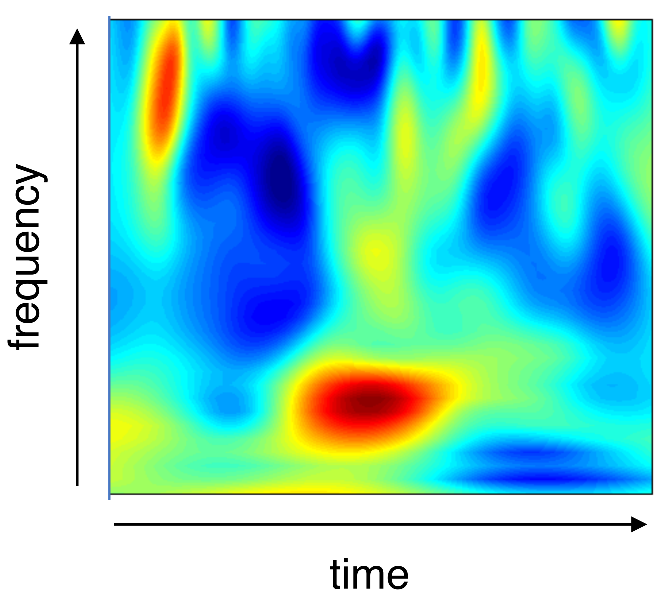
$$t = \frac{\mu_f - \mu_s}{\sqrt{\hat{\sigma}^2\left(1/n_f + 1/n_s\right)}}$$

# Images of MEG/EEG data

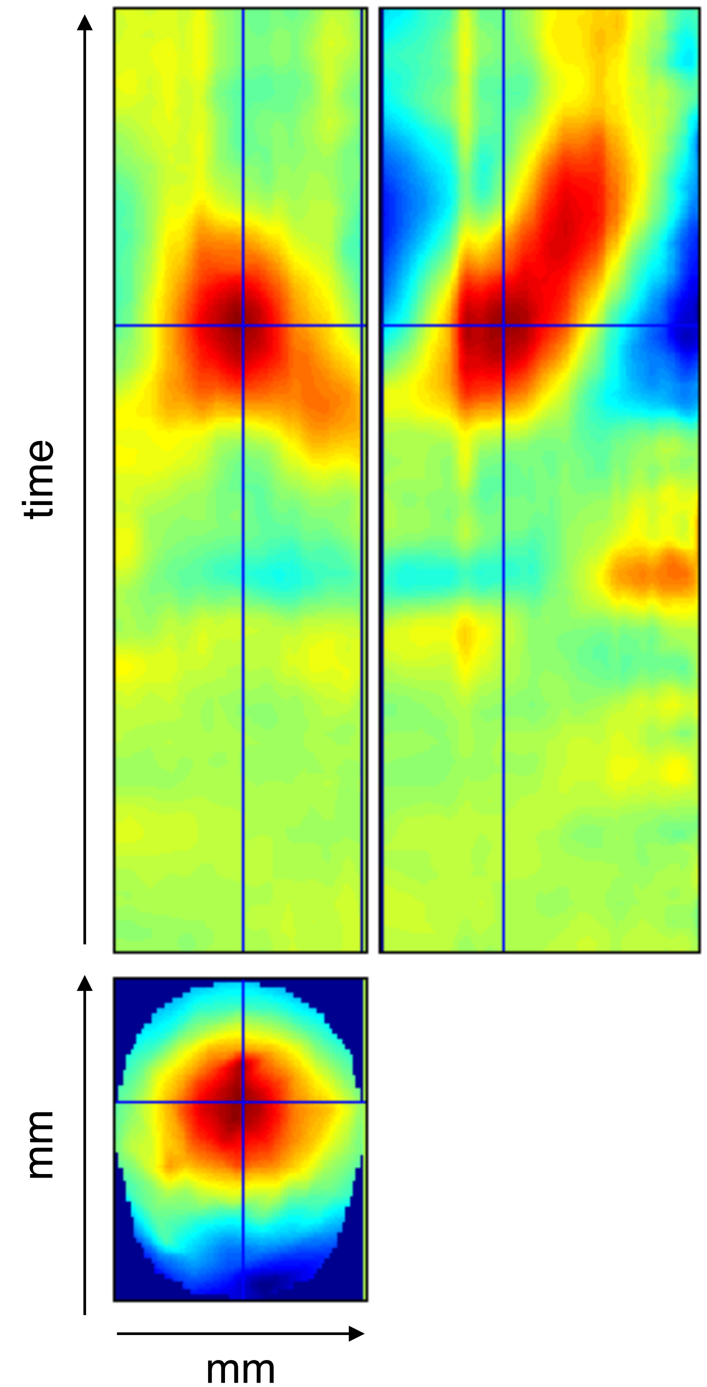
## 1D time



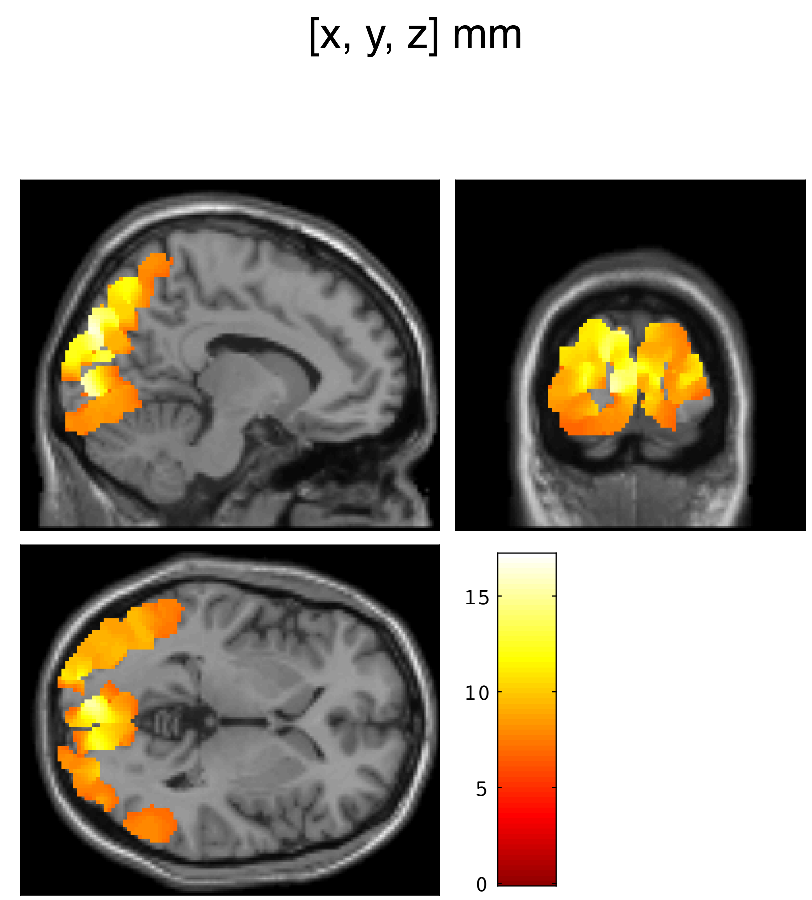

## 2D time-frequency



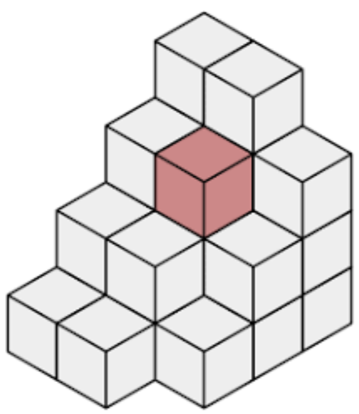

## 3D sensor-time




## 3D source images

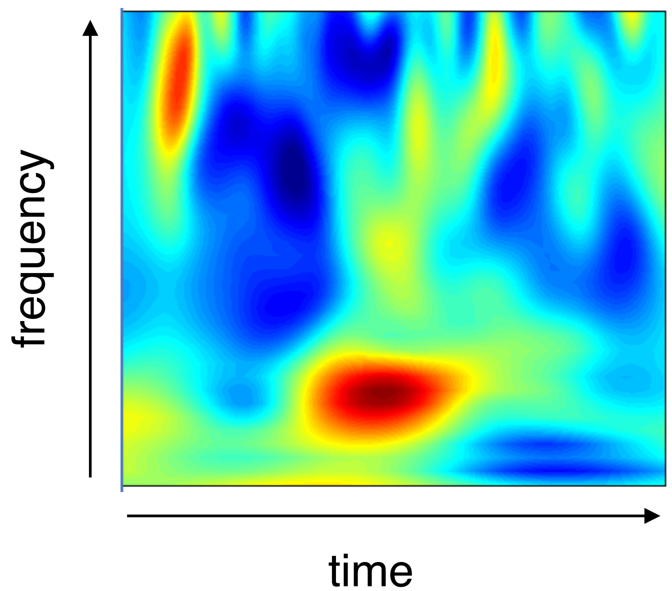


# Mass-univariate statistical framework

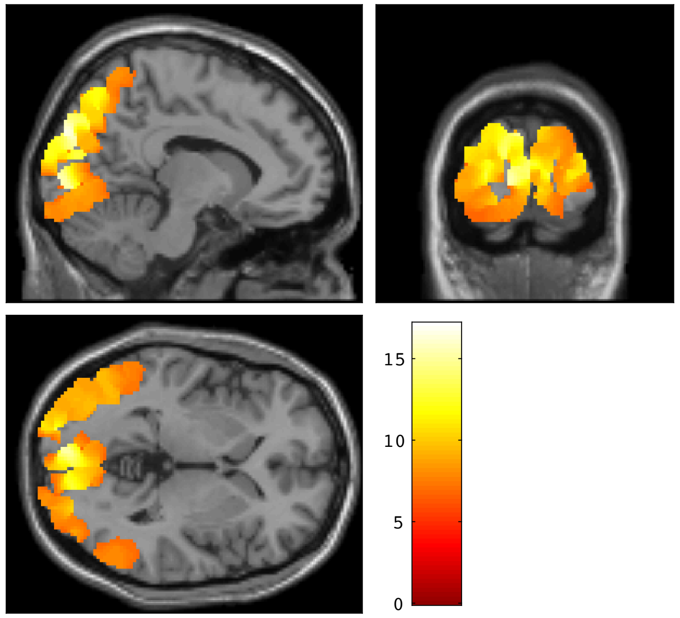


2D pixel or 3D voxel

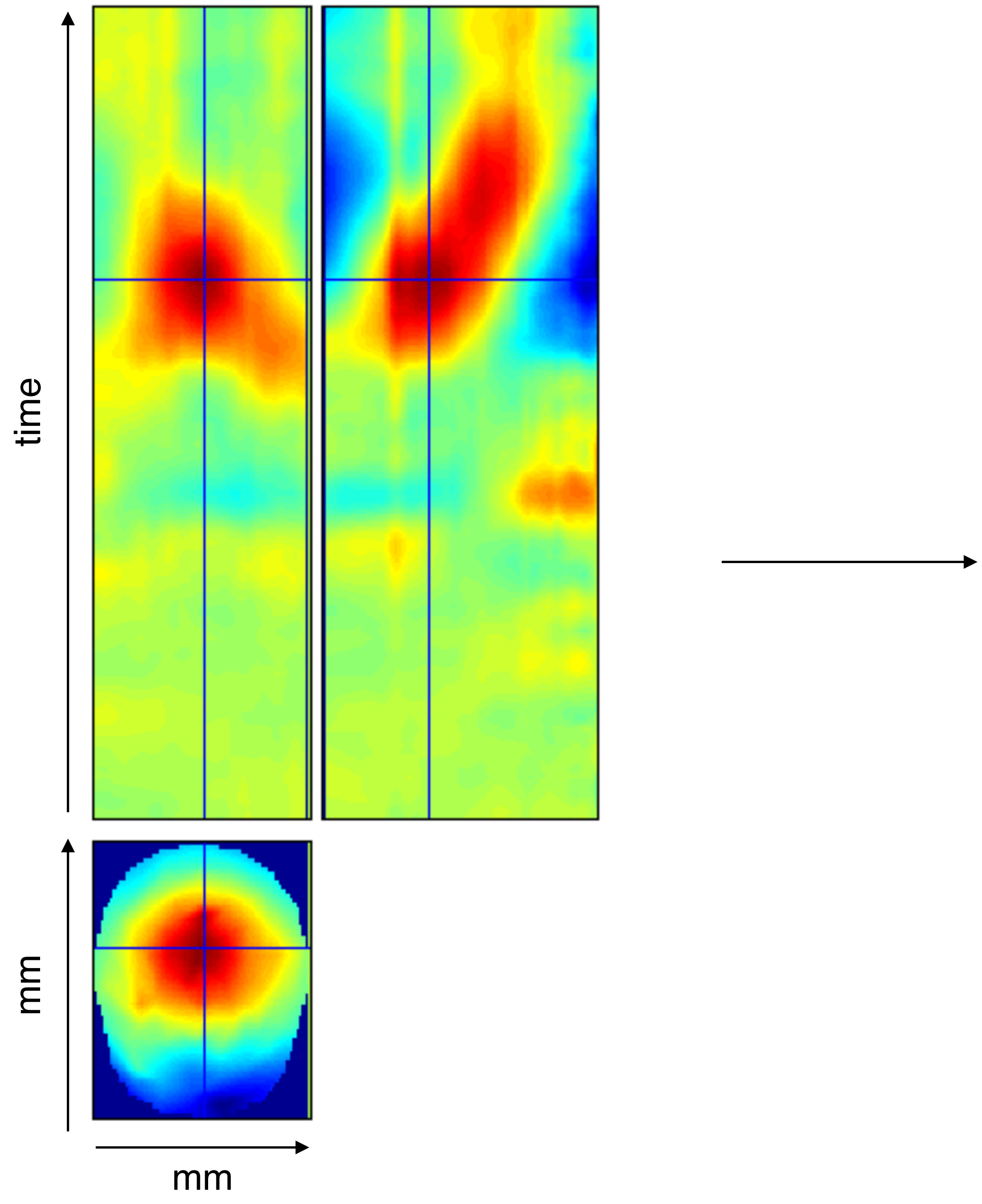
2D time-frequency



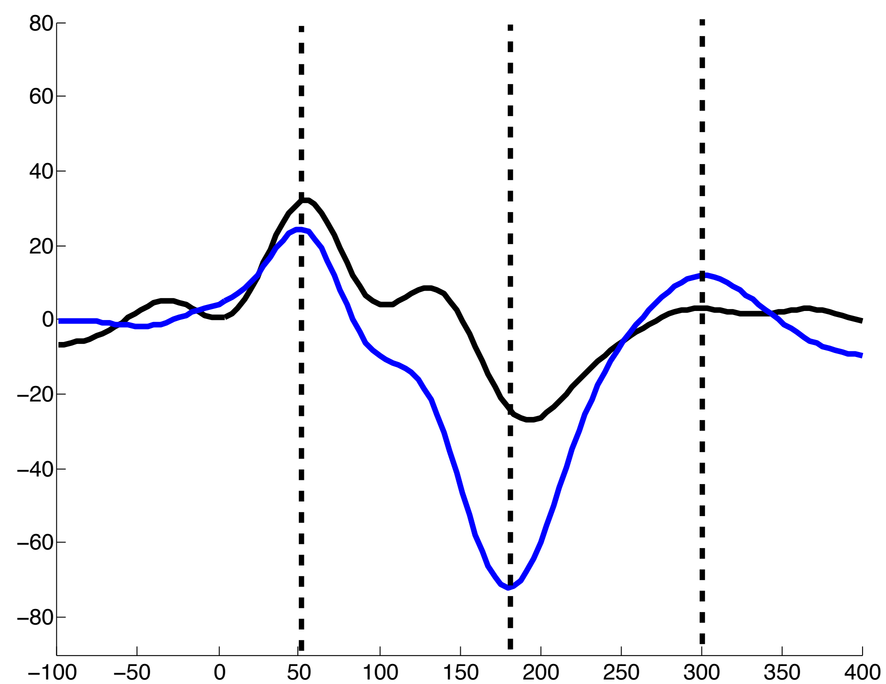

3D source images



# Mass-univariate statistical framework




*Avoid* selection bias ("cherry picking")



# Overview

**Introduction**

SPM for MEG/EEG data

ERP example

**General linear model (GLM)**

Definition and design matrix

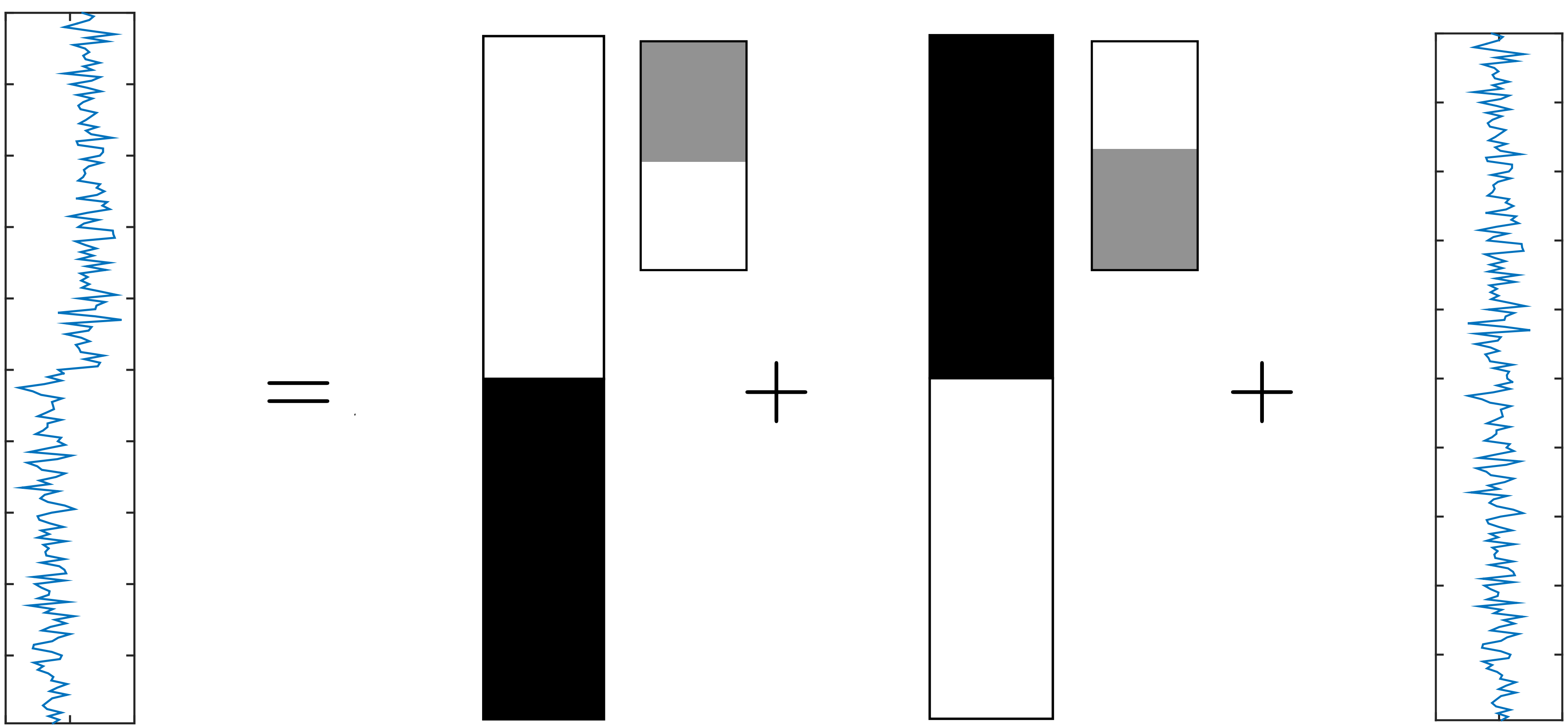
Parameter estimation

**Classical inference**

Contrasts and inference (*t*-tests and *F*-tests)

Correlated regressors and orthogonalisation

# Linear model




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

# General linear model

Linear regression

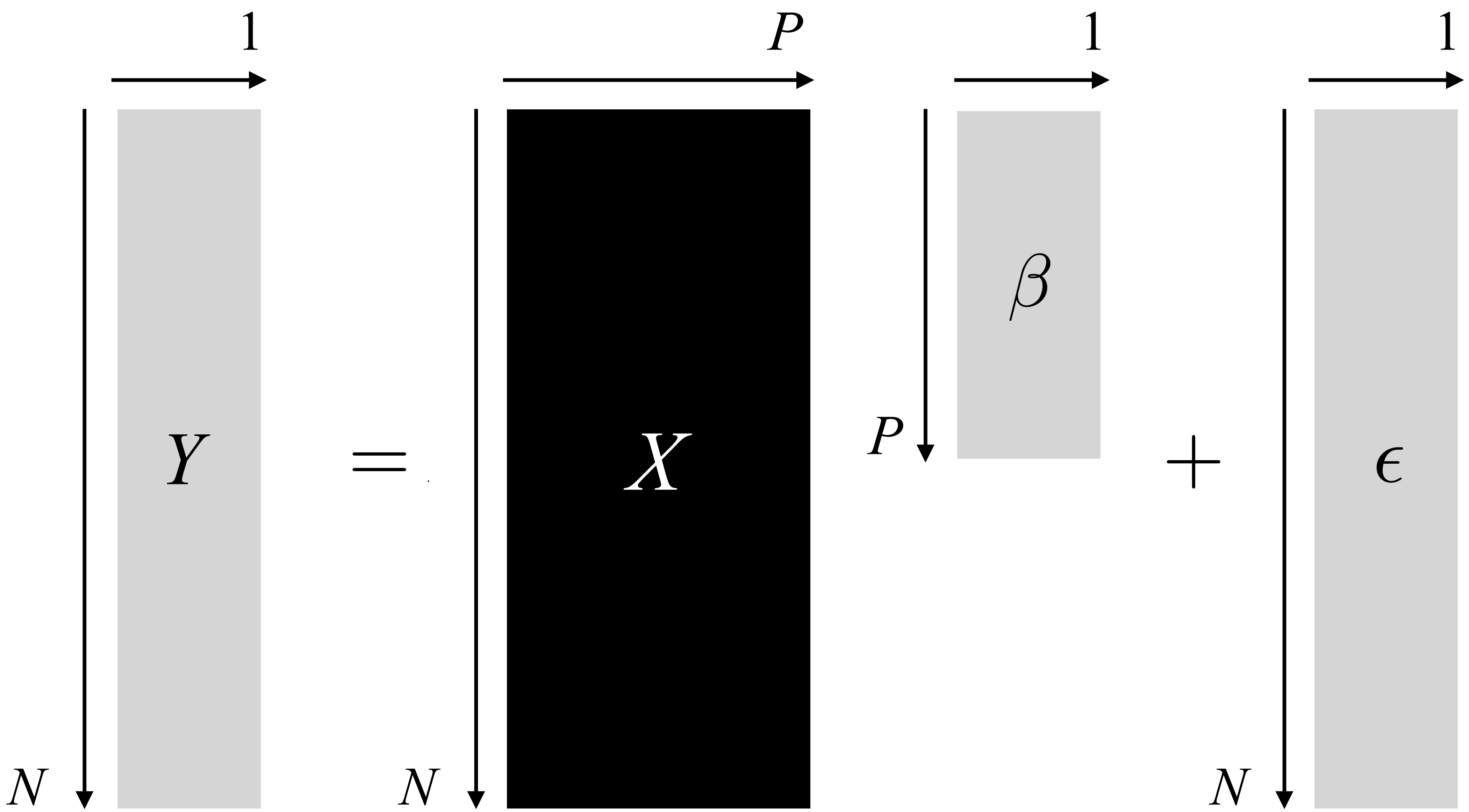
$$y = x_0\beta_0 + x_1\beta_1 + .... + x_p\beta_p + \epsilon$$

Matrix formulation

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

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

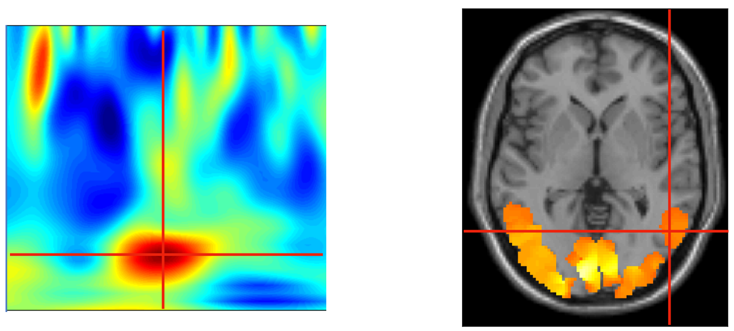
# General linear model

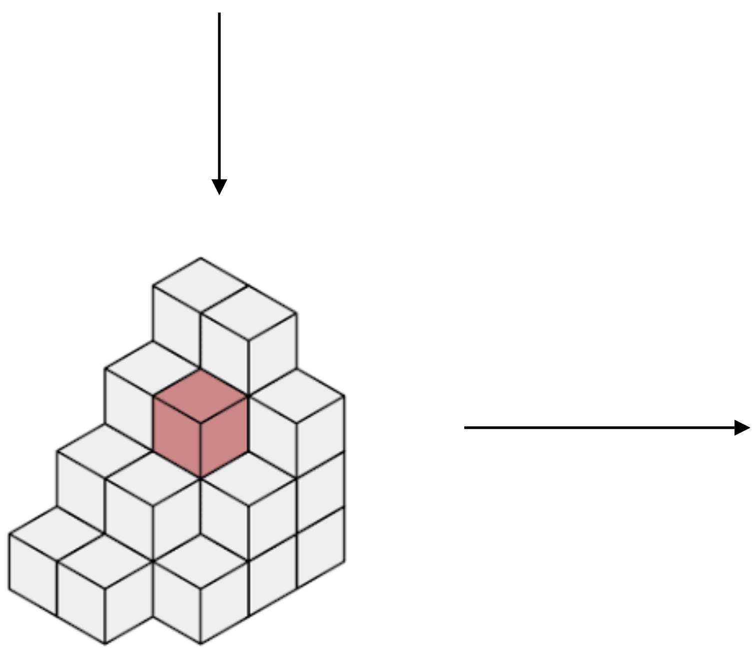



$N$ x $1$ data points, $N$ x $P$ predictors, $P$ x $1$ parameters, and $N$ x $1$ residuals

# Voxel-wise linear model

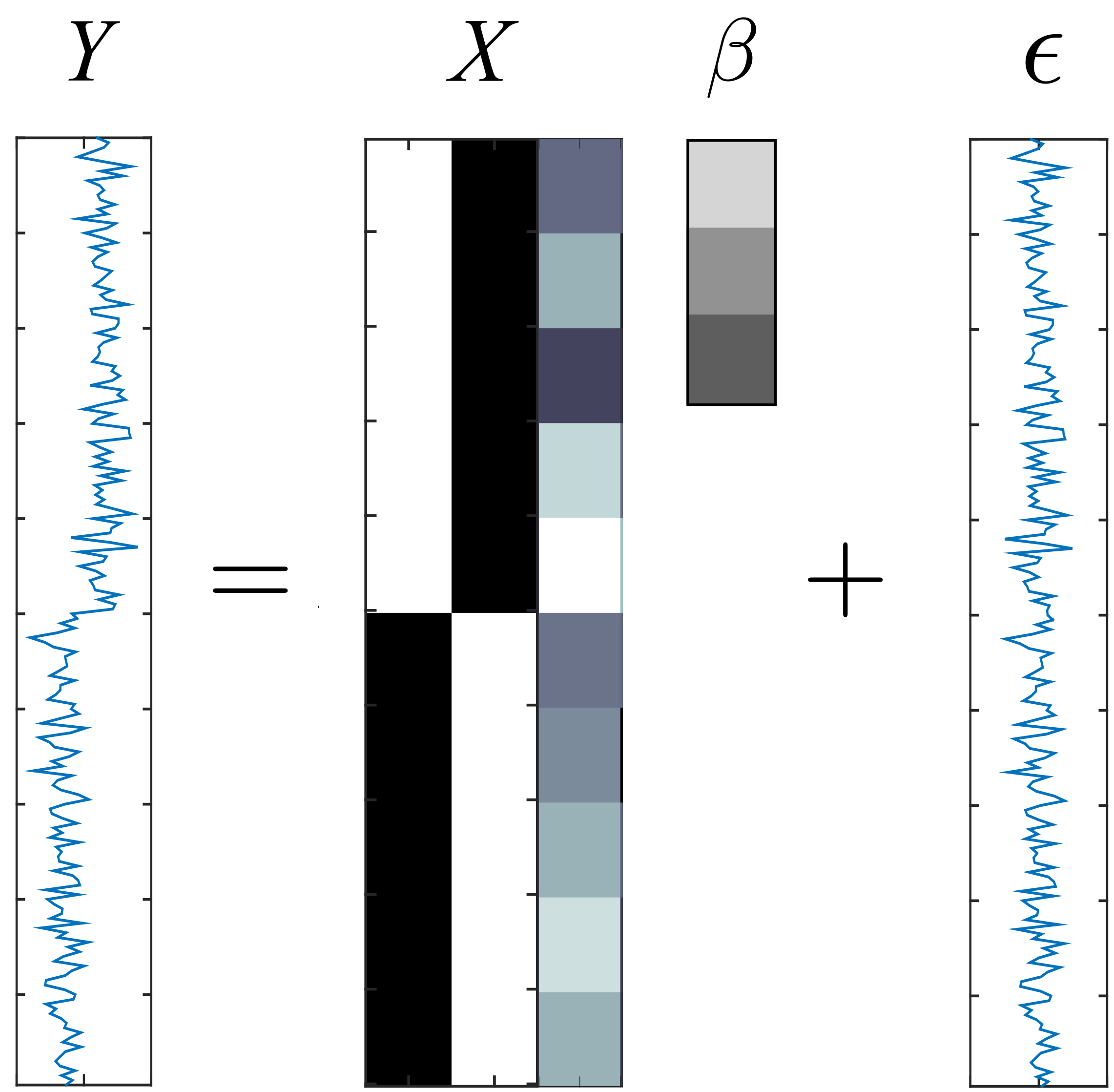
2D or 3D image





2D pixel or 3D voxel

Design matrix encodes
*experimental factors* and *confounds*



# GLM assumptions

Residual errors are independent and identically distributed (i.i.d.)

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

If $X$ has full rank, the moment matrix $(X^\top X)$ is invertible

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

If $X$ is overdetermined with $N > P$, $X$ is rank deficient

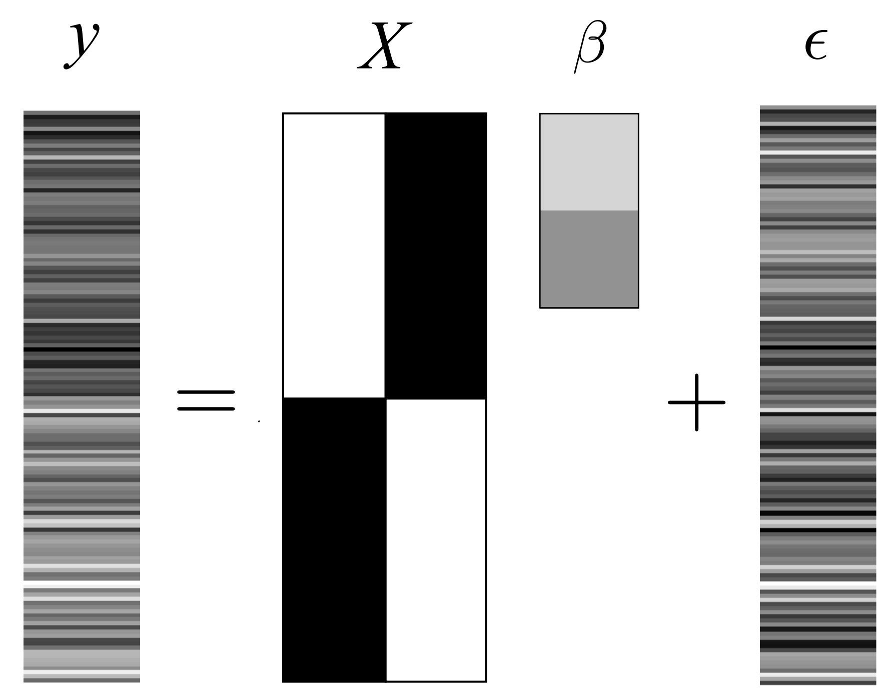
$(X^\top X)$ is only invertible using Moore-Penrose pseudo-inverse

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

# Parameter estimation

Iff residual error is i.i.d. and Gaussian

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




Ordinary least squares (OLS)

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

minimise $S = \sum_i \hat{\epsilon}_i^2$ w.r.t $\hat{\beta}$ (objective function)

$$\frac{\partial S}{\partial \hat{\beta}_p} = 2 \sum_{i=1}^{N} (-x_{(i,p)})(Y_i - x_{(i,1)}\hat{\beta}_1 - \ldots - x_{(i,P)}\hat{\beta}_P) = 0$$

# Parameter estimation

Parameter estimates

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

Predicted responses

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

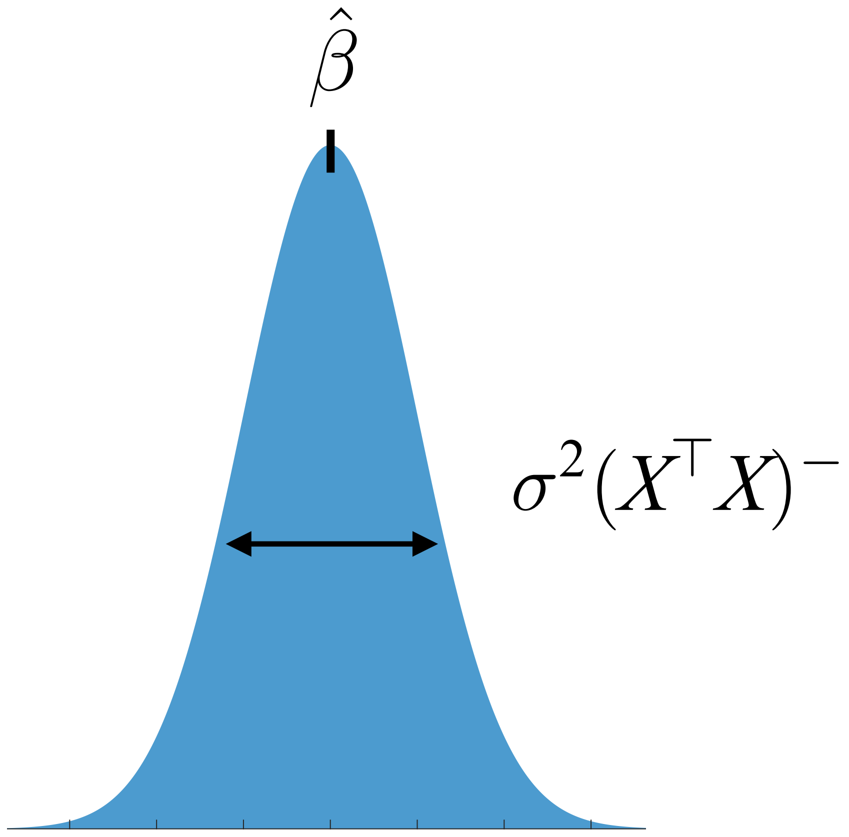
Variance (pooled)

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

$$\hat{\sigma}^2 = \frac{\epsilon^\top \epsilon}{N - \mathrm{rank}(X)}$$

Normally distributed

$$\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X^\top X)^-)$$



# Parameter estimation: *a geometric perspective*

Ordinary least squares

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

Residual forming matrix

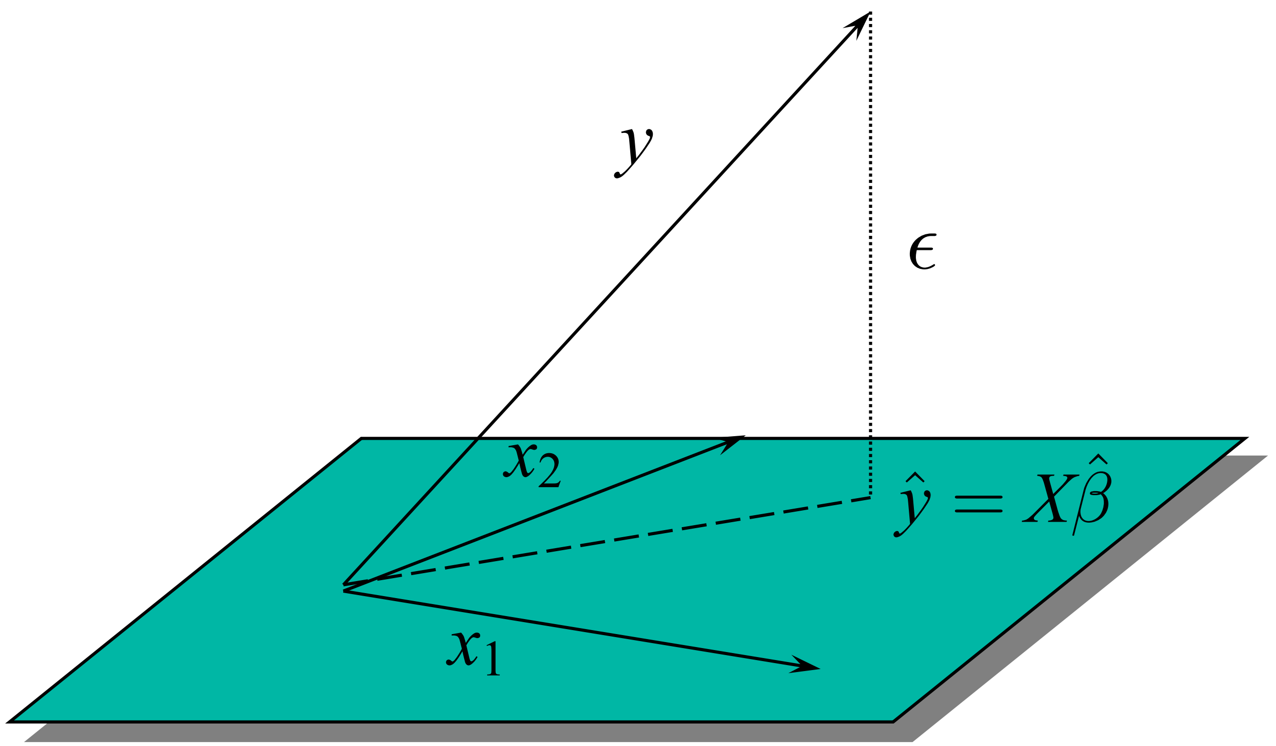
$$\epsilon = Ry$$

$$R = I - P$$

Projection matrix

$$\hat{y} = Py$$

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




Design space of *X*

# Overview

**Introduction**

SPM for MEG/EEG data

ERP example

**General linear model (GLM)**

Definition and design matrix

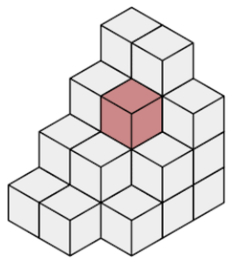
Parameter estimation

**Classical inference**

Contrasts and inference (*t*-tests and *F*-tests)

Correlated regressors and orthogonalisation

# *t*-contrast



voxel-wise

Constrast vector

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

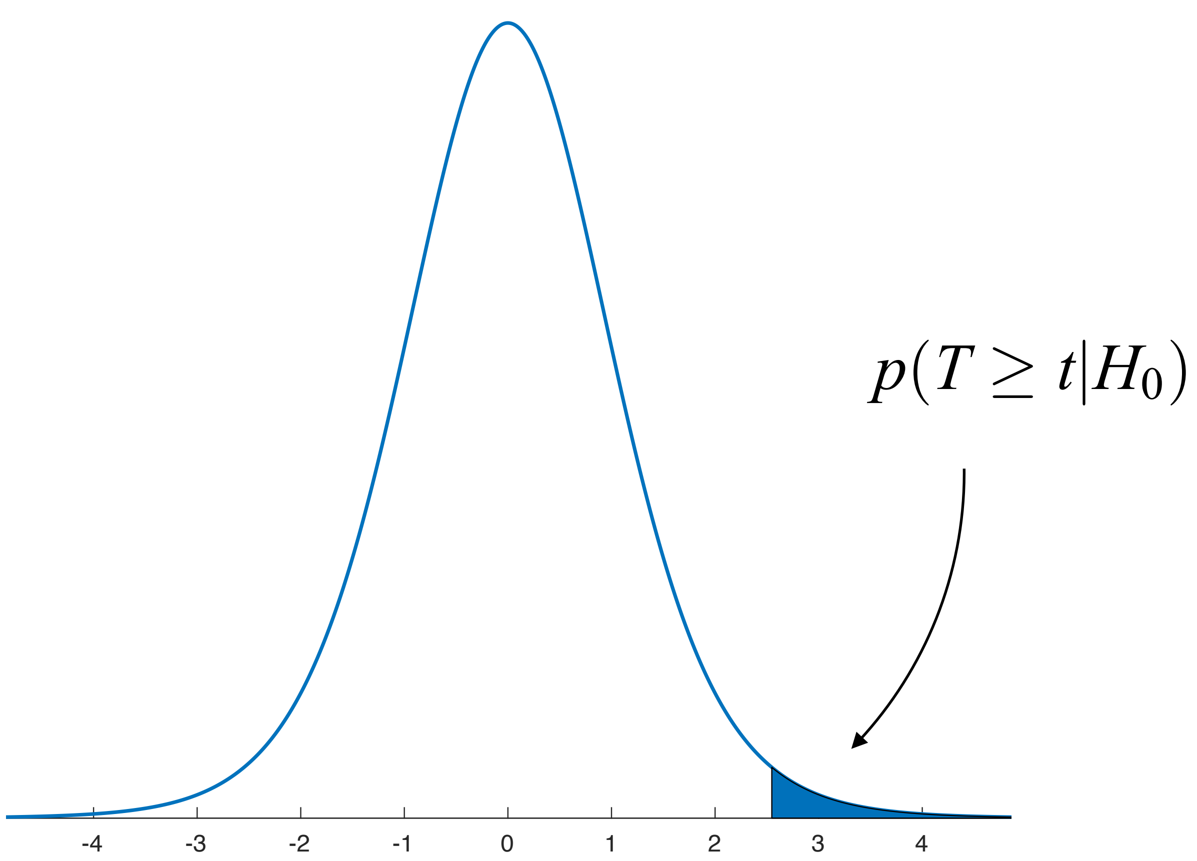
*T* statistic

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

*P* value (one-sided)

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

*T* null distribution



# $t$-statistic

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

Size of effect

Variance of effect

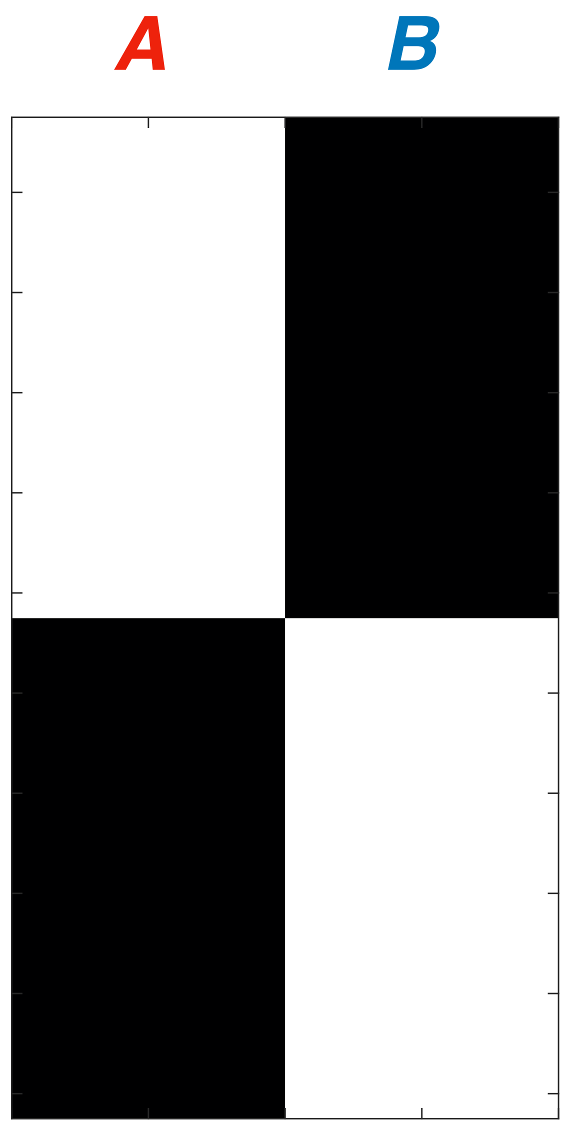
Design efficiency $^{-1}$

Contrast estimate is normally distributed

$$c^\top \hat{\beta} \sim \mathcal{N}(c^\top \beta, \sigma^2 (c^\top (X^\top X)^- c))$$

# Two-sample *t*-test

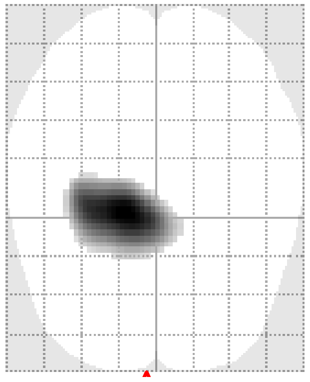
Between-groups design




Constrast vector

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

SPM-$t$



One-sided hypothesis test

$$c^\top \hat{\beta} \leq 0 \quad \text{(null)}$$

$$c^\top \hat{\beta} > 0 \quad \text{(alternative)}$$

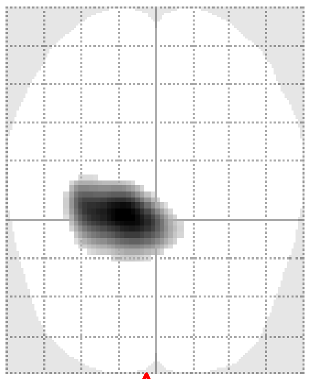
Linear combination of parameters

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

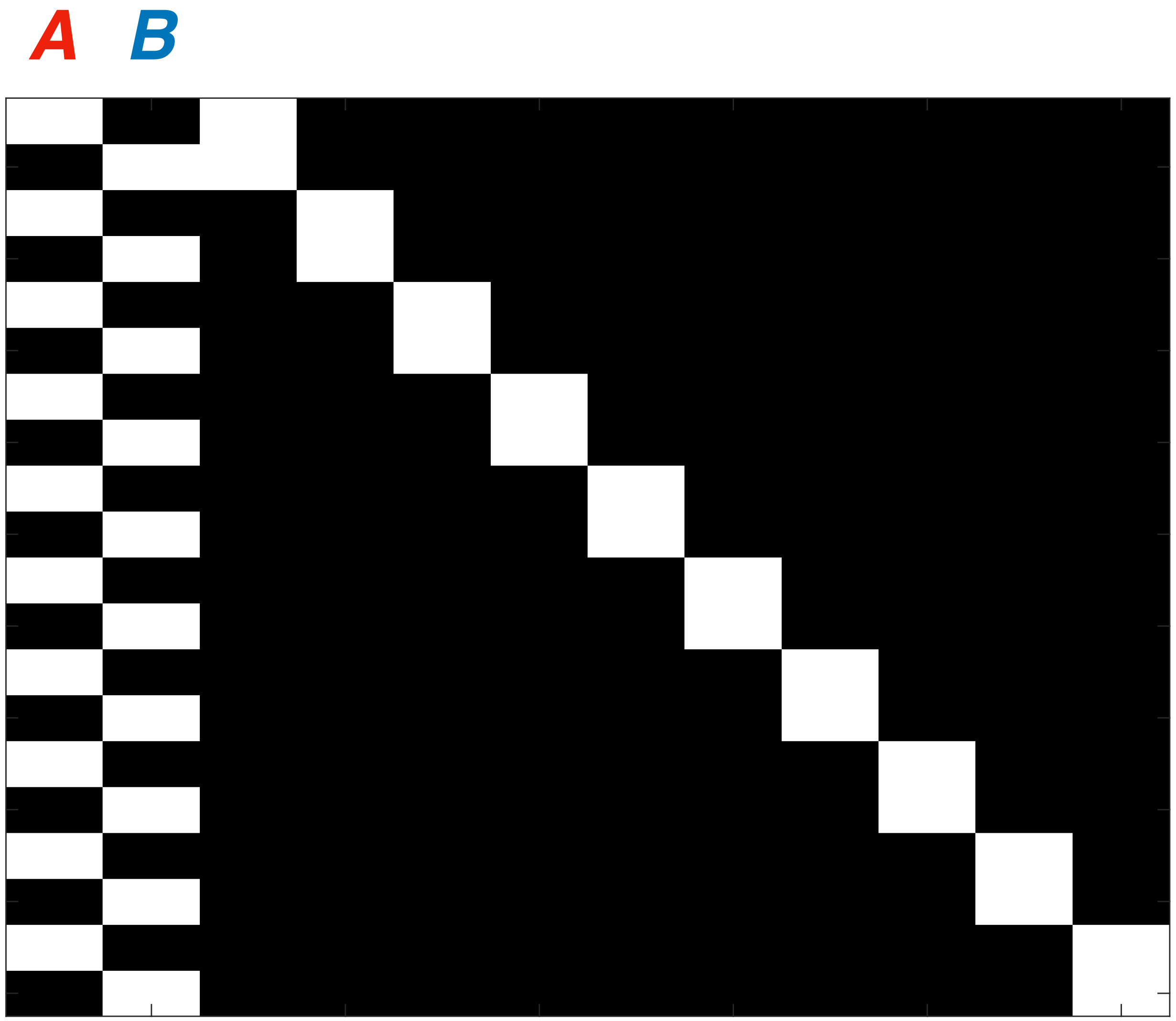
$$\hat{\beta}_1 - \hat{\beta}_2 > 0$$

# Paired $t$-test

SPM-$t$



Repeated-measures design




Constrast

$$c = [1 \text{ -1 } 0 \ \ldots 0]^\top$$

One-sided hypothesis test

$$c^\top \hat{\beta} \leq 0 \quad \text{(null)}$$

$$c^\top \hat{\beta} > 0 \quad \text{(alternative)}$$

# $F$-test: extra sum-of-squares principle

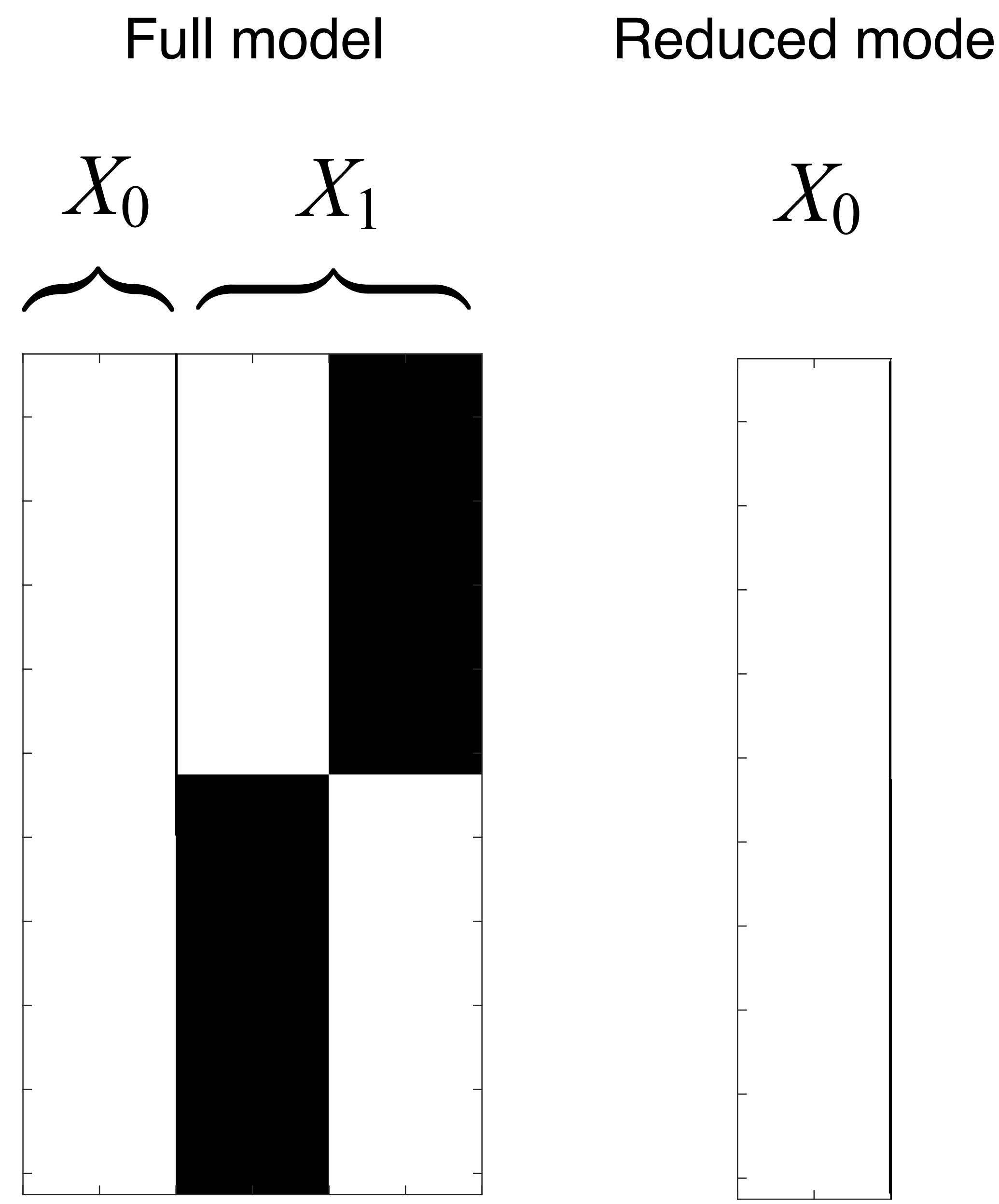
$F$-statistic

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

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

$$RSS = \sum \epsilon^2_{full}$$

$$RSS_0 = \sum \epsilon^2_{reduced}$$



# $F$-contrast

## Reduced model

$$X_0 = Xc_0$$

$$c_0 = I_p - cc^{-}$$ (Contrast orthogonal to $c$)

## Projection matrix

$$M = R_0 - R$$

$$R_0 = I_n - X_0 X_0^{-}$$ (Residual forming matrix of $X_0$)

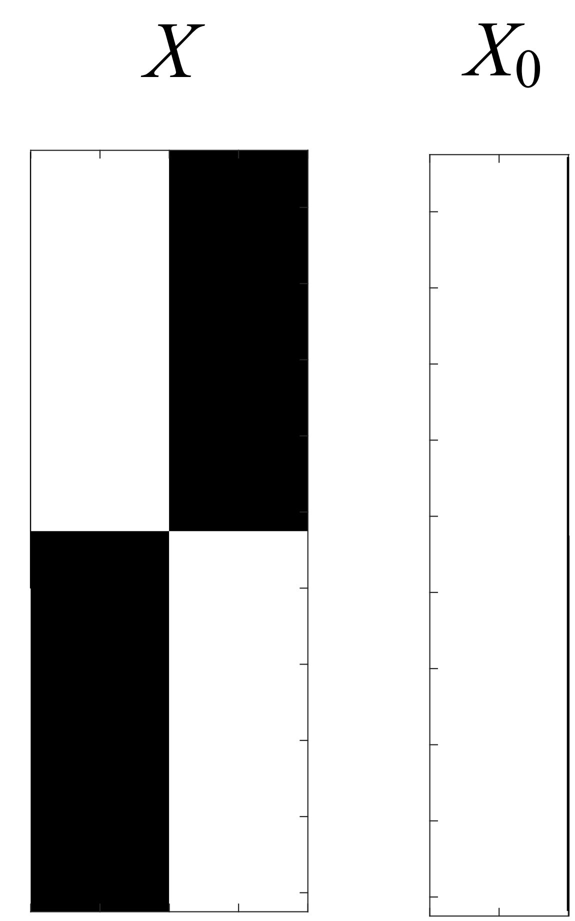
$$R = I_n - XX^{-}$$ (Residual forming matrix of $X$)

## $F$-statistic

$$F = \frac{Y^\top M Y / \nu_1}{Y^\top R Y / \nu_2} \sim F_{\nu_1, \nu_2}$$

## Contrast matrix

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



# *F*-statistic

Variance explained by full model

$$F = \frac{Y^\top M Y / \nu_1}{Y^\top R Y / \nu_2} \sim F_{\nu_1, \nu_2}$$

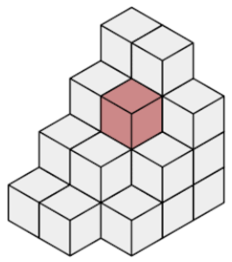
Variance explained by reduced model

Degrees of freedom

$\nu_1 = \mathrm{rank}(X) - \mathrm{rank}(X_0)$ (*Design* degrees of freedom)

$\nu_2 = N - \mathrm{rank}(X)$ (*Error* degrees of freedom)

# $F$-contrast




Constrast matrix

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

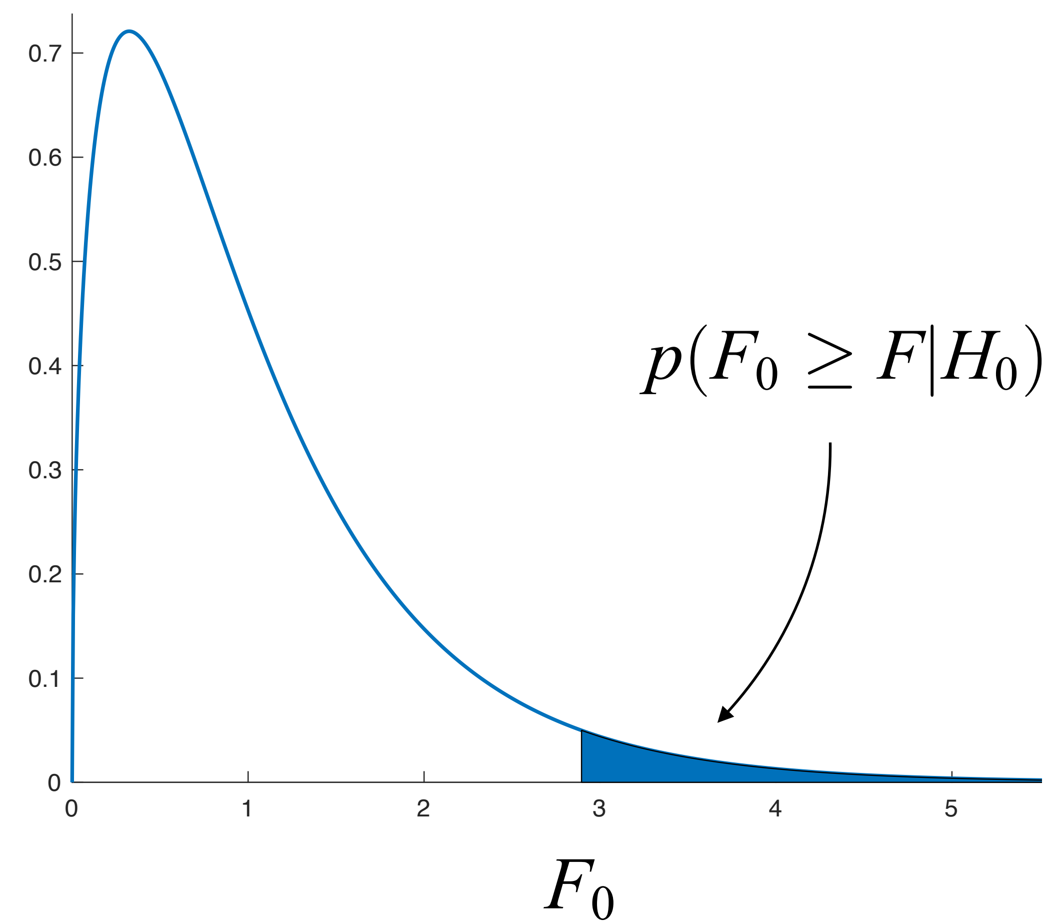
$F$ statistic

$$F = \frac{Y^\top M Y / \nu_1}{Y^\top R Y / \nu_2} \sim F_{\nu_1,\nu_2}$$

$P$ value

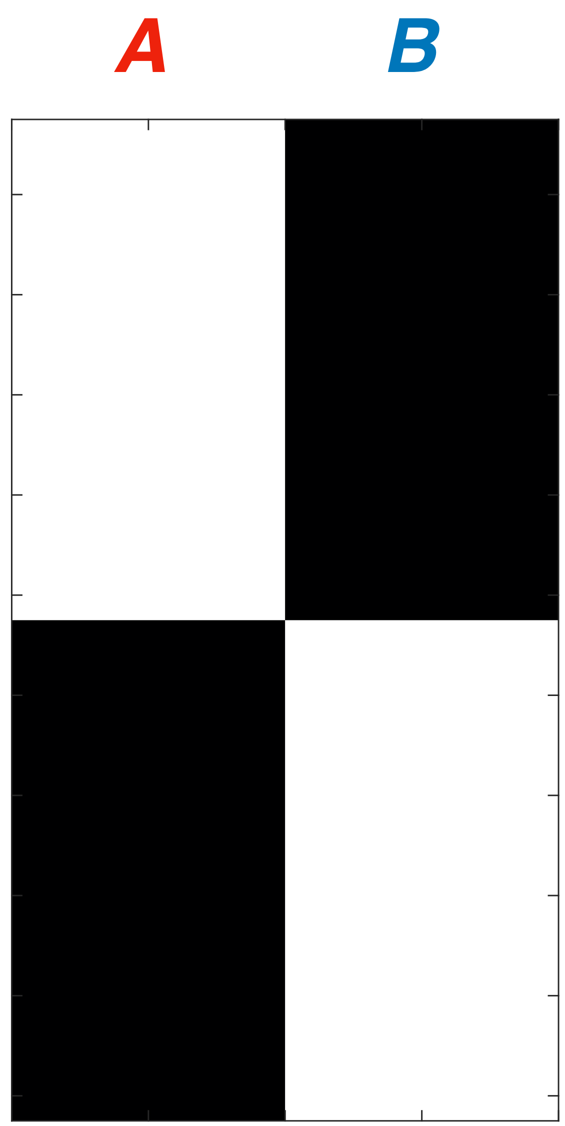
$$p(F_0 \geq F | H_0)$$

$F$ null distribution

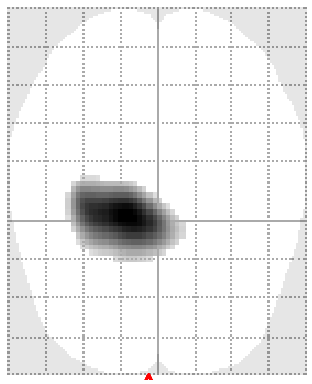


# *F*-test and uni-dimensional contrasts

Between-groups design




SPM-*F*



Constrast vector

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

Two-sided hypothesis test

$$c^\top \hat{\beta} = 0 \quad \text{(null)}$$

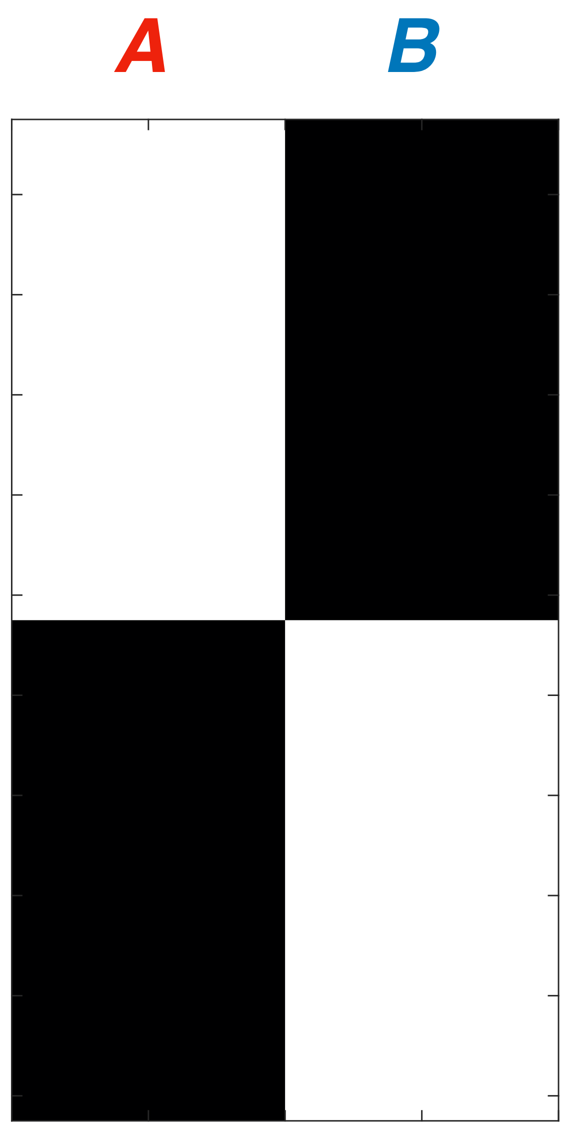
$$c^\top \hat{\beta} \neq 0 \quad \text{(alternative)}$$

Uni-dimensional test of parameters

$$\text{Testing } \hat{\beta}_1 - \hat{\beta}_2 = \hat{\beta}_2 - \hat{\beta}_1$$

# $F$-test and multi-dimensional contrasts

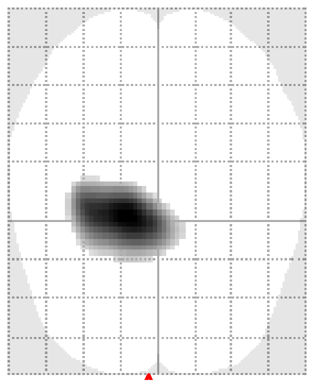
Between-groups design




Constrast matrix

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

SPM-$F$



Two-sided hypothesis test

$$c^\top \hat{\beta} = 0 \quad \text{(null)}$$

$$c^\top \hat{\beta} \neq 0 \quad \text{(alternative)}$$

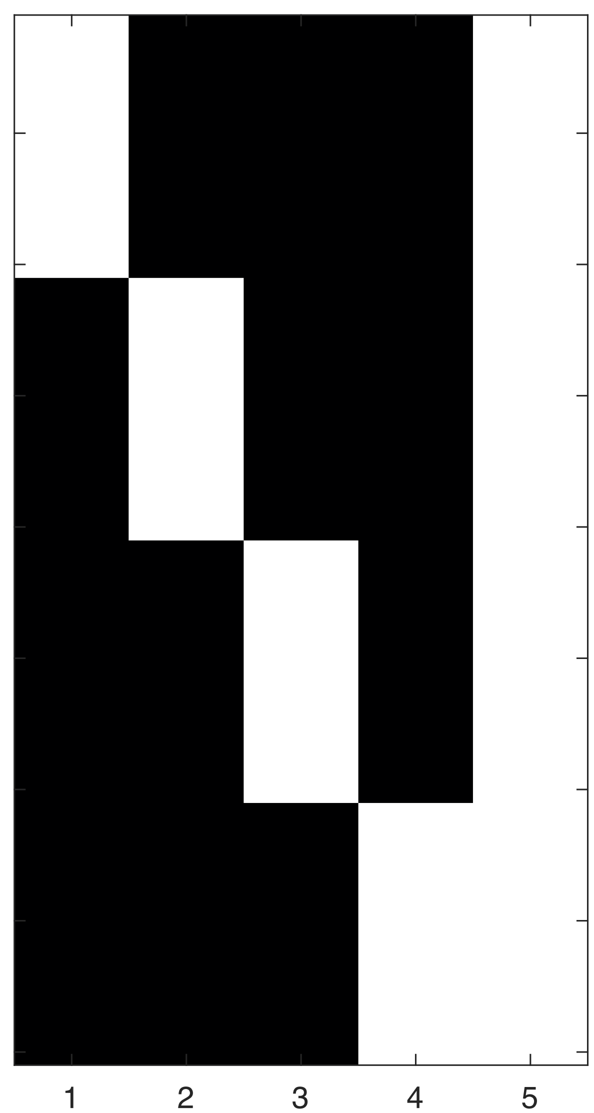
Multi-dimensional test of parameters

$$H_0 \ : \hat{\beta}_1 = \hat{\beta}_2 = 0$$

$$H_A \ : \exists \hat{\beta}_k \in \hat{\beta} \neq 0 \quad \text{(at least one } \hat{\beta}_k \text{)}$$

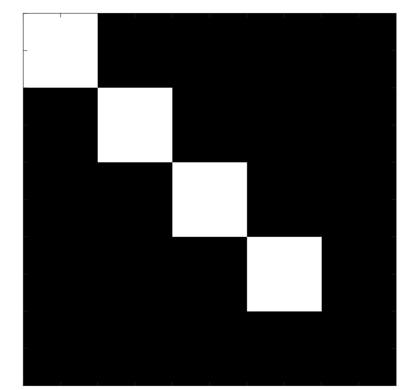
# F-contrast: *any effect*

One-way ANOVA




Constrast matrix

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



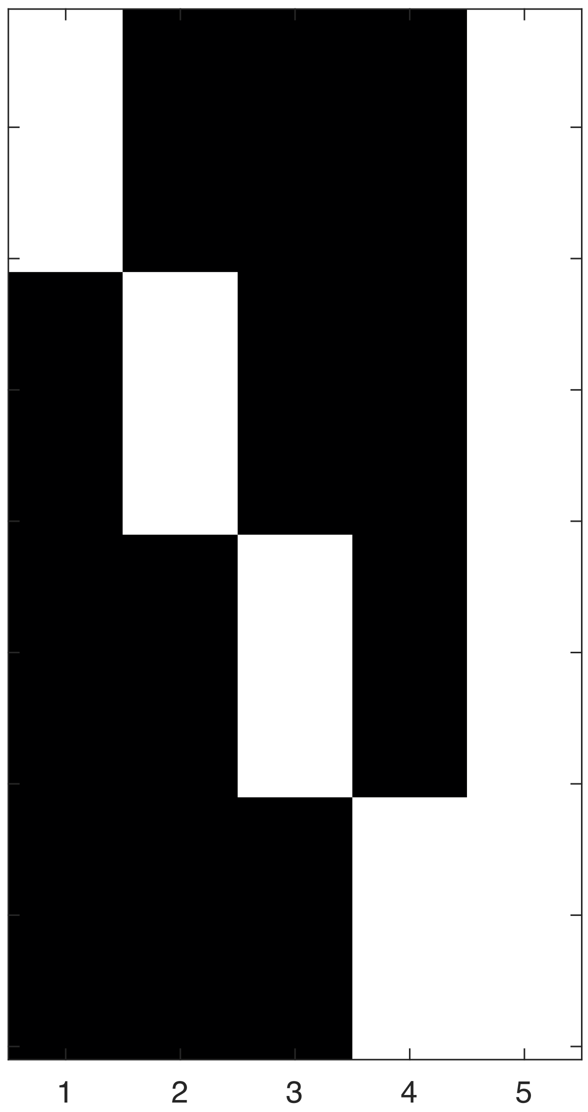
Multi-dimensional hypothesis test

$$H_0 \; : \hat{\beta}_1 = \hat{\beta}_2 = \hat{\beta}_3 = \hat{\beta}_4 = 0$$

$$H_A \; : \exists \hat{\beta}_k \in \hat{\beta} \neq 0$$ (at least one $\hat{\beta}_k$ )

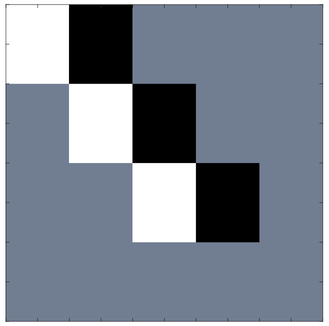
# F-contrast: *any difference*

One-way ANOVA




Constrast matrix

$$C = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{\top}$$



Multi-dimensional hypothesis test

$$H_0 \ : C^{\top}\hat{\beta} = 0$$

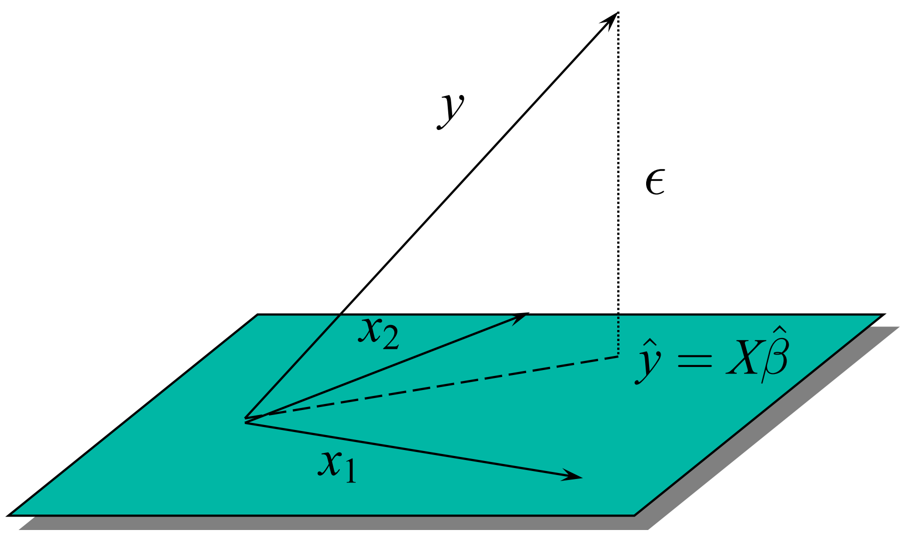
$$H_A \ : \exists \, c_k^{\top}\hat{\beta} \in C^{\top}\hat{\beta} \ \neq 0$$ (at least one contrast)

# GLM summary

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

Special cases of the GLM

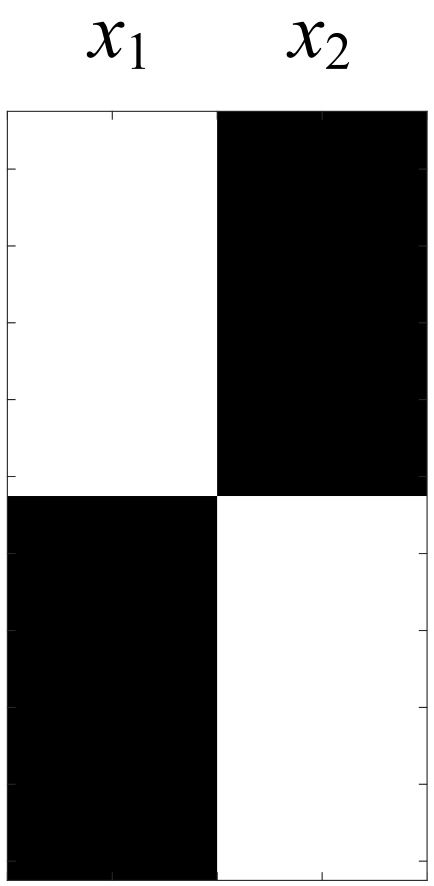
- *t*-test
- *F*-test
- multiple regression
- Analysis of variance (ANOVA)
- Analysis of covariance (AnCova)



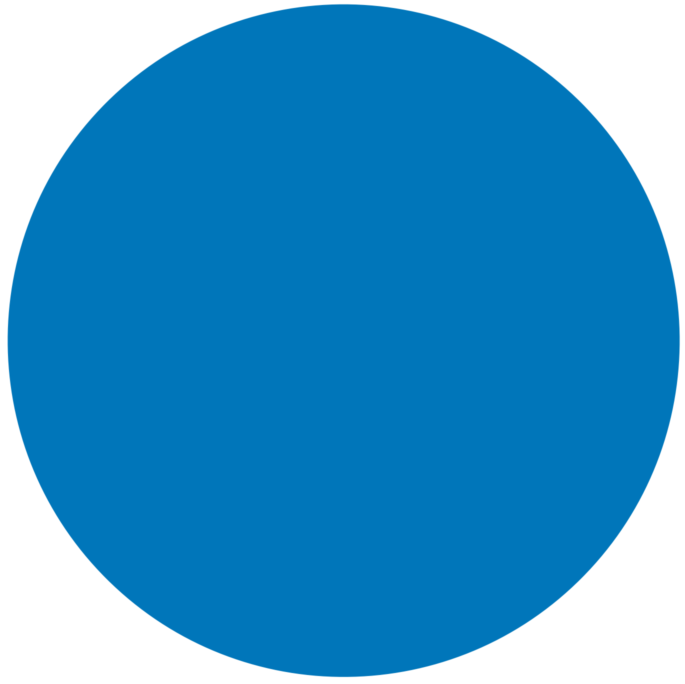

Design space of *X*

# Orthogonal regressors

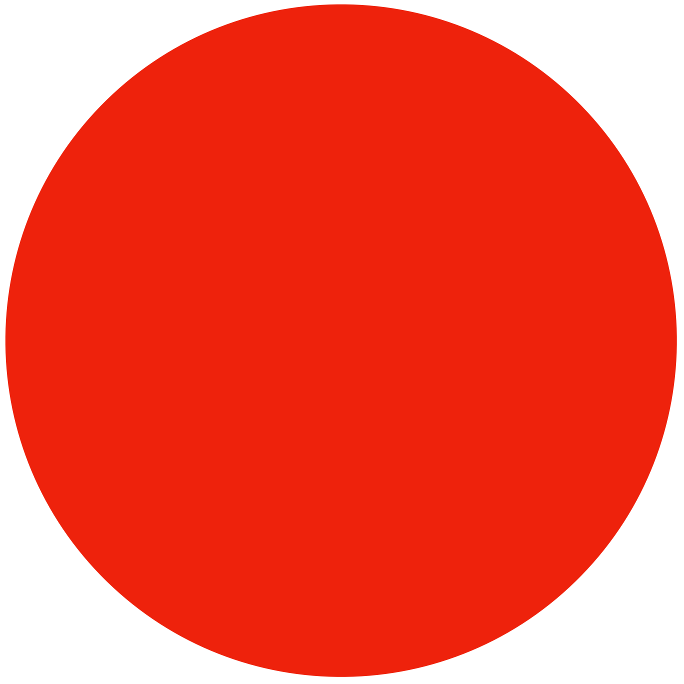
Design matrix




Variance explained by $x_1$

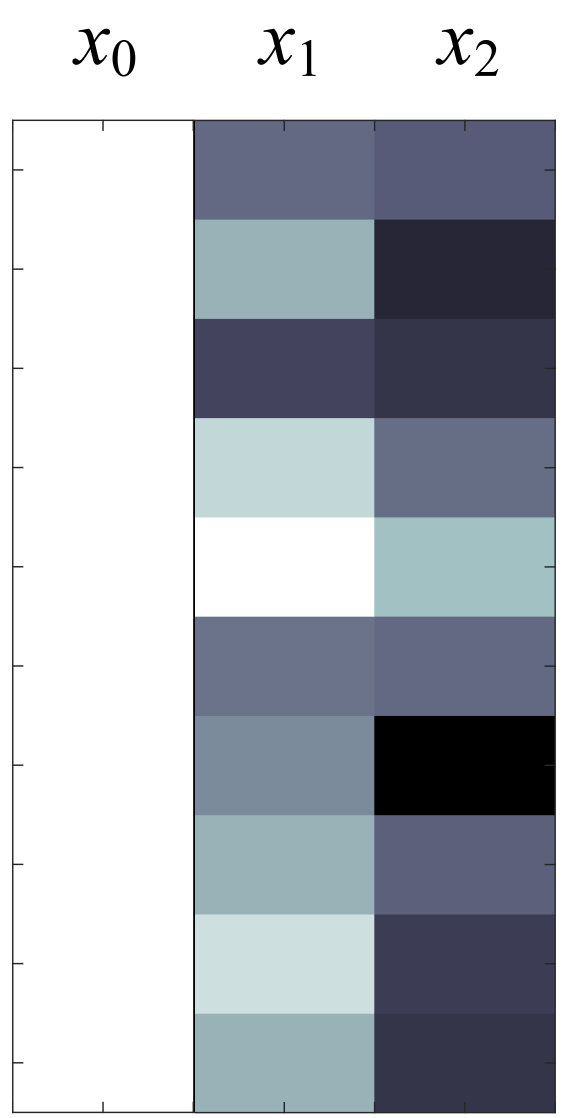


Variance explained by $x_2$

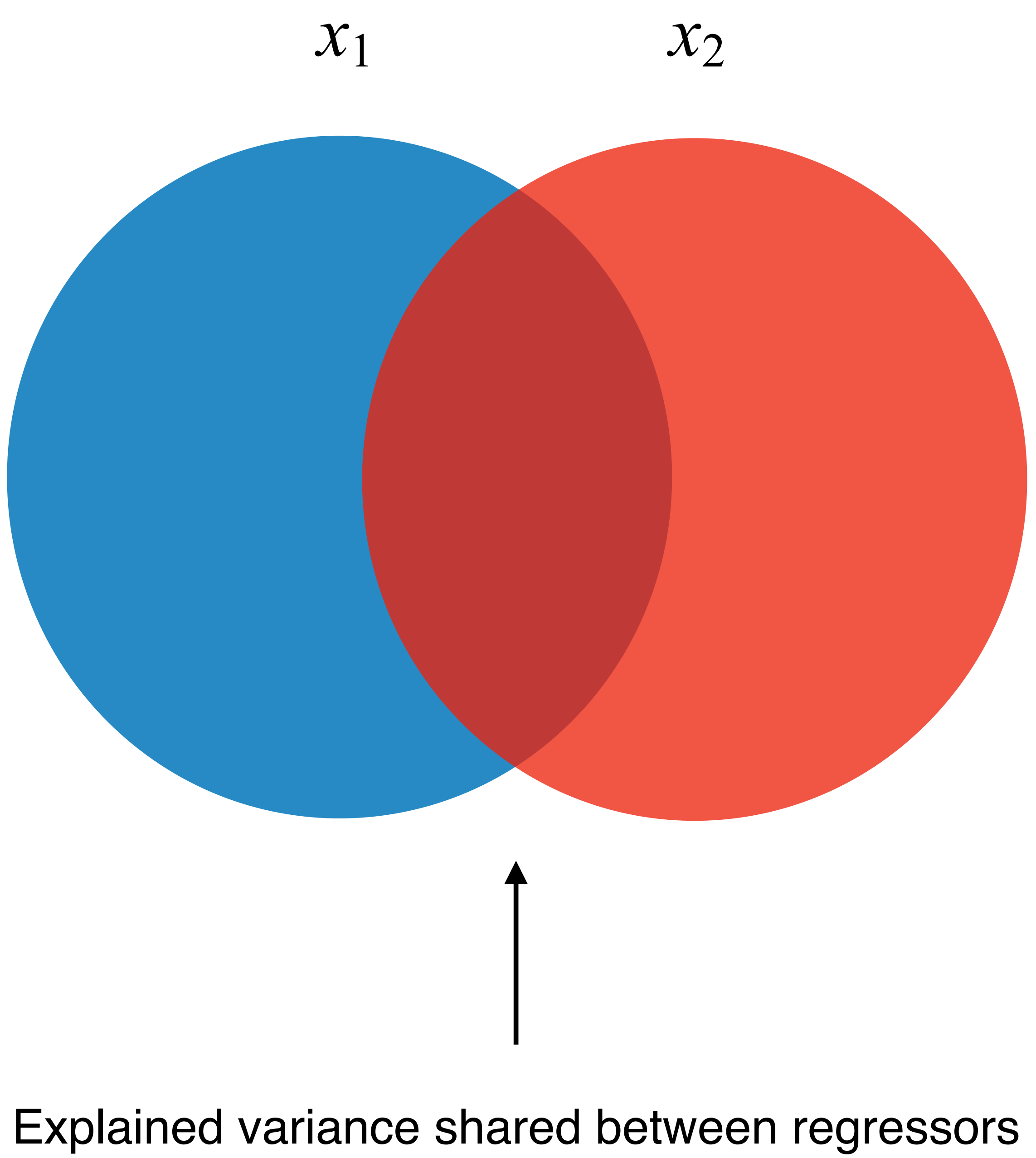


# Correlated regressors

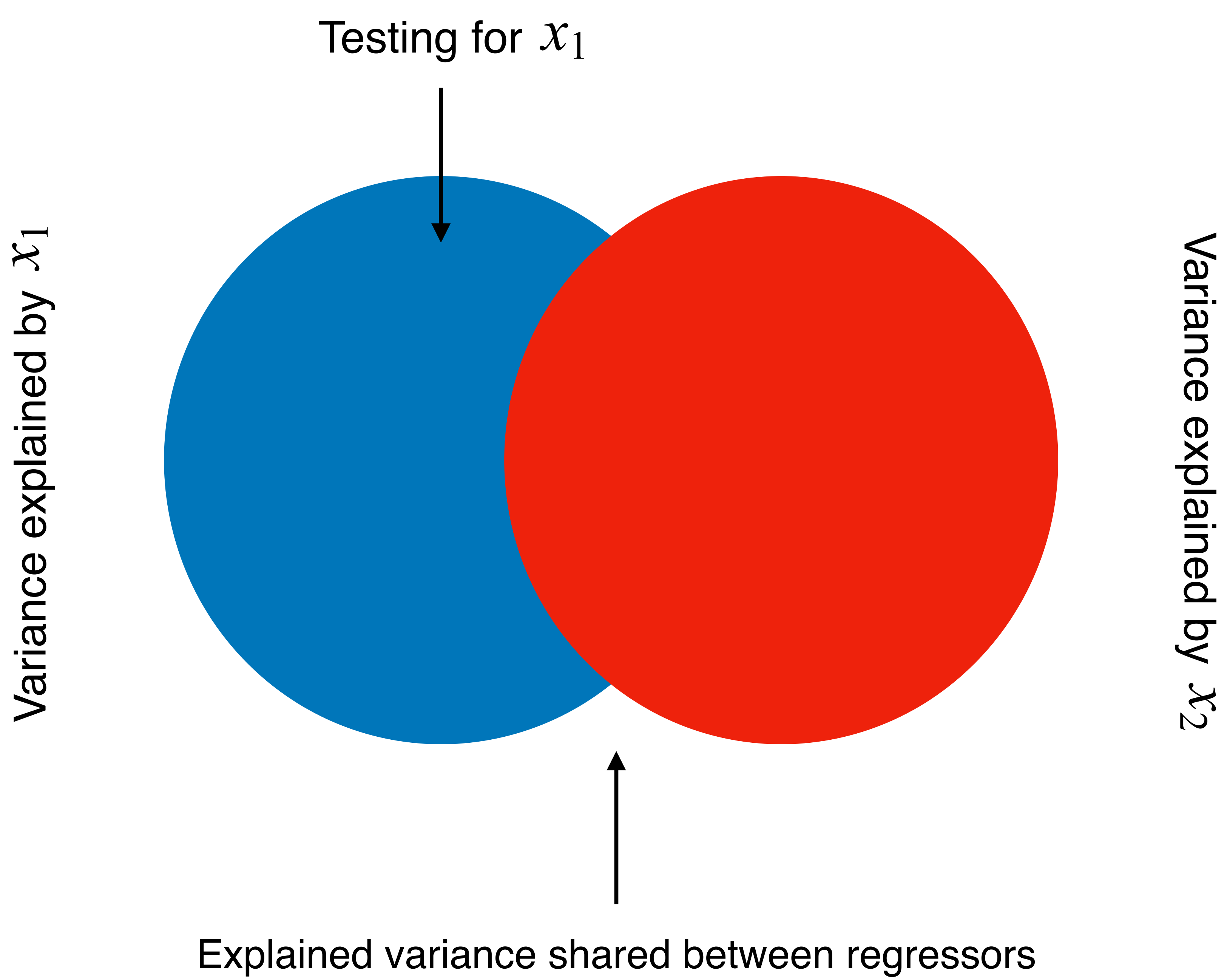
Regression model



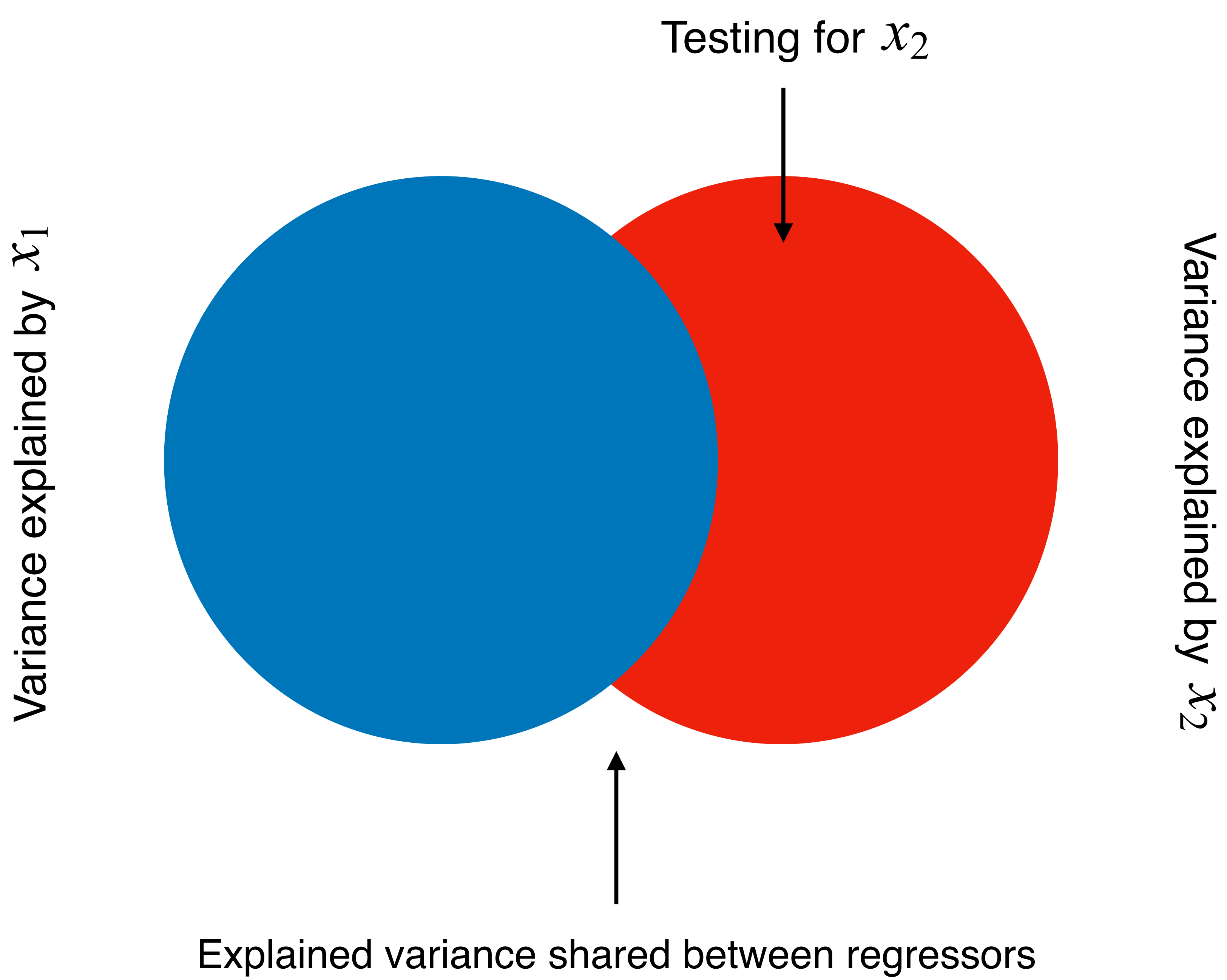

$\cos(\phi) = +/-$



# Correlated regressors

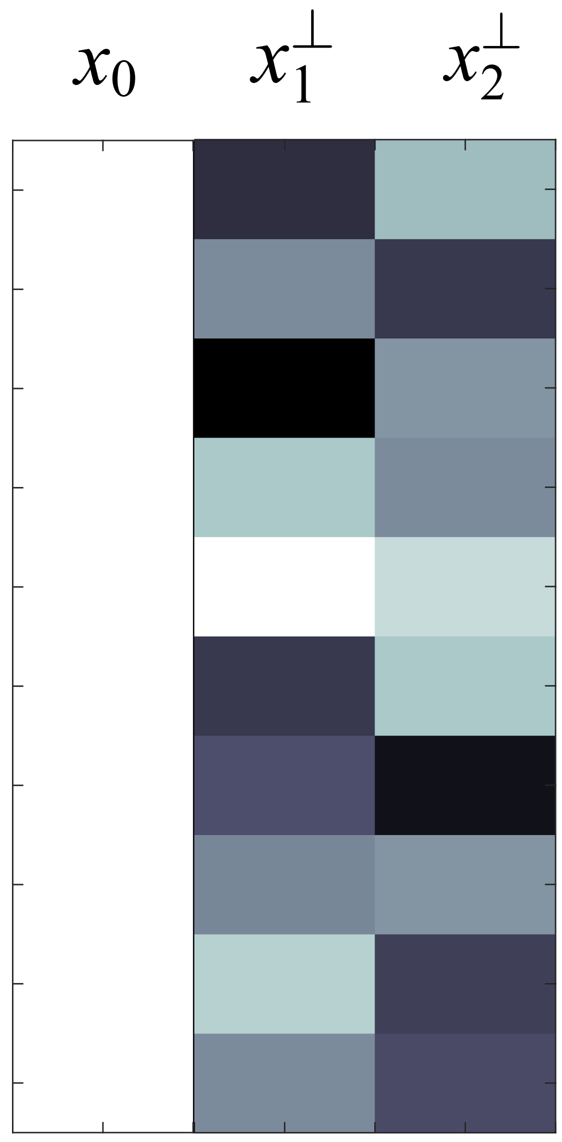


# Correlated regressors

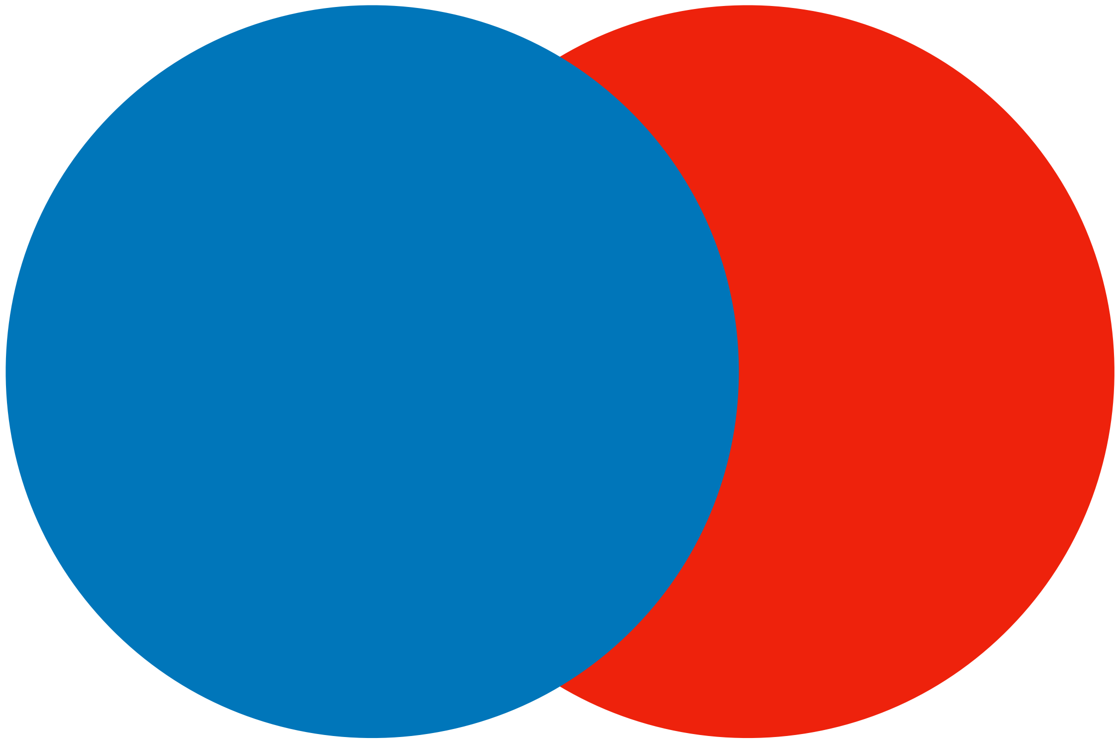


# Orthogonalised regressors

Regression model

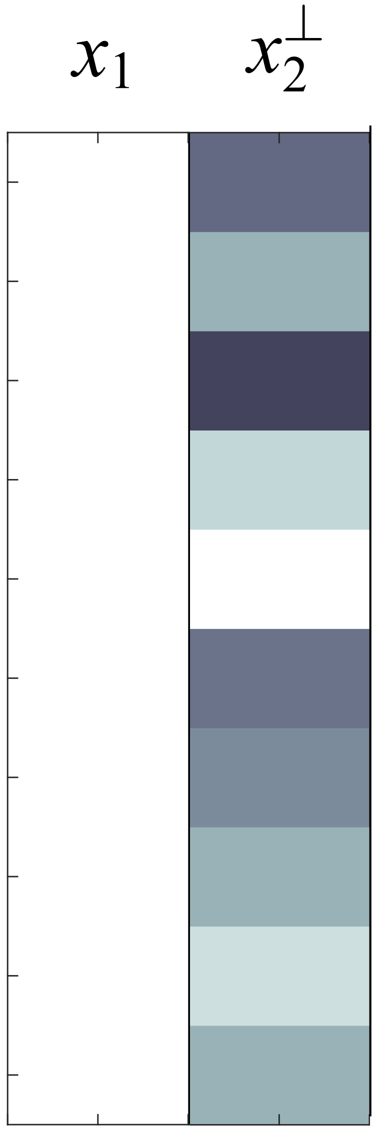



$\cos(\phi) = 0$



# Orthogonalised regressors

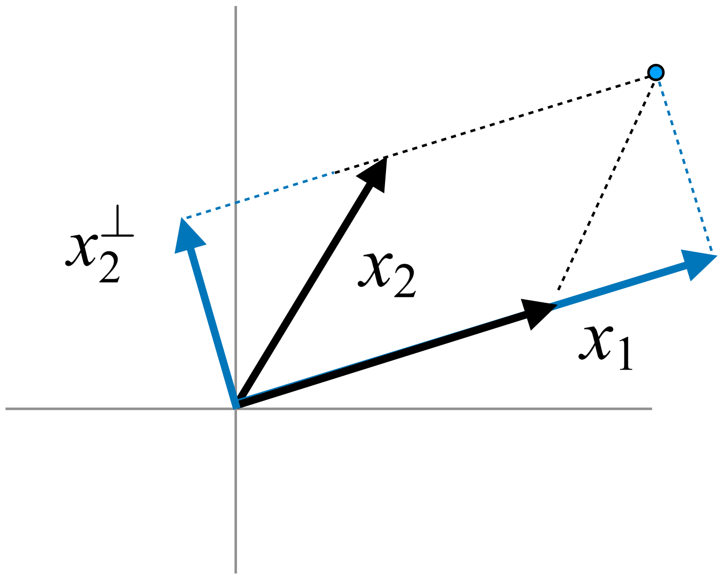
## Regression model




$\cos(\phi) = 0$

## Gram-Schmidt orthogonalisation

$$x_2^\perp = x_2 - \frac{x_1^\top x_2}{x_1^\top x_1}$$



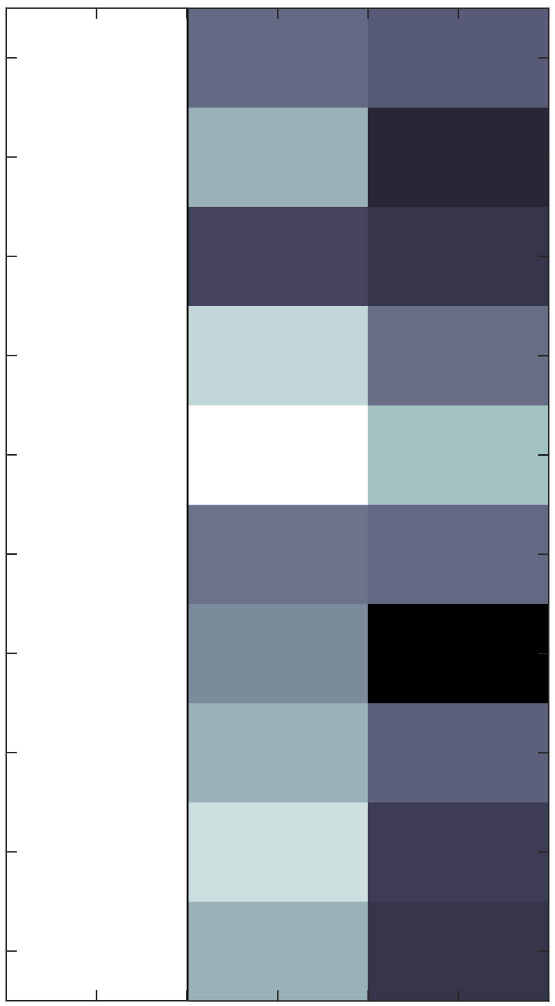

## Unique parameters estimates

$$\begin{bmatrix} \hat{\beta}_1^\perp \\ \hat{\beta}_2 \end{bmatrix} = (X^\top X)^- X^\top y$$

$\hat{\beta}_2 \rightarrow \hat{\beta}_2$

$\hat{\beta}_1 \rightarrow \hat{\beta}_1^\perp$

# Correlated regressors

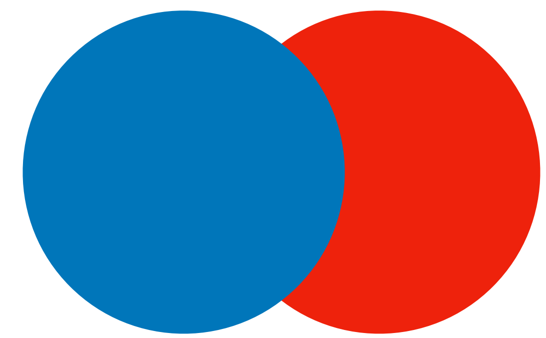



$\cos(\phi) = +/-$

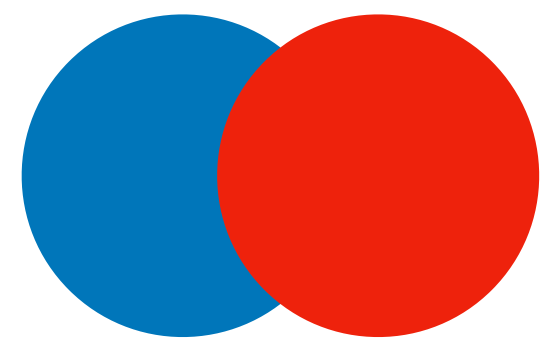
Test full subspace of $Xc$

$$c = (X^\top X)c$$

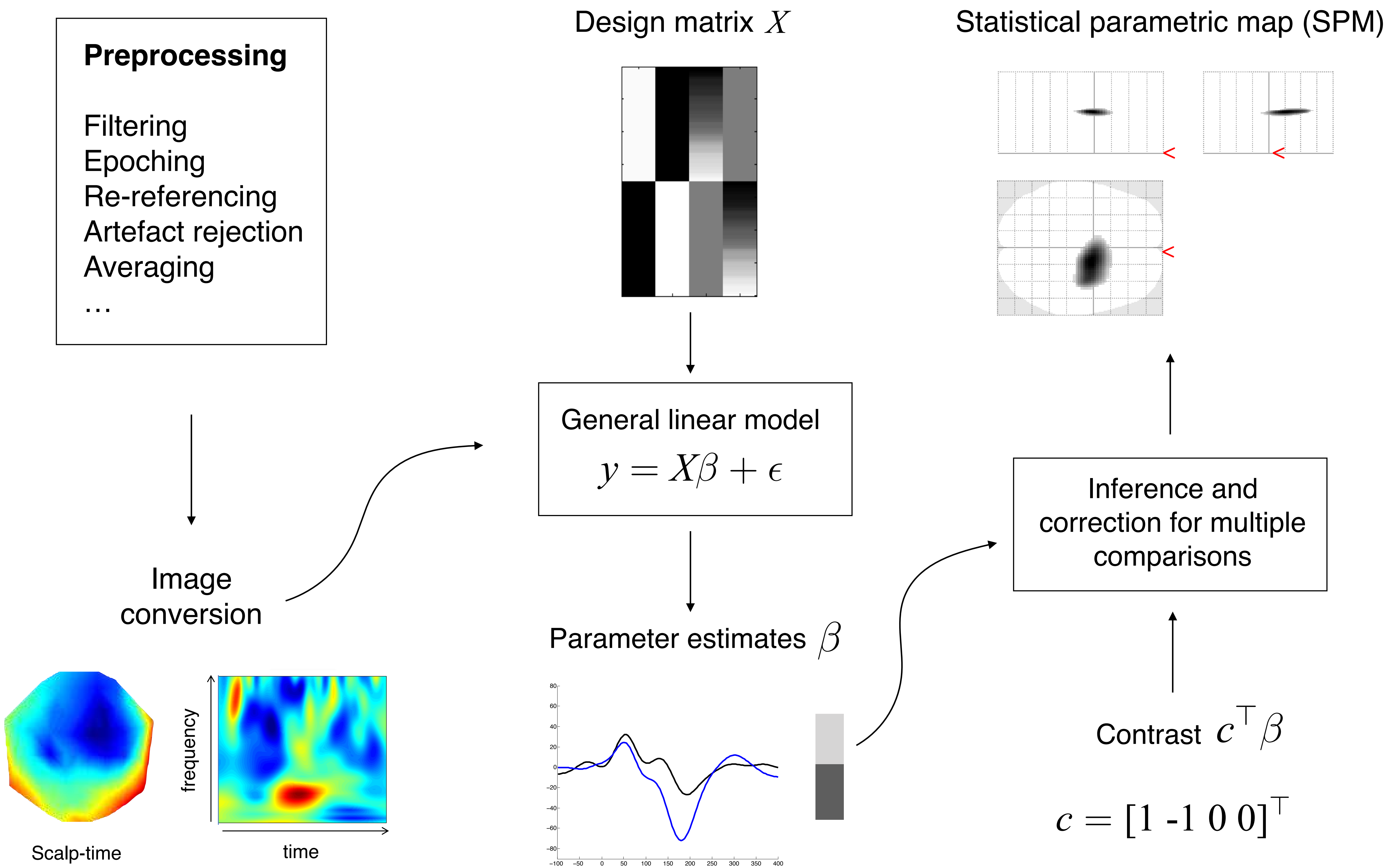
$$c = (X^\top X)[0\ 1\ 0]^\top$$



$$c = (X^\top X)[0\ 0\ 1]^\top$$



# Summary



Preprocessing
Filtering
Epoching
Re-referencing
Artefact rejection
Averaging
…
Image conversion
Scalp-time
frequency
time
Design matrix $X$
General linear model
$y = X\beta + \epsilon$
Parameter estimates $\beta$
Statistical parametric map (SPM)
Inference and correction for multiple comparisons
Contrast $c^\top \beta$
$c = [1\ \text{-}1\ 0\ 0]^\top$

*Thanks to*

SPM methods group and collaborators

Bernadette van Wijk
Christoph Mathys
Christophe Phillips
Gareth Barnes
Guillaume Flandin
James Kilner
Karl Friston
Marta Garrido
Peter Zeidman
Rik Henson
Vladimir Litvak
Will Penny