

The Multivariate General Linear Model

Dr. Martyn McFarquhar

Division of Neuroscience & Experimental Psychology
The University of Manchester

Biennial Edinburgh SPM course, 2019

Contents

1. Review of the univariate GLM
2. The multivariate GLM
3. Hypothesis testing
4. Multiple comparison correction
5. MRM
6. Multi-modal imaging data
7. Discriminant functions analysis
8. Limitations



Review of the univariate GLM

Recall that the **univariate** GLM is expressed as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

The classical **distributional** assumptions are that

$$\mathbf{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$$

The two ramifications of this structure are:

1. **Variance homogeneity** across elements of \mathbf{Y}
2. **No correlation** between elements of \mathbf{Y}

$$\begin{bmatrix} \sigma^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

Review of the univariate GLM

One key type of data that this structure **precludes** is **repeated measurements**

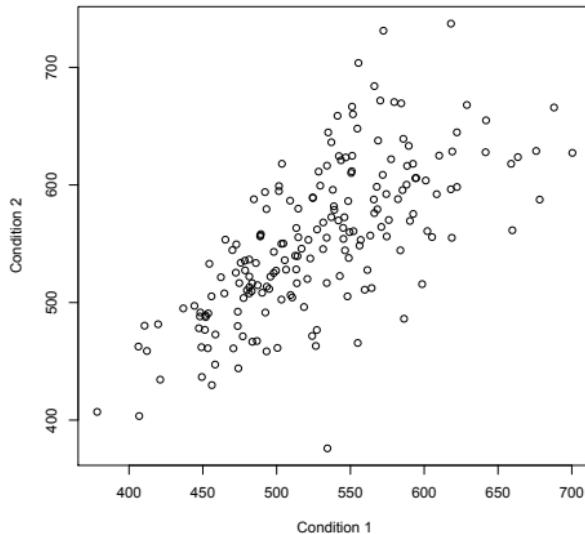
Repeated measurements come in many flavours for fMRI:

- Single-subject:
 - Time-series analysis
 - Multiple voxels
- Group-level
 - Multiple experimental conditions
 - Longitudinal studies
 - Multiple modalities

The **key element** of repeated measurements is that values measured from the **same subject** are likely to be **correlated**

Review of the univariate GLM

Non-imaging example: reaction time



Subjects who are **slower** in **Condition 1** tend to be **slower** in **Condition 2**

The **relative** difference between the conditions can be consistent, even if the **absolute measurements** differ from subject to subject



Review of the univariate GLM

Why does the correlation matter?

For a **hypothesis test** it is important to identify the **correct variance** for the **denominator** of the test statistic:

- Too **small** and we will artificially **inflate** the test statistic
- Too **big** and we will artificially **deflate** the test statistic

For **two random variables** the variance of their **difference** is

$$\mathbf{Var}(X_1 - X_2) = \mathbf{Var}(X_1) + \mathbf{Var}(X_2) - 2\mathbf{Cov}(X_1, X_2)$$

If we think of comparing **parameter estimates** for **repeated measurements** then we need to accommodate the **covariance** (correlation) in order to get this right



Review of the univariate GLM

Accommodating correlation in hypothesis tests

$$\mathbf{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}) \quad t = \frac{\mathbf{L}\hat{\boldsymbol{\beta}}}{\sqrt{\text{Var}(\mathbf{L}\hat{\boldsymbol{\beta}})}} = \frac{\mathbf{L}\hat{\boldsymbol{\beta}}}{\sqrt{\mathbf{L}\text{Var}(\hat{\boldsymbol{\beta}})\mathbf{L}'}}$$

The assumed **covariance structure** of the **data** feeds **directly** into the **denominator** of the test statistic

$$\begin{array}{c} (\mathbf{X}\boxed{\mathbf{V}}^{-1}\mathbf{X})^{-1} \\ \downarrow \\ \sigma^2\mathbf{I} \end{array}$$

If we don't accommodate **correlation** in the **covariance structure** then the **magnitude of the test statistics** will be incorrect



Review of the univariate GLM

Accommodating correlation in the GLM

Have already seen one approach to this — **estimate** the covariance structure and then **whiten** the data

Estimation of the structure often requires **iterative algorithms**

- **Slow** if applied to **each voxel** — algorithms can **fail to converge**
- **Pool voxels** and estimate **one structure** — big assumption
- For **non-marginal models** we need to specify the **random-effects structure** — can be difficult in practise

For **group level** data there is a **more straightforward approach** — the **multivariate GLM**

The multivariate GLM

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ 1 & x_{31} & \cdots & x_{3k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Univariate means that we have a data **vector** — values from a **single variable**

Expanding **Y** to a **data matrix** gives the **multivariate** GLM

The multivariate GLM

$$\begin{bmatrix} y_{11} & y_{21} \\ y_{12} & y_{22} \\ y_{13} & y_{32} \\ \vdots & \vdots \\ y_{1n} & y_{2n} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ 1 & x_{31} & \cdots & x_{3k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_{10} & \beta_{20} \\ \beta_{11} & \beta_{21} \\ \vdots & \vdots \\ \beta_{1k} & \beta_{2k} \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} & \varepsilon_{21} \\ \varepsilon_{12} & \varepsilon_{22} \\ \varepsilon_{13} & \varepsilon_{23} \\ \vdots & \vdots \\ \varepsilon_{1n} & \varepsilon_{2n} \end{bmatrix}$$

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E}$$

Univariate means that we have a data **vector** — values from a **single variable**

Expanding **Y** to a **data matrix** gives the **multivariate** GLM



The multivariate GLM

$$\begin{bmatrix} y_{11} & y_{21} \\ y_{12} & y_{22} \\ y_{13} & y_{32} \\ \vdots & \vdots \\ y_{1n} & y_{2n} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ 1 & x_{31} & \cdots & x_{3k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_{10} & \beta_{20} \\ \beta_{11} & \beta_{21} \\ \vdots & \vdots \\ \beta_{1k} & \beta_{2k} \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} & \varepsilon_{21} \\ \varepsilon_{12} & \varepsilon_{22} \\ \varepsilon_{13} & \varepsilon_{23} \\ \vdots & \vdots \\ \varepsilon_{1n} & \varepsilon_{2n} \end{bmatrix}$$

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E}$$

The **columns** of **Y** give the **repeated measurements** and the **columns** of **B** and **E** give their **parameters** and **errors**

Applying this at **each voxel** gives us a **mass-multivariate** implementation



The multivariate GLM

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E}$$

The important bit for **repeated measurements** is the **distributional assumptions** about the **rows** of \mathbf{Y}

$$\mathbf{Y}_i \sim \mathcal{N}(\mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma})$$

The columns of \mathbf{Y} are allowed to be **correlated** and have **unequal variance**

Completely **unconstrained** — no need to make assumptions about **structure**

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1t} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1t} & \sigma_{2t} & \dots & \sigma_t^2 \end{bmatrix}$$



The multivariate GLM

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E}$$

Furthermore, **estimation** of this model can be done using **ordinary least-squares**

$$\begin{aligned}\hat{\mathbf{B}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} & \hat{\Sigma} &= \frac{1}{n - k}\mathbf{E}'\mathbf{E} \\ \hat{\mathbf{E}} &= \mathbf{Y} - \mathbf{XB}\end{aligned}$$

Computationally, this is **very quick** — no need for **iterative algorithms**

For imaging this is **hugely advantageous** because we can estimate a **unique** and **unconstrained** covariance structure **at every voxel**



Hypothesis testing

We know already that one **general approach** to hypothesis testing in the GLM is to construct a **sum-of-squares**

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

We form an *F*-statistic using

$$F = \frac{Q}{r\sigma^2} = \frac{Q}{r} \div \sigma^2 = \frac{Q}{r} \div \frac{\epsilon'\epsilon}{n - k}$$

Mean square for
the hypothesis

Mean square for
the error



Hypothesis testing

We can modify this for use in **multivariate models** by defining **two** matrices of **contrast weights**

$$\mathbf{Q}_H = (\mathbf{L}\hat{\mathbf{B}}\mathbf{M}')'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')^{-1}(\mathbf{L}\hat{\mathbf{B}}\mathbf{M}')$$

\mathbf{Q}_H is also known as a **sums-of-squares and crossproducts** (SSCP) matrix

We can then do something similar with the **error**

$$\mathbf{Q}_E = \mathbf{M}(\mathbf{E}'\mathbf{E})\mathbf{M}'$$

We can see how the **covariance** estimate is accounted for in the test statistic

$$\hat{\Sigma} = \frac{1}{n - k} \mathbf{E}'\mathbf{E}$$



Hypothesis testing

Before we see what to do with \mathbf{Q}_H and \mathbf{Q}_E let's explore the **contrast matrices**

$$\mathbf{Q}_H = (\mathbf{L}\hat{\mathbf{B}}\mathbf{M}')'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')^{-1}(\mathbf{L}\hat{\mathbf{B}}\mathbf{M}')$$

The **L** matrix acts on the **rows** of **B**

The **M** matrix acts on the **columns** of **B**

L contains weights for the **between-subject** effects

M contains weights for the **within-subject** effects

Note that in **some cases** \mathbf{Q}_H will be a **scalar** or **vector** — as in the univariate GLM



Hypothesis testing

A mixed **between-subject** (2 levels) and **within-subject** (3 levels) design

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \end{bmatrix}$$

Main effect of between

$$\mathbf{L} = [1 \quad -1] \quad \mathbf{M} = [1/3 \quad 1/3 \quad 1/3]$$

Main effect of within

$$\mathbf{L} = [1/2 \quad 1/2] \quad \mathbf{M} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

Between x Within interaction

$$\mathbf{L} = [1 \quad -1] \quad \mathbf{M} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$



Hypothesis testing

Back to \mathbf{Q}_H and \mathbf{Q}_E — we can start thinking about **hypothesis testing** by taking the usual approach of **dividing** the **hypothesis** by the **error**

$$\mathbf{H} = \mathbf{Q}_E^{-1} \mathbf{Q}_H$$

As \mathbf{H} could be a **matrix**, we need some means of reducing \mathbf{H} into a **single number**

Where λ_i is the *i*th **eigenvalue** from the **spectral decomposition** of \mathbf{H}



$$\text{Pillai's trace} = \text{trace}([\mathbf{Q}_H + \mathbf{Q}_E]^{-1} \mathbf{Q}_H) = \sum_{i=1}^q \frac{\lambda_i}{1 + \lambda_i}$$

$$\text{Wilks' lambda} = \frac{|\mathbf{Q}_E|}{|\mathbf{Q}_H + \mathbf{Q}_E|} = \prod_{i=1}^q \frac{\lambda_i}{1 + \lambda_i}$$

$$\text{Hotelling-Lawley trace} = \text{trace}(\mathbf{Q}_E^{-1} \mathbf{Q}_H) = \sum_{i=1}^q \lambda_i$$

$$\text{Roy's largest root} = \frac{\lambda^*}{1 + \lambda^*}$$



Hypothesis testing

Although **4 test statistics** seems unnecessary, there is no **objectively best way** of reducing **H** to a single number

Each of the tests have a **transform** to an **approximate F-variate**

e.g. **Pillai's trace** (V)

$$F_V = \left(\frac{2n + s + 1}{2m + s + 1} \right) \left(\frac{V}{s - V} \right)$$

$$F_V \sim \mathcal{F}_{s(2m+s+1), s(2n+s+1)}$$

This allows us to calculate a **p-value** for **thresholding our results**



Hypothesis testing

In terms of their behaviour, **PT** is the **most conservative**, followed by **WL**, **HTL** and **RLR**

As such

- **Pillai's trace** is the **safest** to use
- **Roy's largest root** is the **least safe** — should probably be avoided
- **Wilks' lambda** is a **good balanced** — generalisation of the likelihood ratio test

Although we are now at a point where we can get a **p-value** for **each** voxel, we are still hit with the usual issue of **multiple comparison correction in imaging**



Multiple comparison correction

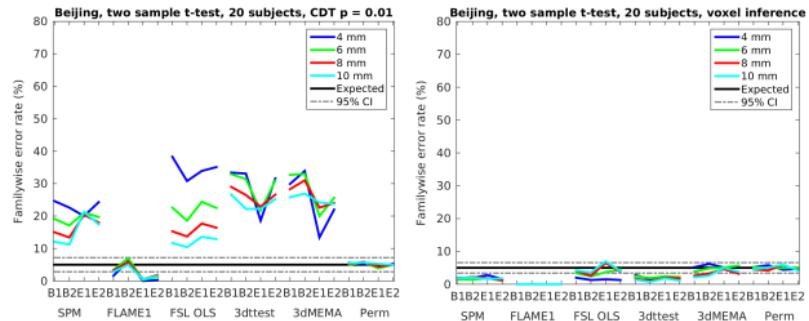
Option 1: Random Field Theory

Solutions for **random fields** of **some** of the multivariate test statistics do exist (e.g. Carbonell, Galan & Worsley, 2008)

Could also work with the **F-approximations** of the test statistics

However, these are **approximations** that give **approximate p-values**

Recent concerns over the use of **random field theory** in certain situations (see Eklund, Nichols & Knutsson, 2016)



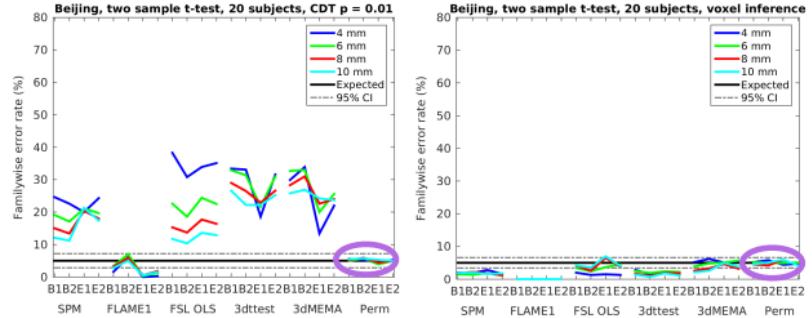
Multiple comparison correction

Option 2: False-discovery rate

We could easily feed the **approximate p-values** into an **FDR** correction — this is controlling a different concept to **FWE**

No means of applying FDR to **clusters** without random-field theory or some other approach

A more general solution is to use
non-parametric permutation tests





Multiple comparison correction

Permutation tests

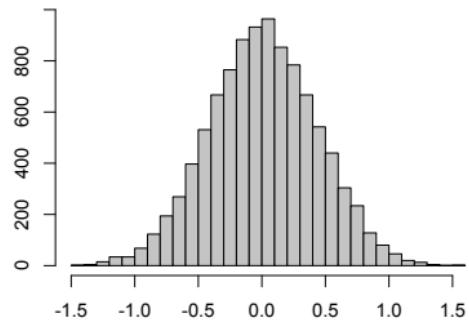
Detail exposition of permutation tests in Winkler *et al.* (2014)

Basic idea is that over many repetitions (e.g. 5,000) we

1. **Randomly re-arrange** the data
2. Fit the **original model** to that data
3. **Re-calculate** the test statistic of interest and **save it**

We build up a **distribution** of the **test statistic** under the null hypothesis that there is **no relationship** between the **data** and the **model**

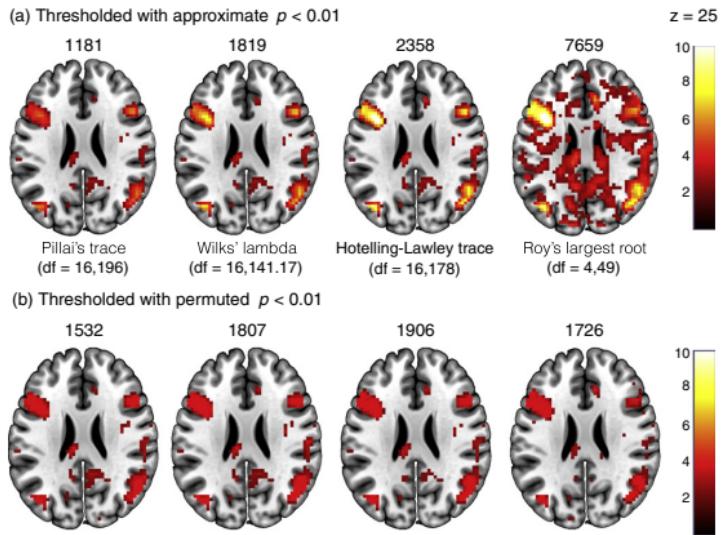
We can then use that distribution to calculate a **non-parametric p-value**



Multiple comparison correction

Permutation tests

This should give us **more accurate** p -values for the 4 statistics

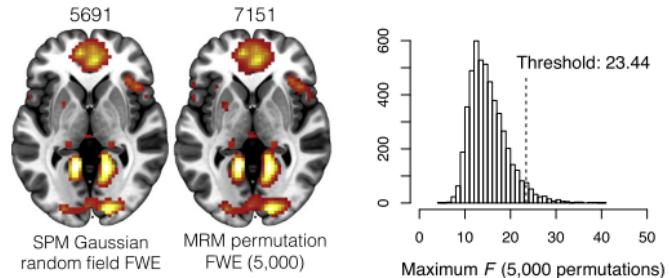


Multiple comparison correction

Permutation tests

For multiple comparison correction, we instead build a distribution of **largest** test statistics **in the whole image**

The *p*-value is then the probability of a test statistic as large, or larger, under the null **anywhere in the image** — controls the **FWE**



This can be expanded to **clusters** by building a distribution of **largest clusters** in the image

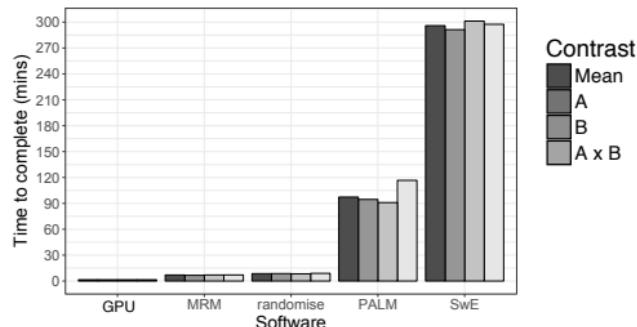
Can also use **non-standard statistics** such as **cluster mass**
(Bullmore *et al.*, 1999)

Multiple comparison correction

Permutation tests

The main challenge is **computational** and not **statistical**

Depends on the **implementation** — both the **language** (MATLAB, C, C++ etc.) and the **code design**



Software	Hours	Mins
SwE	19	46
PALM	6	40
randomise	0	34
MRM	0	27
GPU	0	6

As computer power increases, permutation testing **will get faster**

MRM

[NeuroImage 132 \(2016\) 373–389](#)



Contents lists available at [ScienceDirect](#)

NeuroImage

journal homepage: www.elsevier.com/locate/ynimng



Multivariate and repeated measures (MRM): A new toolbox for dependent and multimodal group-level neuroimaging data



Martyn McFarquhar ^{a,*}, Shane McKie ^a, Richard Emsley ^b, John Suckling ^c, Rebecca Elliott ^a, Stephen Williams ^d

^a Neuroscience & Psychiatry Unit, Stopford Building, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

^b Centre for Biostatistics, Jean MacFarlane Building, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

^c Brain Mapping Unit, Herchel Smith Building for Brain and Mind Sciences, University of Cambridge, Robinson Way, Cambridge CB2 0SZ, UK

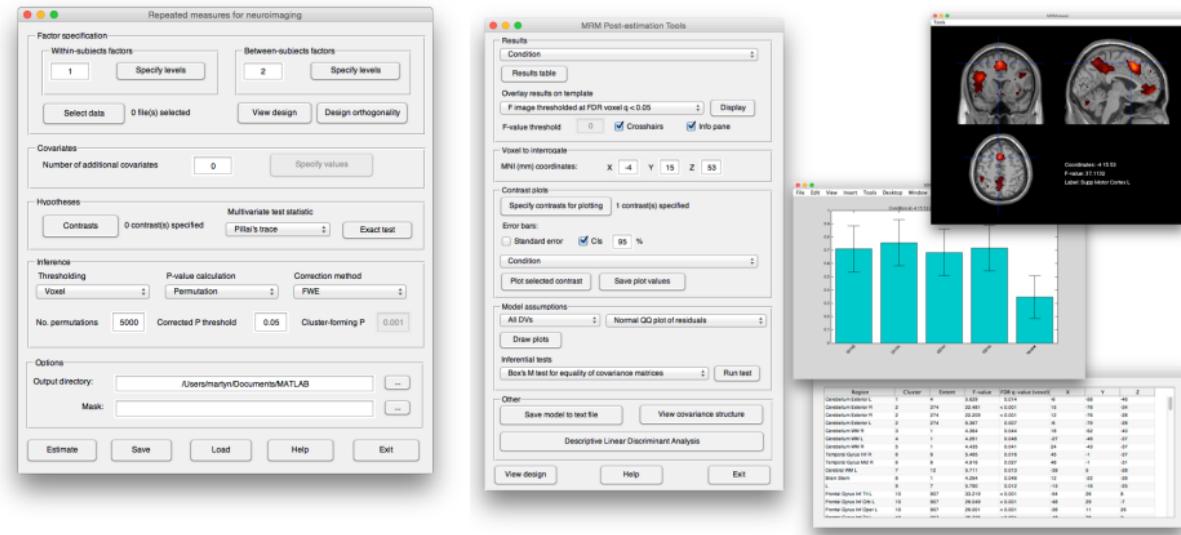
^d Imaging Sciences, Stopford Building, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

<https://github.com/martynmcfarquhar/MRM/releases>

Practical session **this afternoon** using MRM to analyse **repeated measurements** and **multi-modal** imaging data

MRM

MATLAB tool for the specification of **multivariate repeated-measures** and **MANOVA** models of neuroimaging data using **permutation tests**





Multi-modal imaging data

Although we have focussed on **repeated measurements**, the **multivariate GLM** can also be used to analyse **multi-modal** imaging data using a **MANOVA** model

$$\begin{bmatrix} y_{11} & y_{21} \\ y_{12} & y_{22} \\ y_{13} & y_{32} \\ \vdots & \vdots \\ y_{1n} & y_{2n} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ 1 & x_{31} & \cdots & x_{3k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_{10} & \beta_{20} \\ \beta_{11} & \beta_{21} \\ \vdots & \vdots \\ \beta_{1k} & \beta_{2k} \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} & \varepsilon_{21} \\ \varepsilon_{12} & \varepsilon_{22} \\ \varepsilon_{13} & \varepsilon_{23} \\ \vdots & \vdots \\ \varepsilon_{1n} & \varepsilon_{2n} \end{bmatrix}$$

In this instance, the columns of **Y** represent the **different modalities**

The **model form** and **estimation** remain **identical** to a repeated measurements model

The **difference** comes with the **hypothesis tests**

Multi-modal imaging data

Because the **columns** of **Y** may now be on **different scales** the columns of **B** will also be on different scaled

As such, **subtracting** or **averaging** across these columns is **not readily interpretable**

Because of this, in a **MANOVA** the contrast matrix **M** is fixed as an **identity matrix**

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

So a **MANOVA** will only provide **multivariate between-subject** tests



Multi-modal imaging data

How do we **interpret** the results of a **multivariate test**?

Remember that for **multivariate** tests we are assessing a **vector-form** of a **null hypothesis**

Univariate

$$\mathcal{H}_0 : \mu_1 = \mu_2 = \dots = \mu_k = 0$$

Multivariate

$$\mathcal{H}_0 : \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1p} \end{bmatrix} = \begin{bmatrix} \mu_{21} \\ \mu_{22} \\ \vdots \\ \mu_{2p} \end{bmatrix} = \dots = \begin{bmatrix} \mu_{k1} \\ \mu_{k2} \\ \vdots \\ \mu_{kp} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$



Multi-modal imaging data

$$\mathcal{H}_0 : \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1p} \end{bmatrix} = \begin{bmatrix} \mu_{21} \\ \mu_{22} \\ \vdots \\ \mu_{2p} \end{bmatrix} = \dots = \begin{bmatrix} \mu_{k1} \\ \mu_{k2} \\ \vdots \\ \mu_{kp} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

A **significant** multivariate test tells us that **at least one** of these equalities **does not hold**

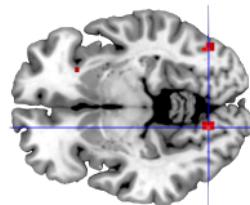
Compared to just doing p **univariate** tests, the multivariate approach:

- Preserves the **Type I error rate**
- Makes use of the **correlation** between the p dependent variables
- Can be **more powerful** than the **univariate** methods
- Can provide more information about the **combination** of dependent variables that **reject** the multivariate null hypothesis



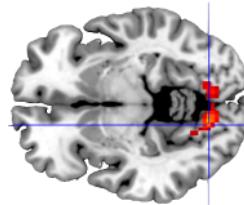
Multi-modal imaging data

For **multi-modal** imaging, we can gain more power to detect **simultaneous** effects when using the **multivariate** approach



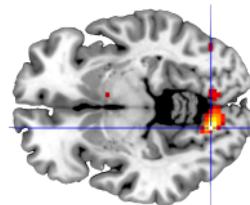
Structural univariate

Coordinates: -15 -73 -7
 F -value: 13.51
 p -value: 0.00108



Functional univariate

Coordinates: -15 -73 -7
 F -value: 15.00
 p -value: 0.00065



Multivariate - structural
+ functional

Coordinates: -15 -73 -7
 F -value: 19.26
 p -value: 0.00001

How we can **explore** a **significant** multivariate effect further?



Discriminant functions analysis

If an **effect** is significant we can say that **at least one** of the **row equalities** does not hold

$$\mathcal{H}_0 : \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1p} \end{bmatrix} = \begin{bmatrix} \mu_{21} \\ \mu_{22} \\ \vdots \\ \mu_{2p} \end{bmatrix} = \dots = \begin{bmatrix} \mu_{k1} \\ \mu_{k2} \\ \vdots \\ \mu_{kp} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

However, it may not hold for **multiple rows** and the **magnitude** of this effect may also differ **across rows**

This may imply doing p **univariate hypothesis tests**, however, there is a **better way** — **descriptive linear discriminant analysis** (dLDA)



Discriminant functions analysis

The aim of dLDA is to convert **each row of \mathbf{Y}** to a **scalar**

$$z_{ij} = a_1 y_{ij1} + a_2 y_{ij2} + \dots + a_p y_{ijp} = \mathbf{a} \mathbf{Y}'_{ij}$$

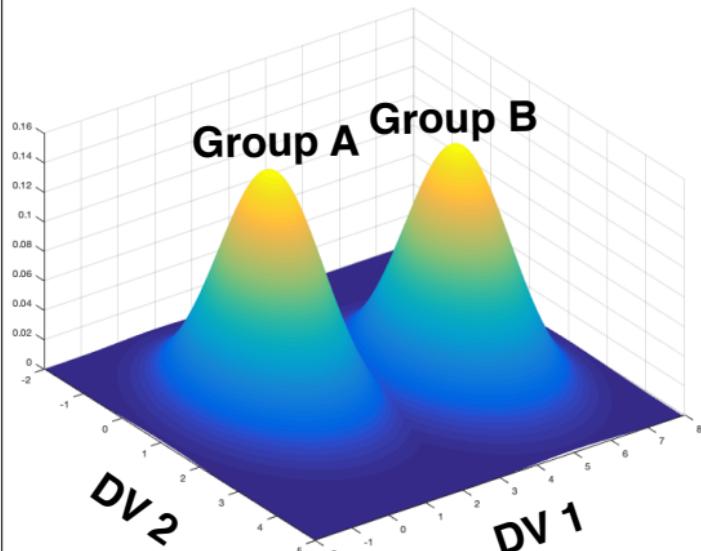
Where \mathbf{Y}_{ij} is the **row** of \mathbf{Y} associated with subject i from group j

The weights in \mathbf{a} form a **discriminant function** and are found so that the **standardised group difference** in z is as **large as possible**

For **two groups** the weights in \mathbf{a} are found to $t(\mathbf{a}) = \frac{z_1 - z_2}{s_z}$
maximise:

The **weights** in \mathbf{a} therefore provide information on the **magnitude** of the **contribution** of each dependent variable to the **multivariate effect of group**

Discriminant functions analysis



The **mean** vector are:

$$\mathbf{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

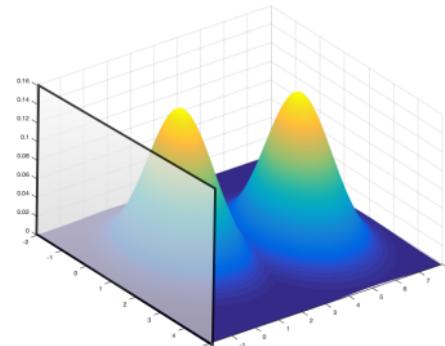
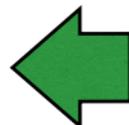
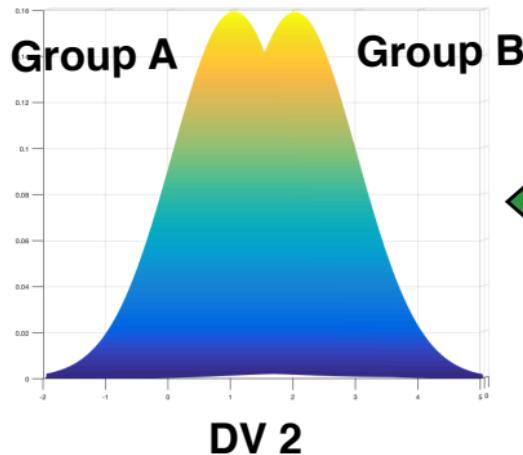
The **multivariate test** is therefore based on

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} -4 \\ -1 \end{bmatrix}$$

There is a group difference for **both DVs** — magnitude is **larger** for DV1

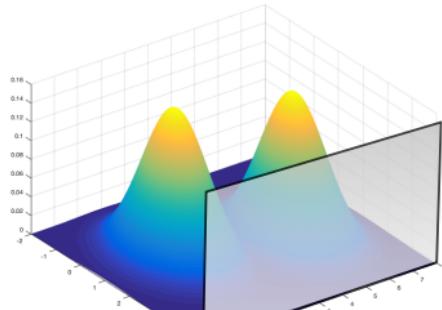
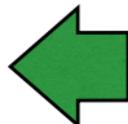
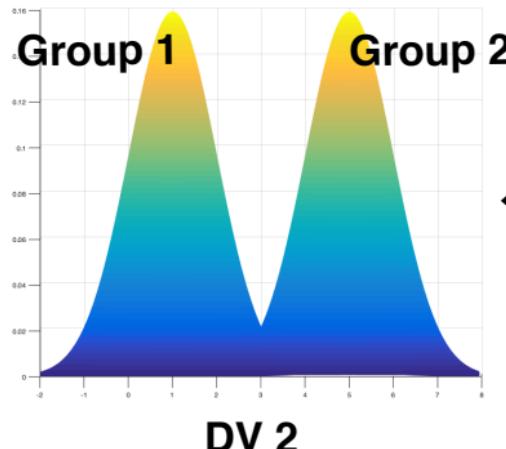
Discriminant functions analysis

Performing **two univariate tests** would be like **projecting** the data onto the **two** main axis separately



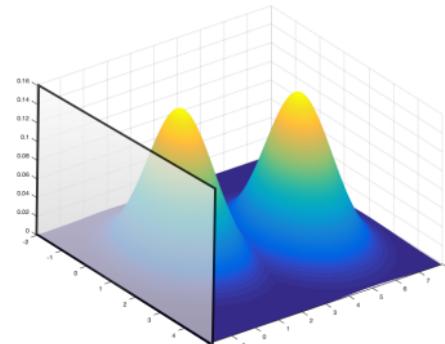
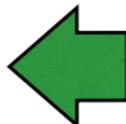
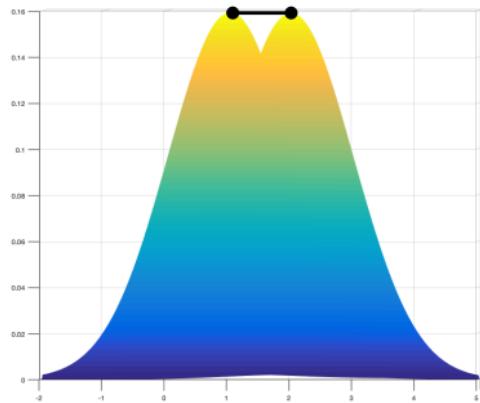
Discriminant functions analysis

Performing **two univariate tests** would be like **projecting** the data onto the **two** main axis separately



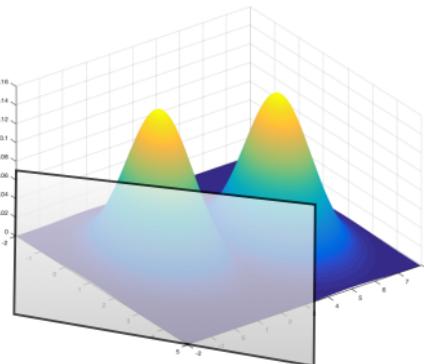
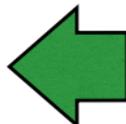
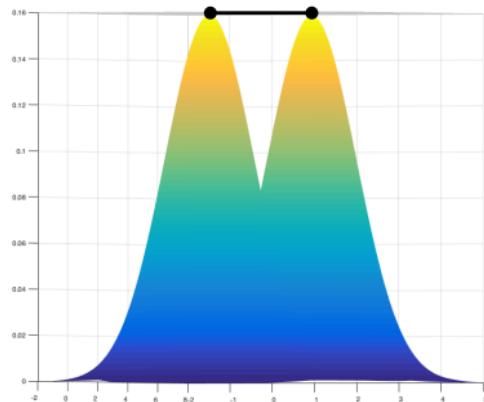
Discriminant functions analysis

What dLDA does is find the axis where the **mean difference** is **maximise**



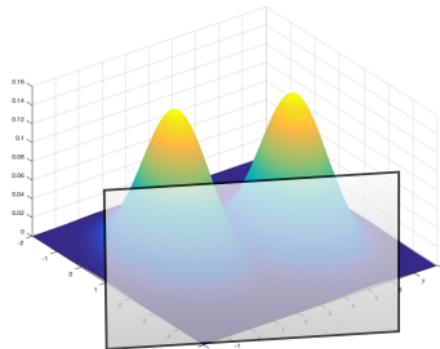
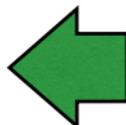
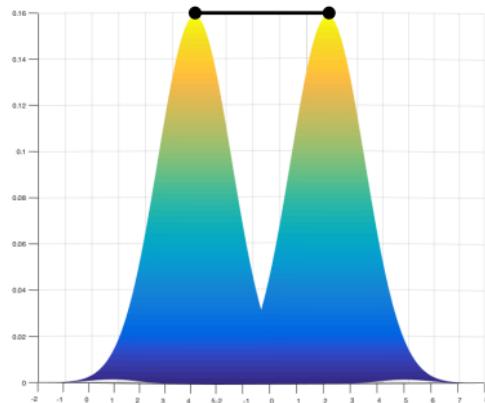
Discriminant functions analysis

What dLDA does is find the axis where the **mean difference** is **maximise**



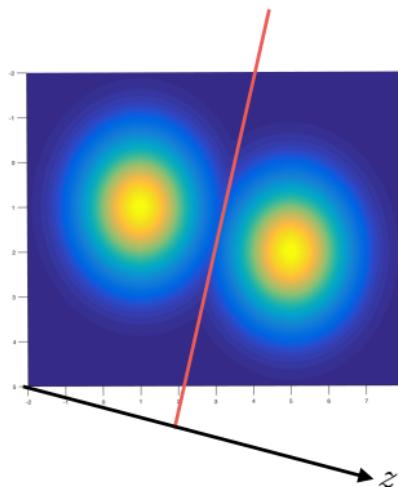
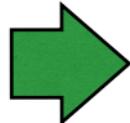
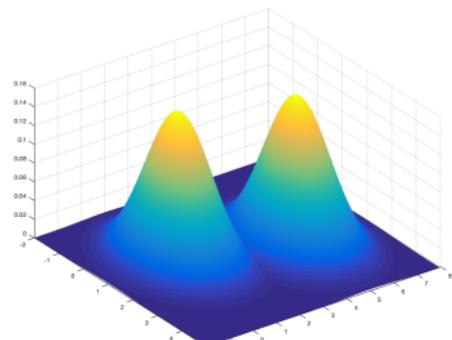
Discriminant functions analysis

What dLDA does is find the axis where the **mean difference** is **maximise**



Discriminant functions analysis

What dLDA does is find the axis where the **mean difference** is **maximise**



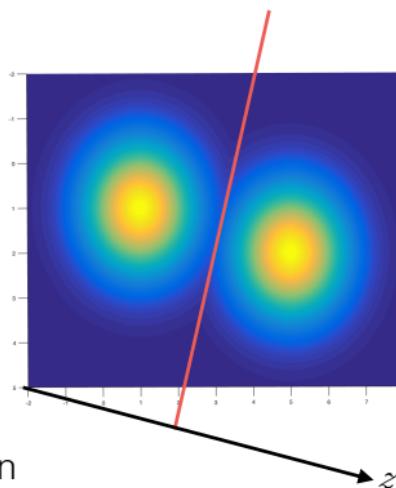
Discriminant functions analysis

What dLDA does is find the axis where the **mean difference** is **maximise**

The **projection** of **Y** to this new axis tells us about the **relative contribution** of the **dependent variables** to the **multivariate group difference**

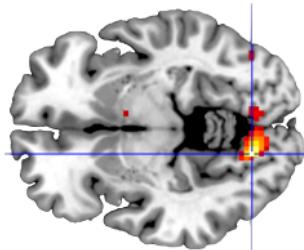
Here, this is **mostly** related to **DV1**, with a **smaller contribution** from **DV2**

This summarises the **multivariate** result in **univariate space**





Discriminant functions analysis



Multivariate - structural
+ functional

Coordinates: -15 -73 -7

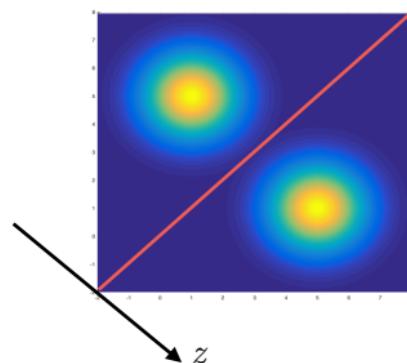
F-value: 19.26

p-value: 0.00001

Performing *d*LDA at the **peak voxel** gives standardised weights of

- **Structural** = 0.826
- **Functional** = 0.850

This suggests that the two modalities are largely contributing **equally** to the **multivariate effect of group**



Limitations

Although the **multivariate GLM** has a lot of potential for imaging, there are some issues:

- Although the **covariance structure** is **unconstrained** it is assumed **identical across groups**
- **Time-varying covariates** cannot be included due to the way the model is structured
- There is no means of including **missing data**, making the approach limited for **longitudinal research**
- Multivariate approaches to **repeated measurements** can be **less powerful** than univariate approaches



Summary

The **multivariate GLM** is a straightforward extension to the familiar **univariate GLM**

This extension allows for **dependent** data with an **unconstrained covariance structure** — estimate uniquely at **each voxel**

Hypothesis testing requires the use of **two** contrast matrices — **L** and **M** — with effective **multiple comparison correction** provided by the use of **permutation tests**

This approach can also be used for both **repeated measurements** and for **multi-modal** imaging datasets, using **MANOVA** models

These models can be followed-up using **dLDA** to investigate the **contribution** of each modality to the **multivariate group effects**

References

- Eklund, A., Nichols, T.E. & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false positive rates, *Proceedings of the National Academy of Sciences*, 113, 7900-05.
- McFarquhar, M., McKie, S., Emsley, R., Suckling, J., Elliott, R. & Williams, S. (2016). Multivariate and repeated measures (MRM): A new toolbox for dependent and multimodal group-level neuroimaging data, *NeuroImage*, 132, 373-89.
- Rencher A.C. & Christensen W.F. (2012) Methods of Multivariate Analysis (3rd ed). John Wiley & Sons, New York.
- Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M. & Nichols, T.E. (2014). Permutation inference for the general linear model, *NeuroImage*, 92, 381-97.