Understanding contrasts

Dr. Martyn McFarquhar

Division of Neuroscience & Experimental Psychology
The University of Manchester

MSc in Neuroimaging for Clinical and Cognitive Neuroscience

Edinburgh SPM course, 2017

What are contrasts?

Informally, a **contrast** is a method of **asking a question** about your model in **mathematical terms**

The GLM **partitions** the data into **model** and **error**

$$\mathbf{Y} = \underbrace{\mathbf{X}\boldsymbol{\beta}}_{\text{model}} + \underbrace{\boldsymbol{\epsilon}}_{\text{error}}$$

The **model** is the **typical** value of the data for given values of the **predictor variables**

This is formed via a **linear combination** of **known** predictor variables and **unknown** parameters

As the **parameters** represent the **relationship** between the **predictors** and the **outcome**, their **estimated values** are the part of the model that we are **most interested in**

The University of Manchester

MANCHESTER 1824

THE UNIVERSITY OF EDINBURGH

Formal definition

General Linear Hypothesis

All hypothesis testing in the GLM is based on the following

$$\mathcal{H}_0 : \mathbf{L}\boldsymbol{\beta} = \mathbf{m}$$

L is the *m* x *k* **matrix** of **weights**

β is the *k* x 1 **vector** of **model parameters**

m is the *m* x 1 **vector** of **proposed values**

The hypothesis is that some **linear combination**(**s**) of the **parameters** equals some **proposed value**(**s**)

In **SPM** the matrix **m** is fixed to contain **all zeros** — each **linear combination** of the parameters in the **rows** of **L** are equal to **0**

Formal definition

General Linear Hypothesis

All hypothesis testing in the GLM is based on the following

$$\mathcal{H}_0 : \mathbf{L}\boldsymbol{\beta} = \mathbf{m}$$

This is a very **flexible** system

$$\mathcal{H}_0 : \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0$$

β_1 is equal to 0

$$\mathcal{H}_0 : \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

β_1 **or** β_2 is equal to 0

$$\mathcal{H}_0 : \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0$$

the **difference** between β_1 and β_2 is 0

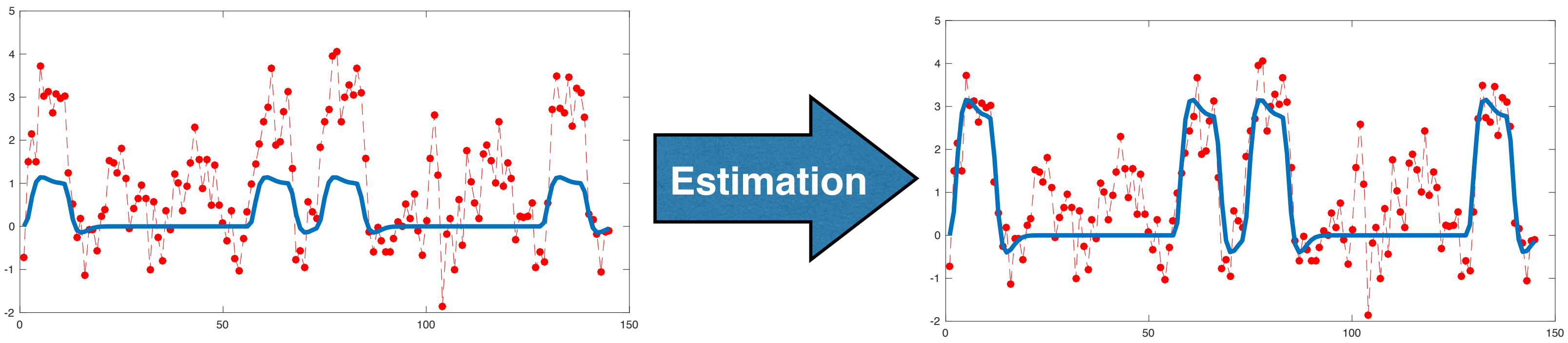
$$\mathcal{H}_0 : \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0$$

the **average** of β_1 and β_2 is 0

Parameters and interpretation

As **contrasts** are **linear combinations** of parameters, their **interpretation** depends on understanding the parameters

1st-level



The parameters **scale** the **predicted shape of the response** to **best fit the data**

Their values tell us about the **magnitude** and **direction** of the **average change from baseline** in each experimental condition

Parameters and interpretation

As **contrasts** are **linear combinations** of parameters, their **interpretation** depends on understanding the parameters

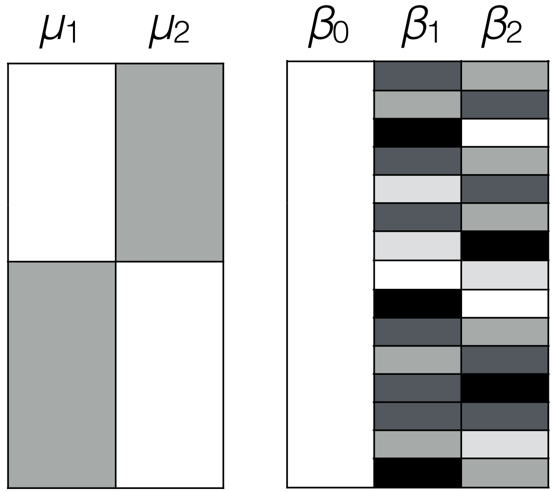
2nd-level

Indicator variables

- Contain only 1 or 0 to model a **factor**
- Parameters are **cell means**

Continuous covariates

- Contain any value **within a range**
- Parameters are **regression slopes**

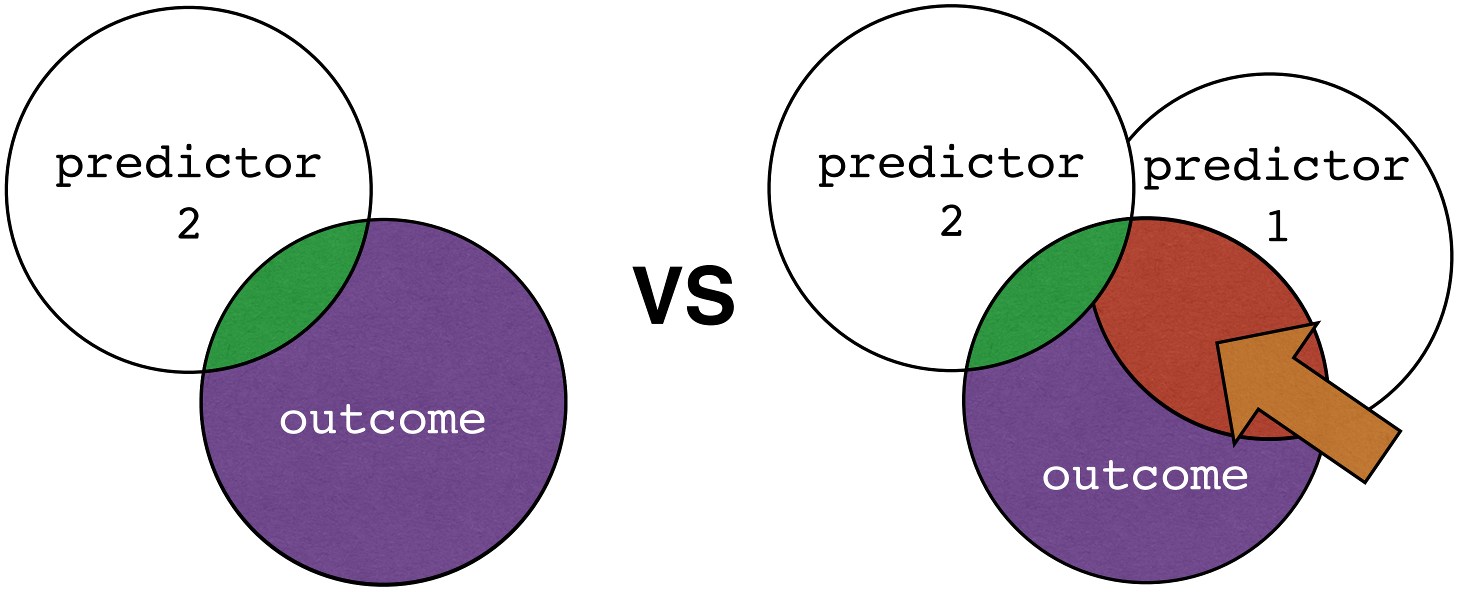


Comparing **cell means** is sensible, but comparing **regression slopes** only makes sense if the predictors are **scaled identically**

Parameters and interpretation

For **multiple** predictor variables, tests on each **parameter** is interpreted as testing the **unique variability** explained **after adjusting for all other variables**

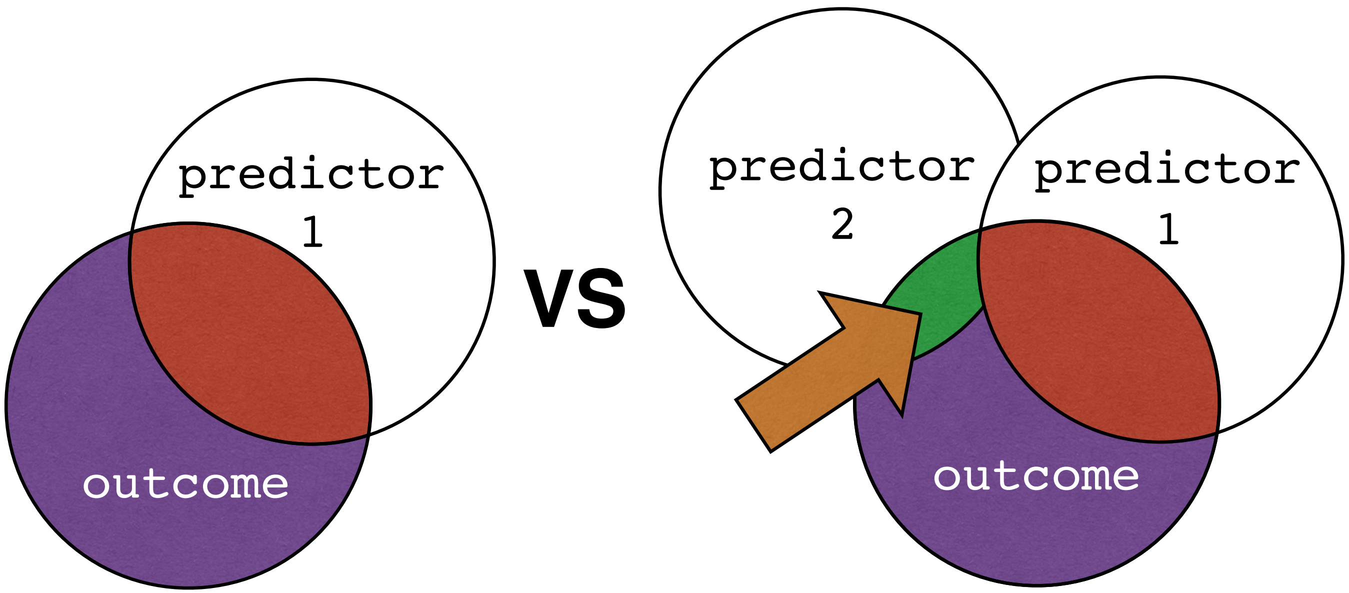
Effect of predictor 1



Parameters and interpretation

For **multiple** predictor variables, tests on each **parameter** is interpreted as testing the **unique variability** explained **after adjusting for all other variables**

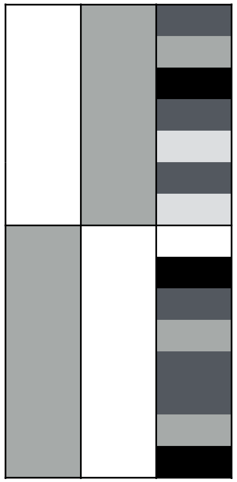
Effect of predictor 2



Parameters and interpretation

For **multiple** predictor variables, tests on each **parameter** is interpreted as testing the **unique variability** explained **after adjusting for all other variables**

For **contrasts**, this means that even if a parameter is given a **weight of 0**, the other parameters are **still adjusted** for its presence



$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

ANCOVA

The individual **cell means** are still **adjusted** for the **covariate**

The value of the **parameter estimates** depends on **other variables in the model**

Contrasts ask questions about the model — they **do not alter the model**

Parameters and interpretation

Parameters are **estimated** using

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

Notice that this **depends** on the **design matrix**

Important consequences

1. The estimates depend on the **scaling** of the columns of **X**
2. If **X** is **rank deficient** (contains redundant columns) then there are an **infinite number of solutions**

If **X** is rank deficient then $(\mathbf{X'X})^{-1}$ **does not exist**

We can still find a **solution** using a **pseudo**-**inverse** — creates **complications** in **specifying** and **interpreting** contrasts

Estimable functions

An **L** matrix defines an **estimable function** of the model parameters if it can be expressed as

$$\mathbf{L} = \mathbf{TX}$$

Our contrast is only **meaningful** if it is formed from a **linear combination** of the **rows** of the **design matrix**

An **estimable contrast** is also one that can be expressed as a **linear combination** of the **model estimates**

$$\begin{aligned}\hat{\mathbf{c}} &= \mathbf{L}\hat{\boldsymbol{\beta}} \\ &= \mathbf{TX}\hat{\boldsymbol{\beta}} \\ &= \mathbf{T}\hat{\mathbf{Y}}\end{aligned}$$

The **model estimates** dictate how **meaningful** values are formed from combining the **predictors** and the **parameters**

A question is only meaningful if it **respects this combination**

Estimable functions

Example - overparameterised 2x2 ANOVA

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$$

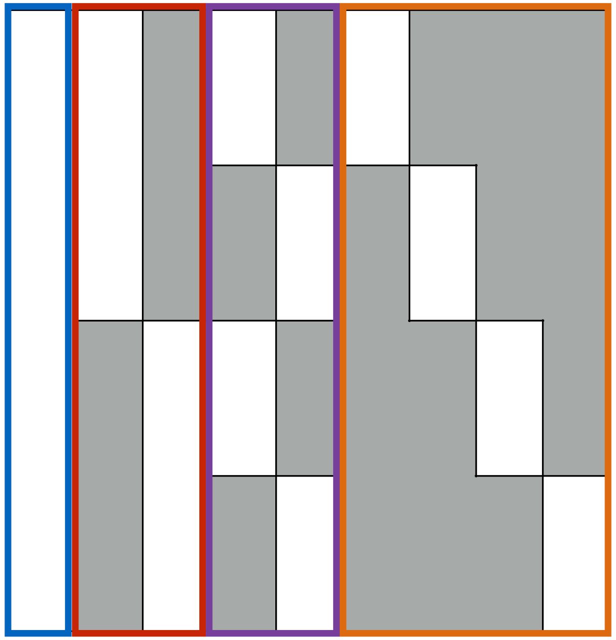
Here the **ANOVA effects** are **explicit** in the model

Constant

Main effect of A

Main effect of B

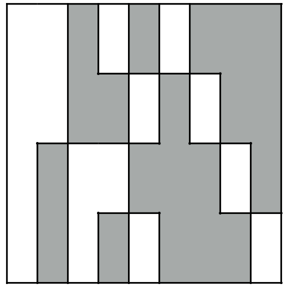
A x B interaction



Estimable functions

Example - overparameterised 2x2 ANOVA

The **design matrix** is **rank-deficient** so estimation requires use of a **pseudo-inverse**



$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{+}\mathbf{X}'\mathbf{Y}$$

The parameters have a **degree of arbitrariness** to their values — certain combinations are **still meaningful**

Find 2 number that sum to 5

$$a + b = 5$$

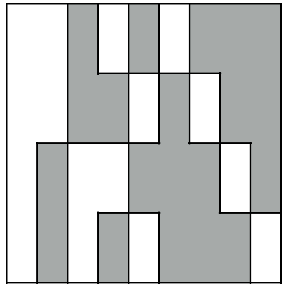
$$a = 2, \quad b = 3 \qquad a = -6.37, b = 11.37$$

$$a = 0.4, b = 4.6 \qquad a = 4322, \quad b = -4317$$

There are **infinite choices** — the individual values are **arbitrary** but the sum **is not**

Estimable functions

Example - overparameterised 2x2 ANOVA



Contrasts must respect the **meaningful combinations of parameters** given by the **rows** of **X**

Main effect of **A**?

$$\mathbf{L} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad ✗$$

$$\left.\begin{aligned} \mathbf{L}_{\mathrm{A1B1}} &= \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \\ \mathbf{L}_{\mathrm{A1B2}} &= \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix} \end{aligned}\right\} \mathbf{L}_{\mathrm{A1}} = \begin{bmatrix} 1 & 1 & 0 & 0.5 & 0.5 & 0.5 & 0.5 & 0 & 0 \end{bmatrix}$$

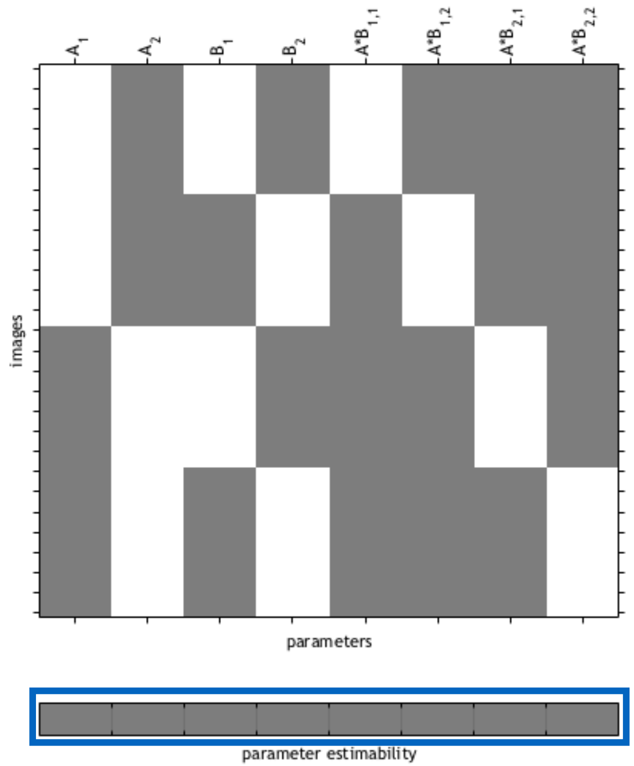
$$\left.\begin{aligned} \mathbf{L}_{\mathrm{A2B1}} &= \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\ \mathbf{L}_{\mathrm{A2B2}} &= \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned}\right\} \mathbf{L}_{\mathrm{A2}} = \begin{bmatrix} 1 & 0 & 1 & 0.5 & 0.5 & 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

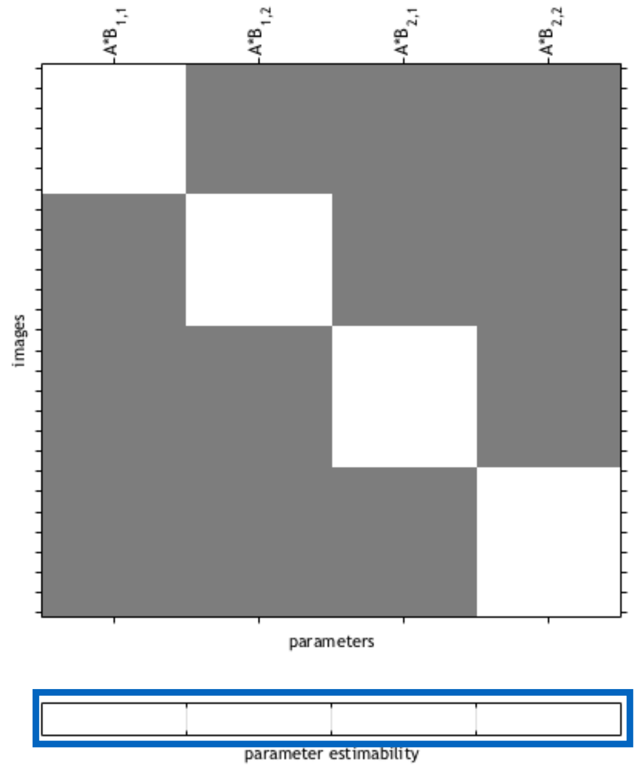
Correct main effect of **A**

$$\mathbf{L}_{\mathrm{A1-A2}} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0.5 & 0.5 & -0.5 & -0.5 \end{bmatrix} \quad ✔$$

Estimable functions

Estimable functions in SPM





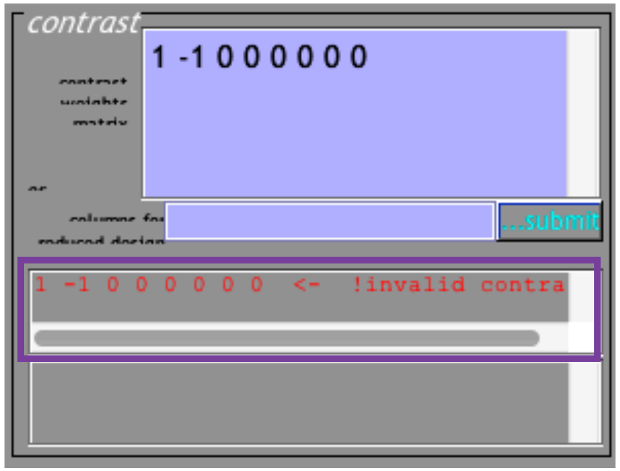
Non-estimable parameters are indicated by a **grey box** below the associated column

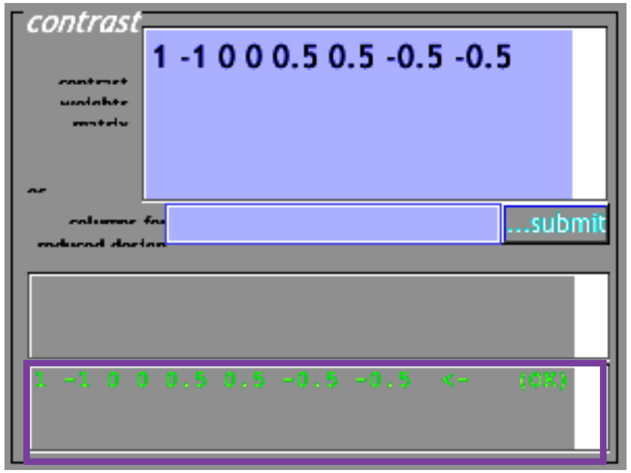
Estimable functions

Estimable functions in SPM

To make sure that the contrast we are testing is **estimable**, SPM will perform an **estimability test**

This ensures that we are testing combinations that **do not depend** on the solution for the parameters (see McFarquhar, 2016)





Estimable functions

Estimable functions in SPM

$$\mathbf{T} = \mathbf{L}\mathbf{X}^{+}$$

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

If **L** is an **estimable function**

$$\mathbf{L} = \mathbf{T}\mathbf{X}$$

$$\mathbf{L} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$
$$\mathbf{T} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$
$$\mathbf{TX} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

If **L** is **not** an **estimable function**

$$\mathbf{L} \neq \mathbf{T}\mathbf{X}$$

$$\mathbf{L} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
$$\mathbf{T} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$
$$\mathbf{TX} = \begin{bmatrix} 0 & \frac{2}{3} & -\frac{2}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$

Contrast interpretation

Most of the time we work with **well parameterised** models and so do not have to worry about **estimability** — check the grey boxes!

Despite this, a contrast can be **estimable** but **misinterpreted**

This largely comes down to

1. Understanding **what** the parameters mean
2. Making sure that the contrast is testing **what you think it is**

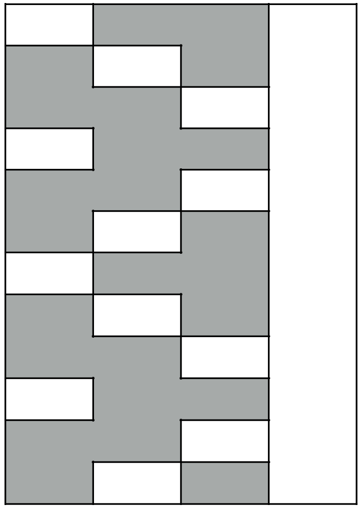
You should always ensure that you are **clear** how to **interpret** the **parameters** and **contrasts** from your model

How can you know what your results **mean** of you don't understand what the model is telling you, or what your questions are?

Contrast interpretation

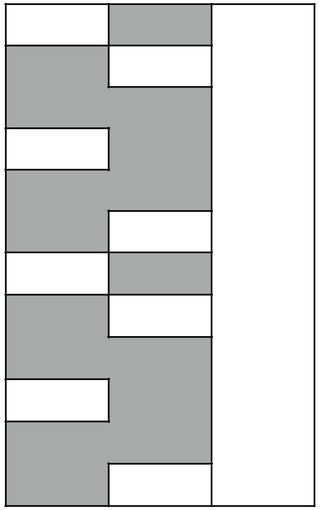
1st-level example — baselines

Model 1
(A, B, rest, const.)



$$\mathbf{L}_{\mathrm{A-B}} = \begin{bmatrix}1 & -1 & 0 & 0\end{bmatrix}$$
$$\mathbf{L}_{\mathrm{A-R}} = \begin{bmatrix}1 & 0 & -1 & 0\end{bmatrix}$$
$$\mathbf{L}_{\mathrm{B-R}} = \begin{bmatrix}0 & 1 & -1 & 0\end{bmatrix}$$
$$\mathbf{L}_{\mathrm{A}} = \begin{bmatrix}1 & 0 & 0 & 1\end{bmatrix}$$
$$\mathbf{L}_{\mathrm{B}} = \begin{bmatrix}0 & 1 & 0 & 1\end{bmatrix}$$

Model 2
(A, B, const.)



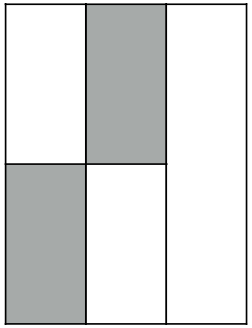
$$\mathbf{L}_{\mathrm{A-B}} = \begin{bmatrix}1 & -1 & 0\end{bmatrix}$$
$$\mathbf{L}_{\mathrm{A-R}} = \begin{bmatrix}1 & 0 & 0\end{bmatrix}$$
$$\mathbf{L}_{\mathrm{B-R}} = \begin{bmatrix}0 & 1 & 0\end{bmatrix}$$
$$\mathbf{L}_{\mathrm{A}} = \begin{bmatrix}1 & 0 & 1\end{bmatrix}$$
$$\mathbf{L}_{\mathrm{B}} = \begin{bmatrix}0 & 1 & 1\end{bmatrix}$$

An **explicit** baseline over an **implicit** baseline alters the **interpretation** of the parameters and the contrasts

Contrast interpretation

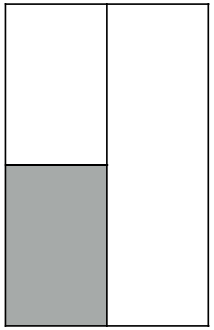
2nd-level example — *t*-test

Model 1
(A, B, const.)



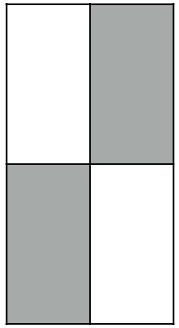
$$\mathbf{L}_{\mathrm{A-B}} = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix}$$
$$\mathbf{L}_{\mathrm{A}} = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$$
$$\mathbf{L}_{\mathrm{B}} = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$$
$$\mathbf{L}_{\mathrm{mean}} = \begin{bmatrix} 0.5 & 0.5 & 1 \end{bmatrix}$$

Model 2
(A, const.)



$$\mathbf{L}_{\mathrm{A-B}} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$
$$\mathbf{L}_{\mathrm{A}} = \begin{bmatrix} 1 & 1 \end{bmatrix}$$
$$\mathbf{L}_{\mathrm{B}} = \begin{bmatrix} 0 & 1 \end{bmatrix}$$
$$\mathbf{L}_{\mathrm{mean}} = \begin{bmatrix} 0.5 & 1 \end{bmatrix}$$

Model 3
(A, B)



$$\mathbf{L}_{\mathrm{A-B}} = \begin{bmatrix} 1 & -1 \end{bmatrix}$$
$$\mathbf{L}_{\mathrm{A}} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$
$$\mathbf{L}_{\mathrm{B}} = \begin{bmatrix} 0 & 1 \end{bmatrix}$$
$$\mathbf{L}_{\mathrm{mean}} = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}$$

Different **parameterisations** lead to the **same** predicted values, but **change the meaning** of the parameters

Need to be **clear** on what the parameters **mean** in order to **correctly interpret a contrast**

Test statistics

Assuming our **contrast** is **estimable** and we are clear on **what it means** we can test the **estimated value** against a proposed **population value** by forming a **test statistic**

In **SPM** we have **two** options for how to test our **contrast value** — as a ***t*-contrast** or an ***F*-contrast**

Each type is used in **different contexts** and it is important to understand their **differences**, and **similarities**, to ensure you use the **most appropriate method** for your questions

t-contrasts

In **SPM** a *t*-contrast is defined by

- An **L** matrix with a **single row**
- Hypothesis testing using a *t*-statistic
- **One-tailed** *p*-values

$$t = \frac{\mathbf{L}\hat{\boldsymbol{\beta}} - \mathbf{m}}{\hat{\sigma}\sqrt{\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'}}$$

Notice that **L** appears in the **numerator** and **denominator** — scaling of **L** does not matter

All lead to the same *t*-statistic:

$$\mathbf{L} = \begin{bmatrix} 1 & -1 \end{bmatrix} \qquad \mathbf{L} = \begin{bmatrix} 2 & -2 \end{bmatrix} \qquad \mathbf{L} = \begin{bmatrix} 1000 & -1000 \end{bmatrix}$$

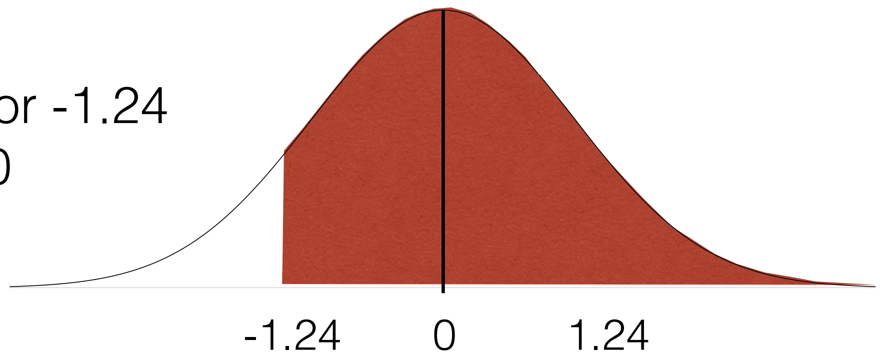
t-contrasts

In **SPM** a *t*-contrast is defined by

- An **L** matrix with a **single row**
- Hypothesis testing using a *t*-statistic
- **One-tailed** *p*-values

The *p*-values are **upper-tail** values — you will only see results for **positive *t*-statistics**

Upper-tail *p* for -1.24
= 0.920



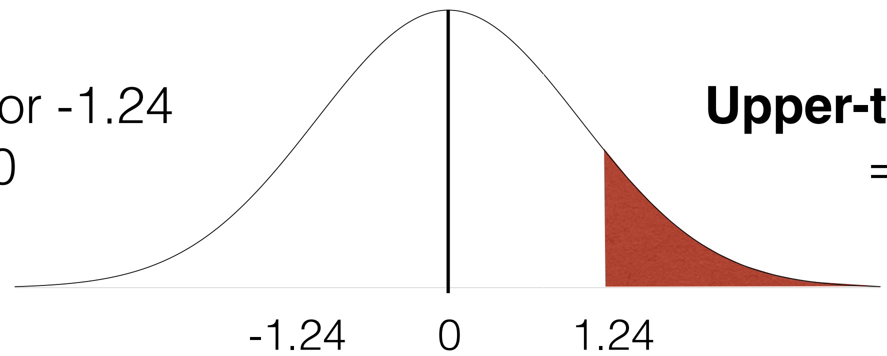
t-contrasts

In **SPM** a *t*-contrast is defined by

- An **L** matrix with a **single row**
- Hypothesis testing using a *t*-statistic
- **One-tailed** *p*-values

The *p*-values are **upper-tail** values — you will only see results for **positive *t*-statistics**

Upper-tail *p* for -1.24 = 0.920



Upper-tail *p* for 1.24 = 0.08

t-contrasts

In **SPM** a *t*-contrast is defined by

- An **L** matrix with a **single row**
- Hypothesis testing using a *t*-statistic
- **One-tailed** *p*-values

The *p*-values are **upper-tail** values — you will only see results for **positive *t*-statistics**

A **positive** *t*-statistic occurs when the **direction of the effect** matches the **direction of the contrast**

$\mathbf{L} = \begin{bmatrix} 1 & 0 \end{bmatrix}$ Only see **positive** effects

$\mathbf{L} = \begin{bmatrix} -1 & 0 \end{bmatrix}$ Only see **negative** effects

One-tailed *p*-values only suitable for **strong directional hypotheses**

F-contrasts

In **SPM** a *F*-contrast is defined by

- An **L** matrix with **multiple rows**
- Hypothesis testing using an *F*-statistic
- **Two-tailed** *p*-values

$$F = \frac{(\mathbf{L}\hat{\boldsymbol{\beta}} - \mathbf{m})'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')^{-1}(\mathbf{L}\hat{\boldsymbol{\beta}} - \mathbf{m})}{r\hat{\sigma}^2}$$

The **numerator** forms a **sum-of-squares** — divided by *r* to form a **mean square**

Multiple rows can be thought of as an **OR** question

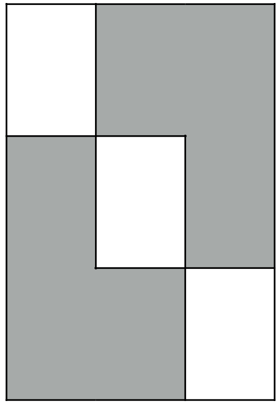
An *F*-contrast with a **single row** is the same as t^2 — in SPM this allows for a **two-tailed** alternative to a *t*-contrast

F-contrasts

Simple example: 1-way ANOVA with a 3-level factor

The **weights** for an F-contrast testing the **main effect** of the factor are:

$$\mathbf{L} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$



Does this seem odd at all?

Should it not be?

$$\mathbf{L} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

F-contrasts

Simple example: 1-way ANOVA with a 3-level factor

Should it not be?

$$\mathbf{L} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ \boxed{1 \quad 0 \quad -1} \end{bmatrix}$$

This row is redundant

This last row is the **sum** of the other rows — value from this row is **not independent of the other rows**

$$\boldsymbol{\beta} = \begin{bmatrix} 10 \\ 15 \\ 12 \end{bmatrix} \quad \mathbf{L}\boldsymbol{\beta} = \begin{bmatrix} -5 \\ 3 \\ -2 \end{bmatrix} \Rightarrow (-5) + 3 = -2$$

The last row **provides no more information** — connection with **numerator degrees of freedom**

Do the weights have to sum to zero?

A potential **source of confusion** relates to wether the **weights** of the contrast **must sum to zero**

Distinction between a **linear combination** and a **contrast**:

- A **contrast** is specifically about **comparing parameters**
- A **linear combination** is more general e.g. averaging, summing etc.

In **SPM** the term **contrast** is used generically for either

In the statistics literature a **contrast** has the **additional requirement** that the weights sum to zero — not necessary for a **linear combination**

Do the weights have to sum to zero?

$$\mathcal{H}_0 : \beta_1 - \beta_2 = 0$$

$$\mathbf{L} = \begin{bmatrix} 1 & -1 \end{bmatrix} \quad \sum l_i = 0 \quad ✔$$

$$\mathbf{L} = \begin{bmatrix} 2 & -2 \end{bmatrix} \quad \sum l_i = 0 \quad ✔$$

$$\mathbf{L} = \begin{bmatrix} 1 & -2 \end{bmatrix} \quad \sum l_i = -1 \quad ✘$$

$$\mathcal{H}_0 : (\beta_1 + \beta_2)/2 - \beta_3 = 0$$

$$\mathbf{L} = \begin{bmatrix} 0.5 & 0.5 & -1 \end{bmatrix} \quad \sum l_i = 0 \quad ✔$$

$$\mathbf{L} = \begin{bmatrix} 1 & 1 & -2 \end{bmatrix} \quad \sum l_i = 0 \quad ✔$$

$$\mathbf{L} = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \quad \sum l_i = 1 \quad ✘$$

In both cases we are looking at **differences** of parameters — these are **true contrasts**

Do the weights have to sum to zero?

The **alternative** is when we look at **individual parameters** or **averages of parameters**

$$\mathcal{H}_0 : \beta_1 = 0$$

$$\mathbf{L} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \sum l_i = 1 \; ✔$$

$$\mathcal{H}_0 : (\beta_1 + \beta_2)/2 = 0$$

$$\mathbf{L} = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix} \sum l_i = 1 \; ✔ \qquad \mathbf{L} = \begin{bmatrix} 1 & 1 \end{bmatrix} \sum l_i = 2 \; ✔$$

If the parameters are **estimable** then these are all valid

Interpretation must be done with **care** — depends on how the model is **parameterised** (e.g. **implicit** vs **explicit** baselines)

Basic contrasts in ANOVA models

Often, it is the **group-level** where our **hypotheses** are focussed

Typically, data collected from **factorial designs** will be analysed using an **AN**alysis **O**f **VA**riance model

For designs with >2 factors, the ANOVA model has a number of effects that we may be interested in — **main effects**, **interactions**, **simple effects**

Important to understand how these are tested using **contrasts**

SPM **defaults** to a **cell means** ANOVA rather than an **overparameterised** ANOVA — no need to worry about **estimability**

Basic contrasts in ANOVA models

2 x 2 ANOVA

Cell means are the **means** from the **intersection** of the factors and the **marginal means** are the **means** from **across** a factor

Factor A

Factor B	1	2	
1	μ_{11}	μ_{21}	$\mu_{.1}$
2	μ_{12}	μ_{22}	$\mu_{.2}$
	$\mu_{1.}$	$\mu_{2.}$	

Cell means

Marginal means for Factor A

Marginal means for Factor B

The **dot** notation indicates a **subscript averaged over**

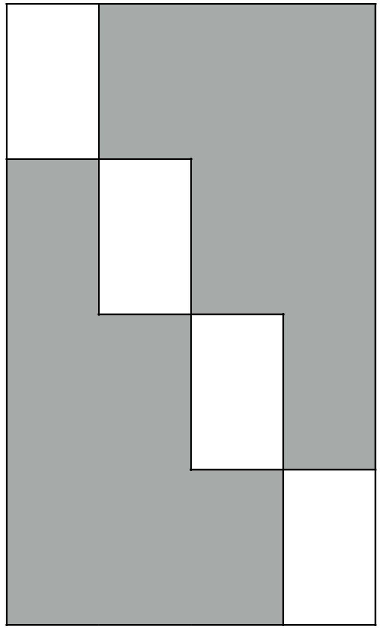
Basic contrasts in ANOVA models

2 x 2 ANOVA

The simplest model is the **cell means** model

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk}$$

Here the **parameters** are the **cell means** of design



The **ANOVA effects** are then formed using **contrasts of the cell means**

Main effect of A $\mathbf{L} = \begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix}$

Main effect of B $\mathbf{L} = \begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}$

A x B interaction $\mathbf{L} = \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}$

The University of Manchester

MANCHESTER 1824

Basic contrasts in ANOVA models

THE UNIVERSITY OF EDINBURGH

2 x 2 ANOVA

The **interaction** contrast

We are looking for a **difference** of **two differences**

$$\mathrm{A} \times \mathrm{B} = \boxed{(\mathrm{A1B1} - \mathrm{A2B1})} - \boxed{(\mathrm{A1B2} - \mathrm{A2B2})}$$

Effect of A at the first level of B

Effect of A at the second level of B

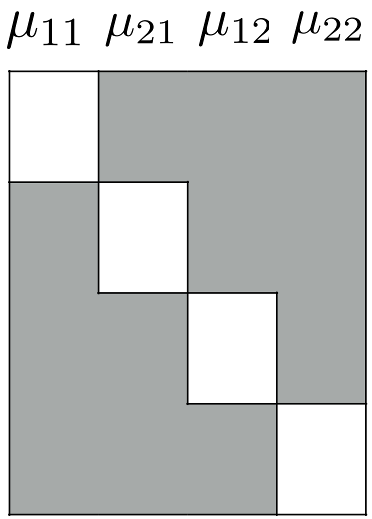
Basic contrasts in ANOVA models

2 x 2 ANOVA

The **interaction** contrast

We are looking for a **difference** of **two differences**

$$\mathrm{A} \times \mathrm{B} = (\mu_{11} - \mu_{21}) - (\mu_{12} - \mu_{22})$$



Basic contrasts in ANOVA models

2 x 2 ANOVA

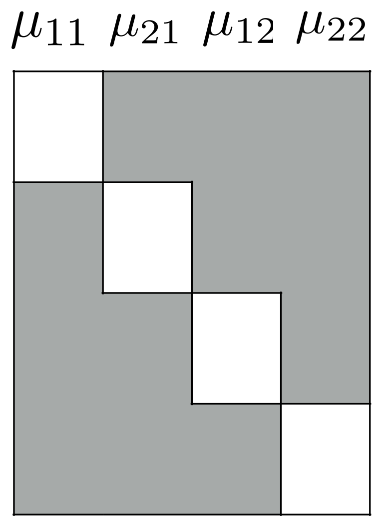
The **interaction** contrast

We are looking for a **difference** of **two differences**

$$\mathrm{A} \times \mathrm{B} = \boxed{(\mu_{11} - \mu_{21})} - \boxed{(\mu_{12} - \mu_{22})}$$

$$\begin{matrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{matrix}$$

$$\mathbf{L} = \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}$$



The **weights** for the **interaction** come from **subtracting** the weights for the **simple effects**

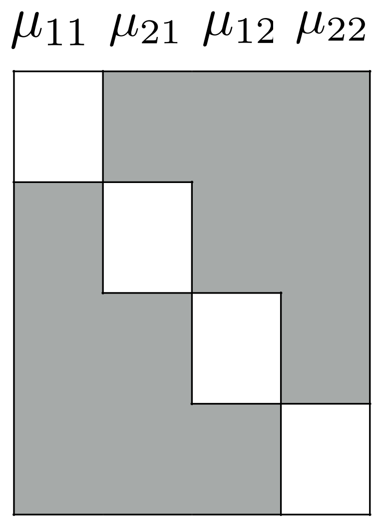
Basic contrasts in ANOVA models

2 x 2 ANOVA

The **interaction** contrast

We are looking for a **difference** of **two differences**

$$\begin{aligned} \mathrm{A} \times \mathrm{B} &= (\mu_{11} - \mu_{21}) - (\mu_{12} - \mu_{22}) \\ &= \mu_{11} - \mu_{21} - \mu_{12} + \mu_{22} \end{aligned}$$



$$\mathbf{L} = \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}$$

Alternatively, we can **remove the brackets** from the expression for the interaction

Basic contrasts in ANOVA models

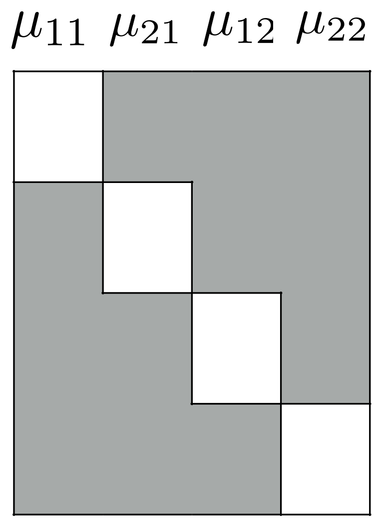
2 x 2 ANOVA

The **interaction** contrast

We are looking for a **difference** of **two differences**

$$
\begin{aligned}
\mathrm{A} \times \mathrm{B} &= (\mu_{11} - \mu_{21}) - (\mu_{12} - \mu_{22}) \\
&= \boxed{\mu_{11}} \boxed{- \mu_{21}} \boxed{- \mu_{12}} \boxed{+ \mu_{22}}
\end{aligned}
$$

$$
\mathbf{L} = \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix}
$$



Alternatively, we can **remove the brackets** from the expression for the interaction

MANCHESTER 1824

The University of Manchester

THE UNIVERSITY OF EDINBURGH

Advanced contrasts in ANOVA models

Unbalanced designs

In other statistical packages, we have a choice of **sums-of-squares** when dealing with **unbalanced** ANOVA designs with >2 factors — **Type I**, **Type II** and **Type III**

The **equally weighted means** contrasts we use in **SPM** are equivalent to **Type III sums-of-squares**

There is debate about how **sensible** the **Type III** tests are — some authors suggest we should be using **Type II** instead (e.g. Venables, 1998; Langsrud, 2003; Fox, 2008; Fox and Weisberg, 2011)

The University of Manchester

MANCHESTER 1824

Advanced contrasts in ANOVA models

Type I

Sequential sums of squares where each effect is tested after adding to the model

$$y_{ijk} = \mu + \epsilon_{ijk}$$
$$y_{ijk} = \mu + \alpha_i + \epsilon_{ijk}$$

Type II

Respects the **principle of marginality** where **main effects** are tested **assuming any interaction is 0**

$$y_{ijk} = \mu + \beta_j + \epsilon_{ijk}$$
$$y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$$

Type III

Ignores the **principle of marginality** so **main effects** are tested **after correcting for interactions**

$$y_{ijk} = \mu + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$
$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

Advanced contrasts in ANOVA models

Implementation using contrasts

Both **Type I** and **Type II** are tricky to calculate **weights** for — no facility in **SPM** to do this for us

$$
\begin{aligned}
\mathbf{L}_{\mathrm{I}} &= \begin{bmatrix} 0 & 1 & -1 & -0.167 & 0.167 & 0.5 & 0.5 & -0.667 & 0.333 \end{bmatrix} \\
\mathbf{L}_{\mathrm{II}} &= \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0.6 & 0.4 & -0.6 & -0.4 \end{bmatrix} \\
\mathbf{L}_{\mathrm{III}} &= \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0.5 & 0.5 & -0.5 & -0.5 \end{bmatrix}
\end{aligned}
$$

Not particularly **intuitive** when written down — both **Type I** and **Type II** weights depend on the **number of subjects in each cell**

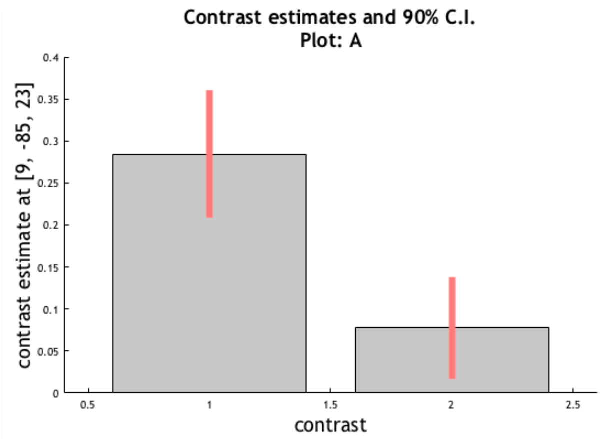
See McFarquhar (2016) for **MATLAB code** and **more detailed discussion** on these approaches

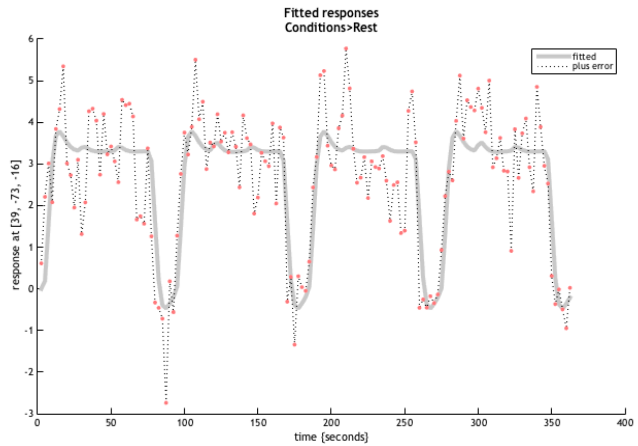
Other uses of contrasts

Contrasts are **generic methods** of selecting **linear combinations** of the **model parameters**

Contrasts are also used in SPM for

- Selecting parameters to **plot** in **bar charts**
- Partitioning the model to **plot** only **effects of interest**





Understanding how this works is **crucial** to getting the most out of the SPM plotting facilities

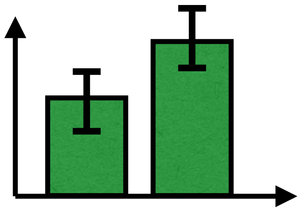
Plotting results with contrasts

Rules for making plots using contrasts

1. Each **row** will be a **separate bar** in the plot
2. **Height** of the bar is the **combination** of parameters **in that row**

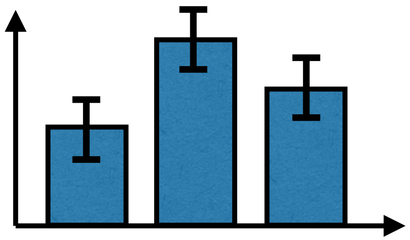
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

Bar 1: value of parameter 1
Bar 2: average of parameters 2 and 3



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

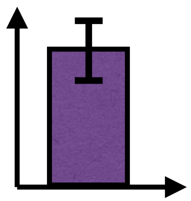
Bar 1: value of parameter 1
Bar 2: value of parameter 2
Bar 3: value of parameter 3



```
>> eye(3)
ans =
   1  0  0
   0  1  0
   0  0  1
```

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

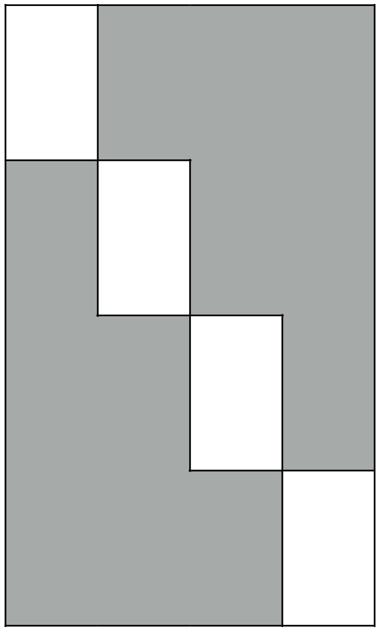
Bar 1: difference between parameters 1 and 2



Plotting results with contrasts

2 x 2 ANOVA

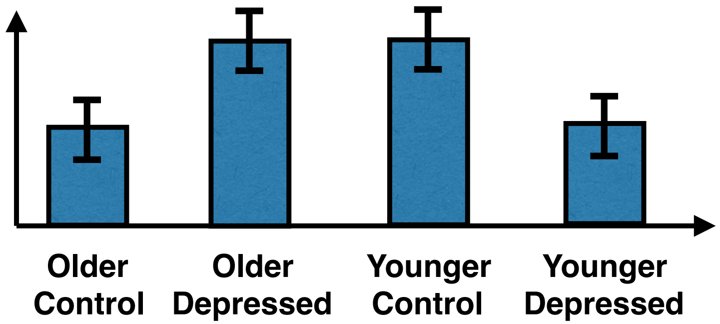
Age (older/younger) x **Diagnosis** (control/depressed)



To plot the **interaction** we want to select **each cell mean**

$$\text{Plot A} \times \text{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

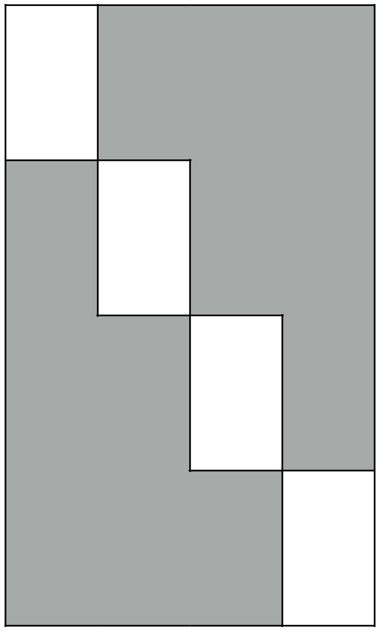
4 bars for the **cell means**



Plotting results with contrasts

2 x 2 ANOVA

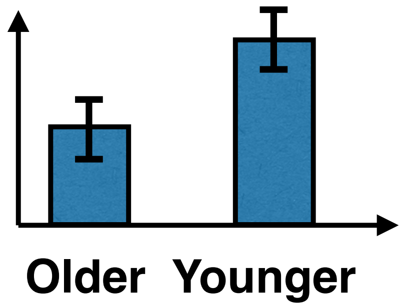
Age (older/younger) x **Diagnosis** (control/depressed)



To plot the **main effects** we average over the cells of the **other factor** to form the **marginal means**

$$\text{Plot Age} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

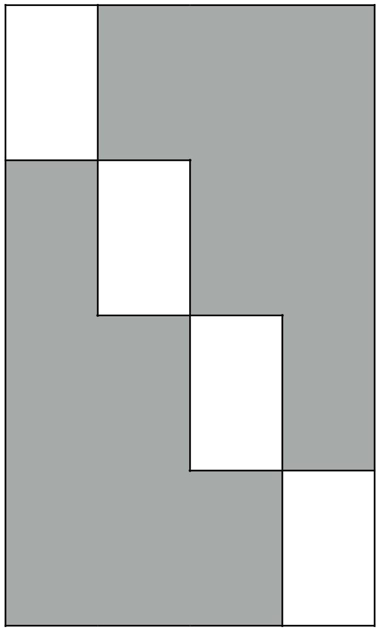
2 bars each **averaged across Diagnosis**



Plotting results with contrasts

2 x 2 ANOVA

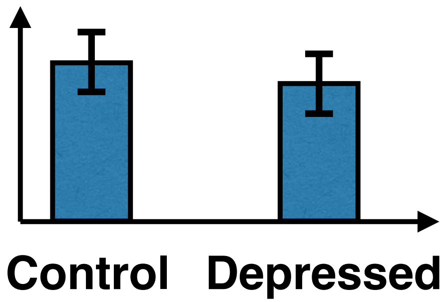
Age (older/younger) x **Diagnosis** (control/depressed)



To plot the **main effects** we average over the cells of the **other factor** to form the **marginal means**

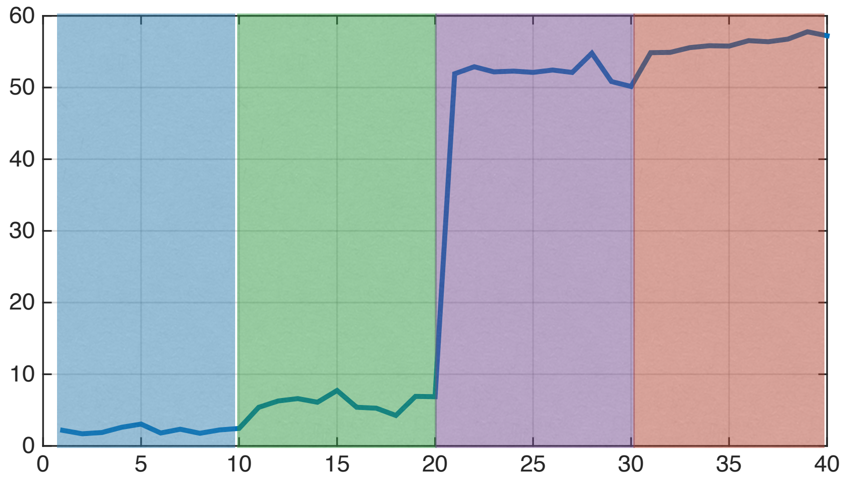
$$\text{Plot Diagnosis} = \begin{bmatrix} 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \end{bmatrix}$$

2 bars each **averaged across Age**



Partitioning effects with contrasts

When attempting to **visualise** effects in our model, sometimes it can be difficult to see what is going on because the **effect of interest** may be **obscured** by another **effect/confound**



A1B1

A1B2

A2B1

A2B2

Clear **main effect of A** — is there a **main effect of B**?

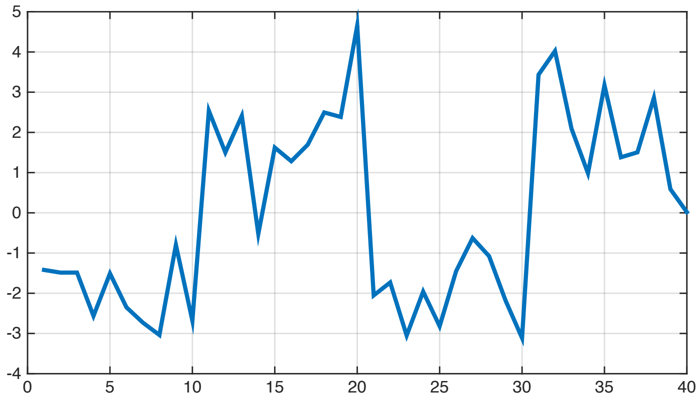
At present, any effect of B is **obscured** by the effect of A

Partitioning effects with contrasts

We can use the **contrast** for the **effect of B** to **remove** from the data those effects that are **not related** to the **effect of B**

$$\mathbf{L} = \begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}$$
$$\mathbf{L}_0 = \mathbf{I}_4 - \mathbf{L}'(\mathbf{L}')^+$$
$$\mathbf{X}_0 = \mathbf{X}\mathbf{L}_0$$
$$\mathbf{R}_0 = \mathbf{I}_n - \mathbf{X}_0\mathbf{X}_0^+$$
$$\mathbf{Y}_{\text{adj}} = \mathbf{R}_0\mathbf{Y}$$

We use the **contrast** to **partition the model** into **effect of interest** and **effects of no interest** and use this to **remove** the **effects of no interest** from the data



This is the **model fit** for the partition of the **effects of interest** + error

Partitioning effects with contrasts

We can use the **contrast** for the **effect of B** to **remove** from the data those effects that are **not related** to the **effect of B**

$$\mathbf{L} = \begin{bmatrix}1 & -1 & 1 & -1\end{bmatrix}$$

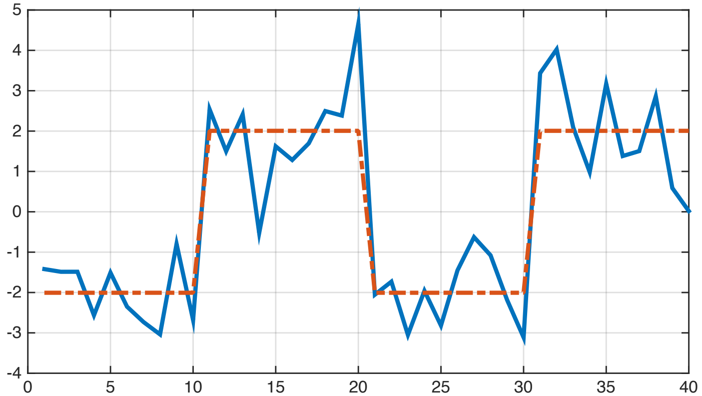
$$\mathbf{L}_0 = \mathbf{I}_4 - \mathbf{L}'(\mathbf{L}')^+$$

$$\mathbf{X}_0 = \mathbf{X}\mathbf{L}_0$$

$$\mathbf{R}_0 = \mathbf{I}_n - \mathbf{X}_0\mathbf{X}_0^+$$

$$\mathbf{Y}_{\text{adj}} = \mathbf{R}_0\mathbf{Y}$$

We use the **contrast** to **partition the model** into **effect of interest** and **effects of no interest** and use this to **remove** the **effects of no interest** from the data



This is the **model fit** for the partition of the **effects of interest** + error

Subtracting the **residuals** gives us the **model fit**

Partitioning effects with contrasts

We can use the **contrast** for the **effect of B** to **remove** from the data those effects that are **not related** to the **effect of B**

$$\mathbf{L} = \begin{bmatrix}1 & -1 & 1 & -1\end{bmatrix}$$

$$\mathbf{L}_0 = \mathbf{I}_4 - \mathbf{L}'(\mathbf{L}')^+$$

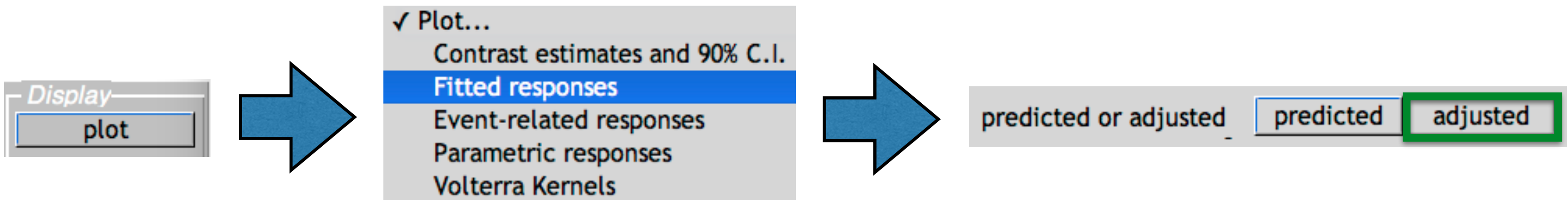
$$\mathbf{X}_0 = \mathbf{X}\mathbf{L}_0$$

$$\mathbf{R}_0 = \mathbf{I}_n - \mathbf{X}_0\mathbf{X}_0^+$$

$$\mathbf{Y}_{\text{adj}} = \mathbf{R}_0\mathbf{Y}$$

We use the **contrast** to **partition the model** into **effect of interest** and **effects of no interest** and use this to **remove** the **effects of no interest** from the data

This is what **SPM** does when you plot **fitted responses**



The University of Manchester

MANCHESTER 1824

Summary

The use of **contrast weights** allows us to defining **linear combinations** of the parameters estimates

Typically this is used to define a **hypothesis** about the **parameters** — important to understand **what** the parameters **mean**

Contrasts must respect the **meaningful** combinations of parameters **dictated** by the rows of **X** — **only** an issue with **rank-deficient** designs

Contrasts are **flexible** as they can

- Provide alternative means of testing hypotheses in **unbalanced** designs
- Select parameters for **plotting**
- Partition the design into **effects of interest/no interest**

References

McFarquhar, M. (2016) Testable Hypotheses for Unbalanced Neuroimaging Data. *Frontiers in Neuroscience: Brain Imaging Methods*, 10, 270.

Poline, J-B., Kherif, F., Pallier, C. & Penny, W. (2007) Contrasts and classical inference. In Friston, K.J. *et al.* (Eds.) *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. London: Academic Press.

Poline, J-B., Brett, M. (2012). The general linear model and fMRI: does love last forever? *NeuroImage*, 62, 871-80.