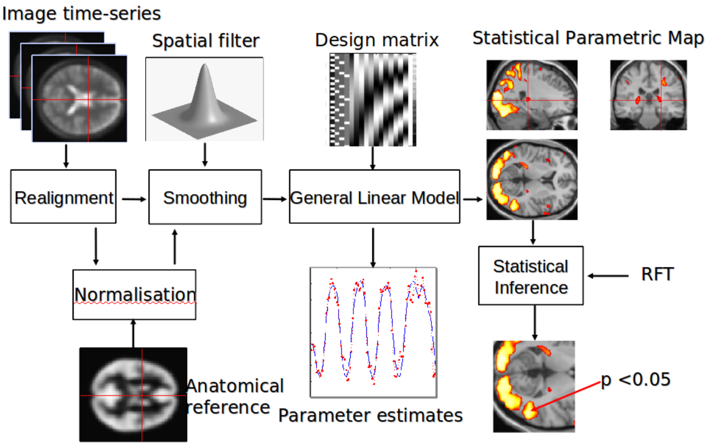
Data Representation & Pre-processing

John Ashburner

Wellcome Trust Centre for Neuroimaging
UCL Institute of Neurlogy
London, UK

1. functional MRI
   - Within subject fMRI: temporal aspects
   - Within subject fMRI: spatial aspects
   - Multiple Subject fMRI
2. Anatomical MRI
3. Advanced issues

SPM approach



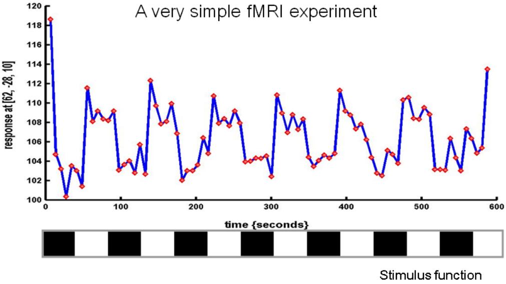
Blood Oxygenation Level Dependent signal

- Brain activity requires active pumping of ions across cell membranes.
- Requires energy from glucose (+ oxygen).
- Greater glucose requirements lead to increased blood flow.
- Leads to higher concentrations of oxyhaemoglobin relative to deoxyhaemoglobin.
- Deoxyhaemoglobin appears darker than oxyhaemoglobin in T_2^* weighted images.

Summary: increased brain activity leads to (very slightly) increased BOLD signal.

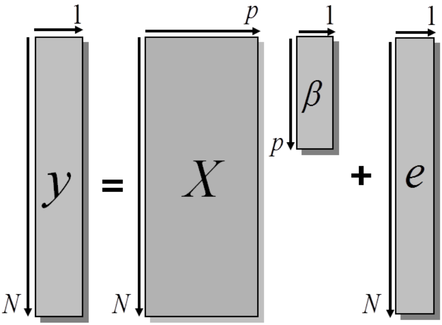
fMRI data

- Several volumes of data are acquired while the subject is stimulated in some way.
- Each volume takes in the region of 2 seconds to acquire.



General linear model

Mass-univariate analysis: voxel-wise GLM



$$y = X\beta + e$$

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

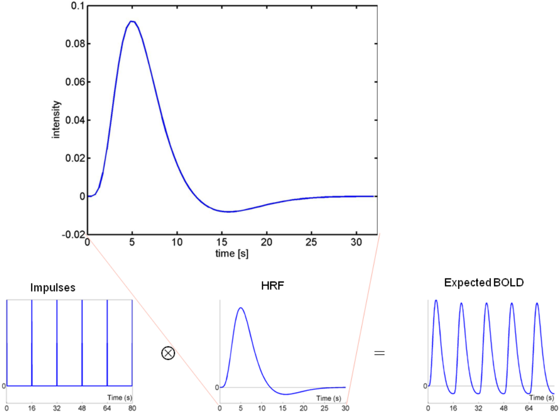
Model is specified by
1. Design matrix X
2. Assumptions about e

N: number of scans
p: number of regressors

The design matrix embodies all available knowledge about experimentally controlled factors and potential confounds.

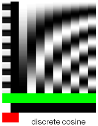
Haemodynamic response function

- BOLD signal is delayed and dispersed relative to stimuli.
- This is normally modelled by convolving our stimuli functions with a "canonical haemodynamic response function".

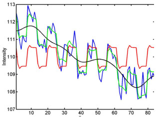


fMRI noise is not i.i.d.

- Low frequency drifts in the baseline intensity.
  - Usually attributed to aliasing effects.
  - Generally regressed out by including low frequency basis functions in a GLM.
- Serial correlations.
  - From normal brain activity, MRI physics etc.
  - This covariance often modelled with a 1st order autoregressive process.

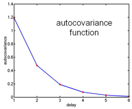


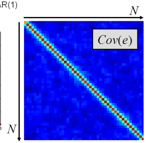
discrete cosine transform (DCT) set



$$e_t = ae_{t-1} + \varepsilon_t \text{ with } \varepsilon_t \sim N(0,\sigma^2)$$

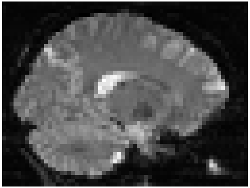
1st order autoregressive process: AR(1)





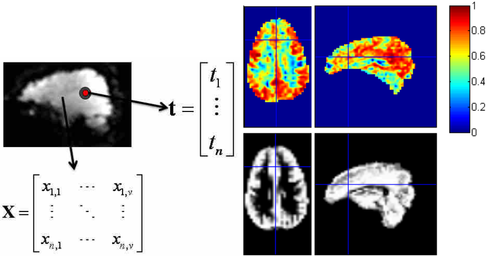
Slice acquisitions

- fMRI usually collected a slice at a time.
- Slice order may be
  - Sequential. eg
    - 1, 2, 3, 4, 5, 6, 7, 8, 9
    - 9, 8, 7, 6, 5, 4, 3, 2, 1
  - Interleaved. eg
    - 1, 3, 5, 7, 9, 2, 4, 6, 8
    - 9, 7, 5, 3, 1, 8, 6, 4, 2
- One approach is to interpolate the data in time.
- Not a perfect solution (aliasing effects).



Spatial covariance

- For brain mapping, covariance among different brain regions is usually ignored.
- Resting-state fMRI shows that there is a great deal of spatial inter-dependence among BOLD signal.

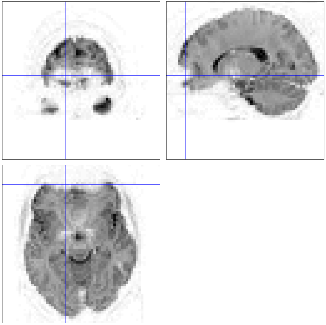


Chu et al. *Measuring the Consistency of Global Functional Connectivity using Kernel Regression Methods.*

IEEE International Workshop on Pattern Recognition in NeuroImaging, 2011.

EPI data

- Most fMRI involves echo-planar imaging (EPI).
- Image resolution is typically around 2 mm isotropic, with volumes of about $64 \times 64 \times 64$ voxels.
- Different magnetic susceptibilities between air and tissue lead to nonuniform magnetic field.
    - Distortions in the phase encode direction.
    - Signal dropout due to de-phasing through slice.



Head motion

- Subject's rarely stay still.
- Head motion by one voxel can make the difference between a voxel containing brain, and containing very little signal.
- This is a 100% signal change – much greater than the 3% or so signal change from BOLD.
- Over hundreds of scans, the tiniest movement can show up as an "activation".
- A partial solution is to realign the volumes, based on optimising the parameters of a rigid-body transform.
- Many artifacts remain in the data.

Motion artifacts

Artifacts that can not be corrected by rigid-body alignment:

- Interpolation artifacts.
- Aliasing from gaps between slices.
- Slices not acquired simultaneously – rapid movement not accounted for.
- MRI artifacts may not move according to the rigid-body model
    - Spatial distortions.
    - Intensity dropout.
    - Nyquist ghosts artifact.
    - etc

Functions of the estimated motion parameters can be modelled as confounding effects in subsequent analyses.

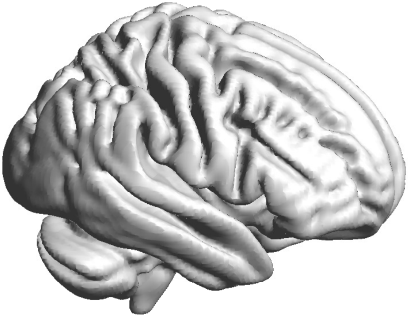
Multi-subject fMRI

Findings are most interesting if they can be generalised to the population.

- Differentiate between intra-subject variance and inter-subject variance.
- Inference about whether findings generalise to the population should be based on inter-subject variance.
- First need to bring homologous functional regions into alignment across subjects.

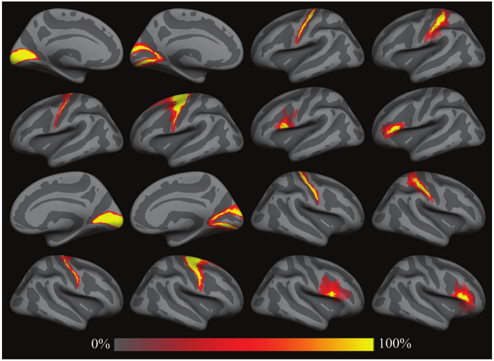
Inter-subject Variability

- Brain shapes and sizes vary greatly from subject to subject.
- Cortical folding patterns vary, although some sulci/gyri are found in all subjects.
- "Spatial normalisation" is used to bring subjects' data into alignment.
- Alignment is not exact, so spatial blurring usually used to compensate.



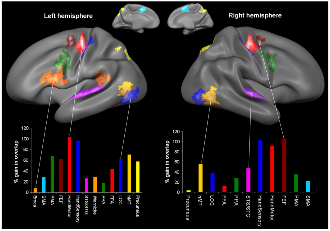
Structure predicts function

Aligning cortical folds leads to good alignment of Brodmann areas.



B Fischl, N Rajendran, E Busa, J Augustinack, O Hinds, B.T.T Yeo, H Mohlberg, K Amunts and K Zilles. Cortical Folding Patterns and Predicting Cytoarchitecture Cerebral Cortex 18:1973–1980 (2008).

Overlap of functional data is increased using nonlinear registration.

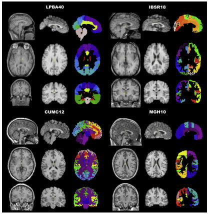


M Frost and R Goebel. Measuring structural-functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. NeuroImage 59(2):1369–1381 (2012).

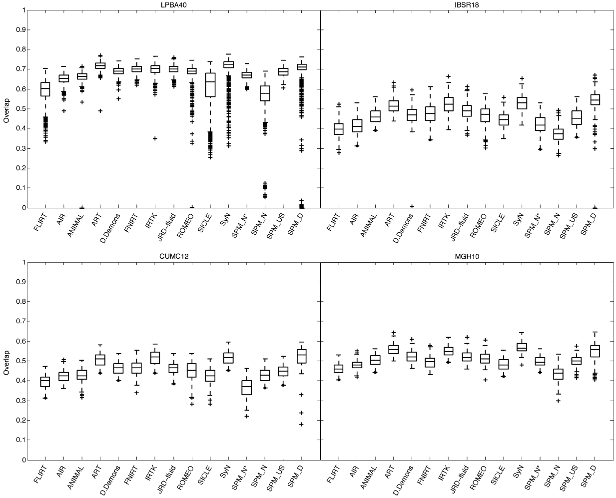
"Spatial Normalisation"

- Involves warping data from subjects to some common reference space.
    - Allows signal in homologous structures to be combined/compared.
    - Locations of "activations" may be reported by their x,y,z coordinates.
- Many possible models for alignment between subjects.
    - Generally involve some form of optimisation.
    - Minimise an objective function consisting of two terms:
        1. Image matching term.
        2. Regularisation term.
- Different alignment models result in different findings.

Volumetric alignment of structure

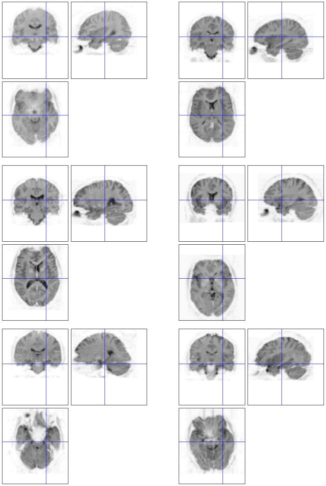


A Klein, J Andersson, BA Ardekani, J Ashburner, B Avants, M-C Chiang, GE Christensen, DL Collins, J Gee, P Hellier, JH Song, M Jenkinson, C Lepage, D Rueckert, P Thompson, T Vercauteren, RP Woods, JJ Mann and RV Parsey. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage 46(3):786–802 (2009).

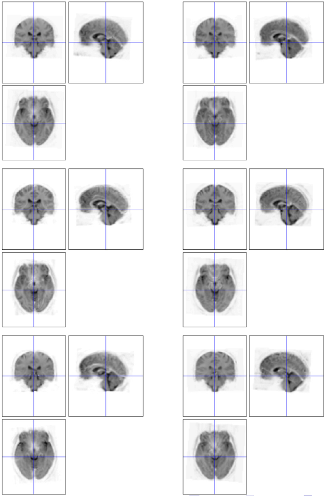


Spatial Normalisation

Original data from six subjects.

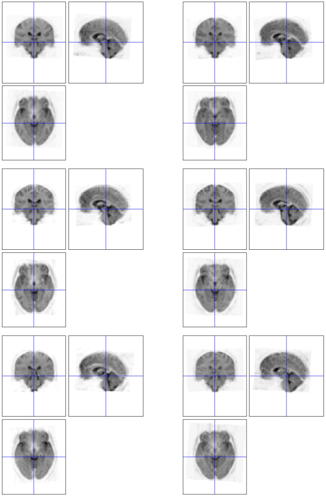


Spatially normalised data.

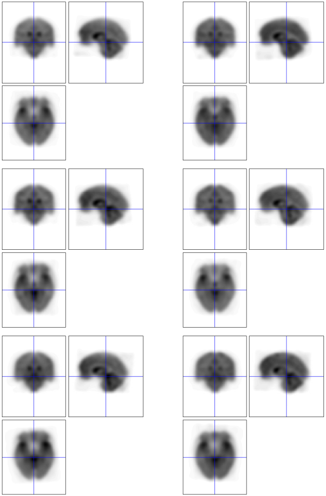


Smoothing

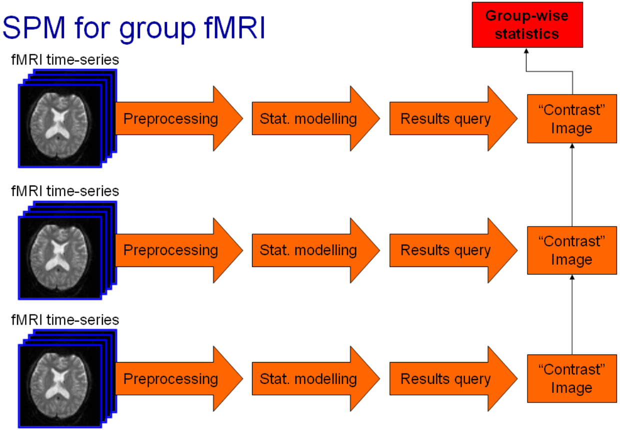
Spatially normalised data.



Smoothed data.



Accounting for inter-subject variance



1 functional MRI

2 Anatomical MRI
- Types of anatomical MRI
- Example data
- Example features

3 Advanced issues

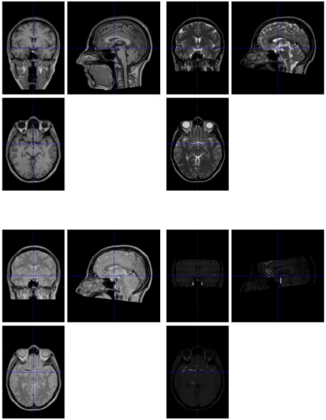
Anatomical MRI

- Most of the variability among subjects (accessible via imaging) is anatomical in nature.
- Many of the fMRI differences we see between populations of subjects could probably be explained anatomically.

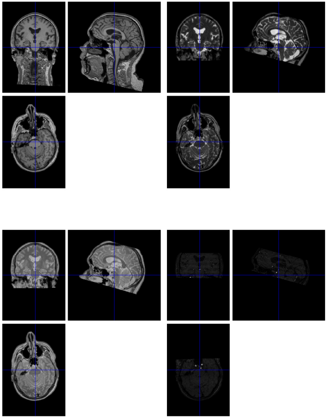
- Most MRI is not quantitative, so signal intensities mean relatively little.
- Characterising anatomical differences involves relating shapes, sizes, lengths etc among populations.

Different types of scan

T1w, T2w, PDw & MRA.

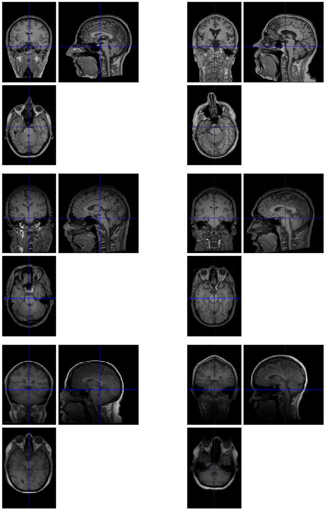


T1w, T2w, PDw & MRA.



Variability among scanners

- Images vary considerably from scanner to scanner, even when they are supposed to be the same.
- Here are T1-weighted scans of six subjects collected on three different scanners.
- More difficult to genaralise findings across scanners.

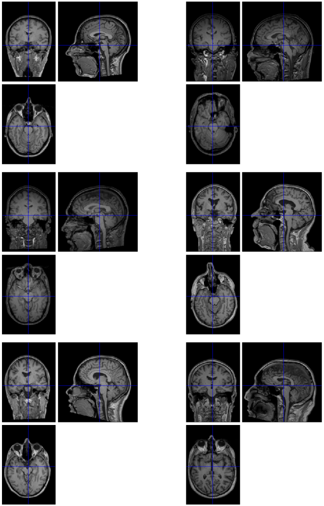


Example Data

Used 550 T1w brain MRI from IXI (**I**nformation e**X**traction from **I**mages) dataset.
`http://www.brain-development.org/`

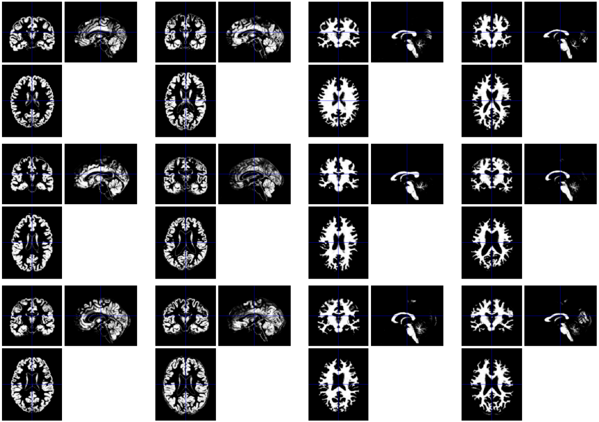
Data from three different hospitals in London:

- Hammersmith Hospital using a Philips 3T system
- Guy's Hospital using a Philips 1.5T system
- Institute of Psychiatry using a GE 1.5T system



Grey and White Matter

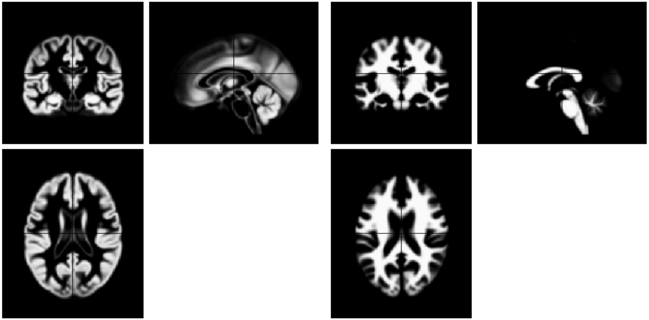
Segmented into GM and WM. Approximately aligned via rigid-body.



Ashburner, J & Friston, KJ. *Unified segmentation*. NeuroImage 26(3):839–851 (2005).

Diffeomorphic Alignment

All GM and WM were diffeomorphically aligned to their common average-shaped template.

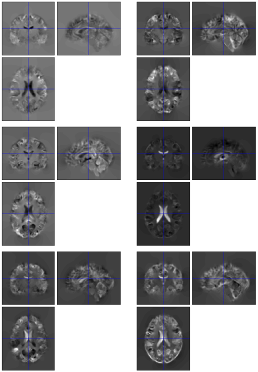


Ashburner, J & Friston, KJ. *Diffeomorphic registration using geodesic shooting and Gauss-Newton optimisation*. NeuroImage 55(3):954–967 (2011).
Ashburner, J & Friston, KJ. *Computing average shaped tissue probability templates*. NeuroImage 45(2):333–341 (2009).

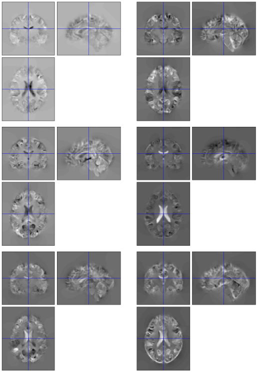
Volumetric Features

A number of features were used for pattern recognition.
Firstly, two features relating to relative volumes.
Initial velocity divergence is similar to logarithms of Jacobian determinants.

Jacobian Determinants

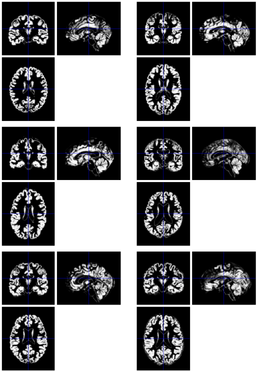


Initial Velocity Divergence

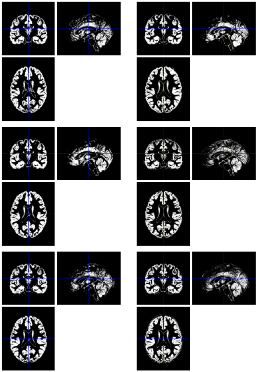


Grey Matter Features

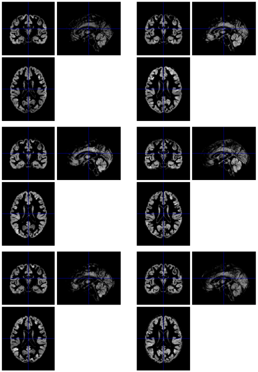
Rigidly Registered GM



Nonlinearly Registered GM



Registered and Jacobian Scaled GM

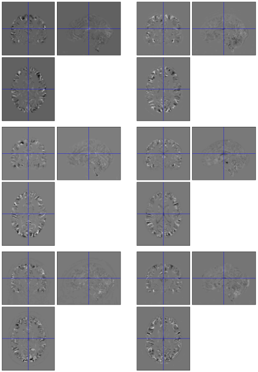


"Scalar Momentum" Features

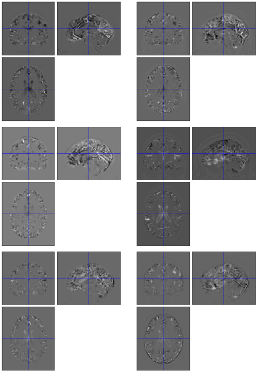
Scalar momentum actually has two components because GM was matched with GM and WM was matched with WM.

Singh, Fletcher, Preston, Ha, King, Marron, Wiener & Joshi (2010). *Multivariate Statistical Analysis of Deformation Momenta Relating Anatomical Shape to Neuropsychological Measures*. T. Jiang et al. (Eds.): MICCAI 2010, Part III, LNCS 6363, pp. 529–537, 2010.

First Momentum Component

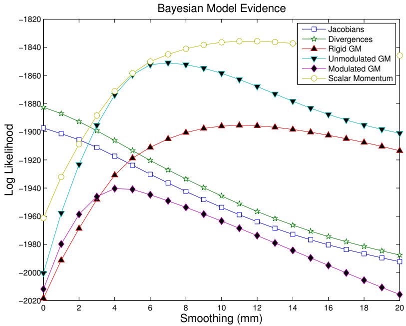


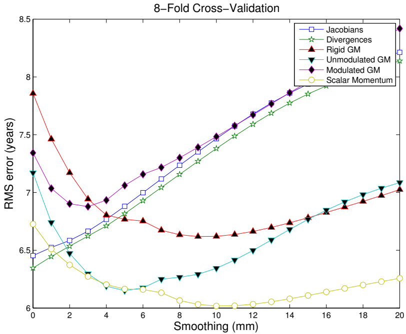
Second Momentum Component



Age Regression

Linear Gaussian Process Regression to predict subject ages.

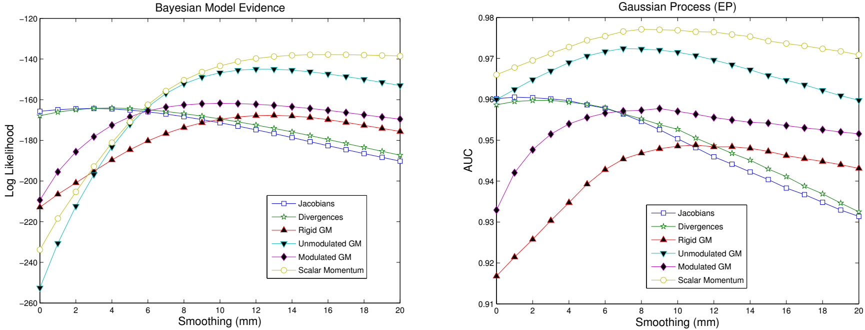




Rasmussen, CE & Williams, CKI. *Gaussian processes for machine learning*. Springer (2006).

Sex Classification

Linear Gaussian Process Classification (EP) to predict sexes.



Rasmussen, CE & Williams, CKI. *Gaussian processes for machine learning*. Springer (2006).

Conclusions

- Scalar momentum (with about 10mm smoothing) appears to be a useful feature set.
- Jacobian-scaled warped GM is surprisingly poor.
- Amount of spatial smoothing makes a big difference.
- Further dependencies on the details of the registration still need exploring.

1 functional MRI

2 Anatomical MRI

3 Advanced issues
- Prior knowledge
- Modelling nonlinearities
- Measuring similarity

No Free Ducklings

No Free Lunch theorem says that learning is impossible without prior knowledge.
http://en.wikipedia.org/wiki/No_free_lunch_in_search_and_optimization

Ugly Duckling theorem says that things are all equivalently similar to each other without prior knowledge.
http://en.wikipedia.org/wiki/Ugly_duckling_theorem

By Ryan Ebert from Portland, US (Flickr) [CC BY 2.0], via Wikimedia Commons. https://creativecommons.org/licenses/by/2.0/

What prior knowledge do we have about the variability among people that can be measured using MRI?
How do we use this knowledge?

Kernel Matrices

Linear kernel matrices are often computed from the raw features:

$$\mathbf{K} = \mathbf{X}\mathbf{X}^T$$

A simple spatial feature selection may be considered as the following, where $\mathbf{\Sigma}_0$ is a (scaled) diagonal matrix of ones and zeros:

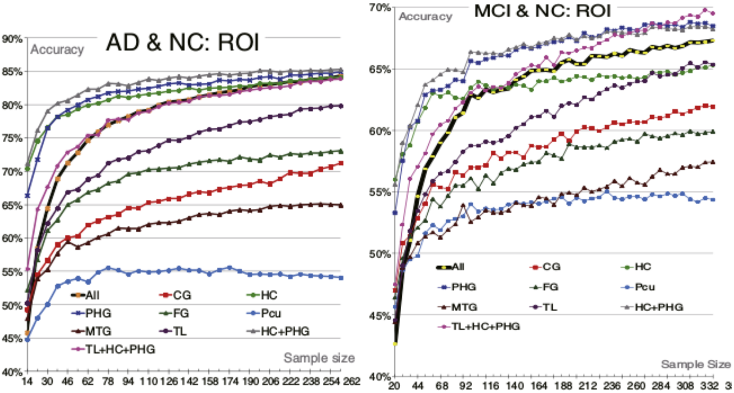
$$\mathbf{K} = \mathbf{X}\mathbf{\Sigma}_0\mathbf{X}^T$$

$\mathbf{\Sigma}_0$ may be more complicated, for example encoding spatial smoothing, high-pass filtering or any number of other things.

Prior knowledge about brain regions involved

- The best way would be to augment the training data with data from previous studies.
- Lack of data-sharing means this is generally not possible, so we need to extract information from publications.
- The neuroimaging literature is mostly blobs.
- These give pointers about how best to weight the data ($\boldsymbol{\Sigma}_0 = diag(\mathbf{s}), s_i \in \mathbb{R}^+$).

Weighting suspected regions more heavily



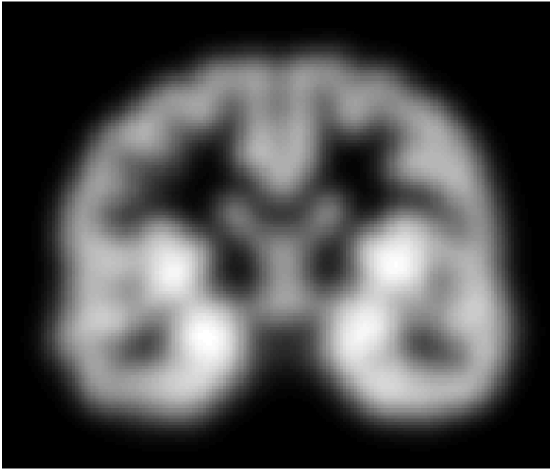
Chu et al. "Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images". NeuroImage 60:59–70 (2012).

Spatial smoothing

If we know that higher frequency signal is more likely to be noise.

$$\mathbf{K} = \mathbf{X}\mathbf{\Sigma}_0\mathbf{X}^T$$

$\mathbf{\Sigma}_0$ no longer diagonal.



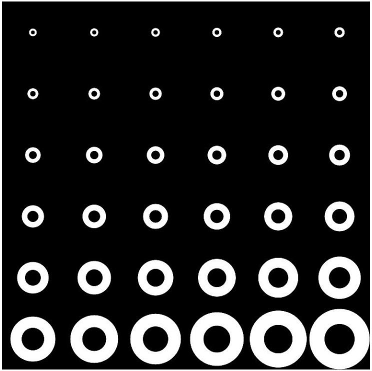
Modelling out nonlinearities

Instead of using nonlinear pattern recognition methods...

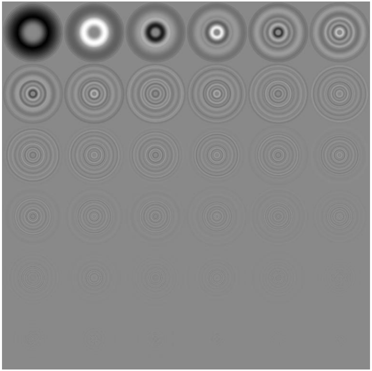
- Capture nonlinearities by appropriate preprocessing.
- Allows nonlinear effects to be modelled by a linear classifier.
- Gives more interpretable characterisations of differences.
- May lead to more accurate predictions.

One mode of geometric variability

Simulated images



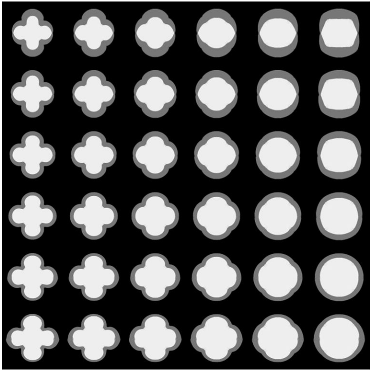
Principal components



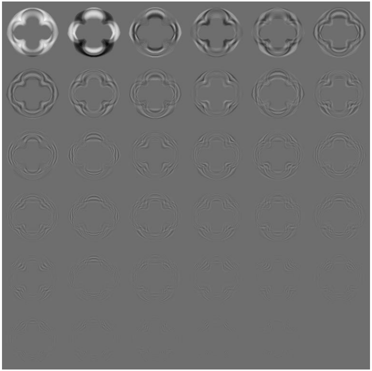
A suitable model would reduce this variability to a single dimension.

Two modes of geometric variability

Simulated images



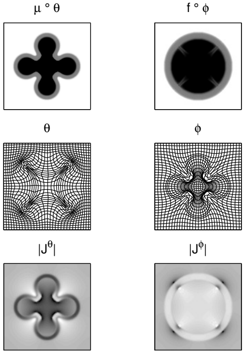
Principal components



A suitable model would reduce this variability to two dimensions.

Distance between individuals

- Image registration finds shortest distance between images.
- Often formulated to minimise the sum of two terms:
  - Distance between the image intensities.
  - Distance of the deformation from the identity.
- The sum of these can give a distance.



Non-Euclidean geometry

- Distances are not always measured along a straight line.
- Sometimes we want distances measured on a manifold.
- Shortest path on a manifold is along a *geodesic*.

Linear interpolation

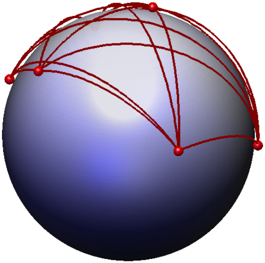
Nonlinear interpolation

Metric distances

Distances should satisfy the properties of a *metric*:

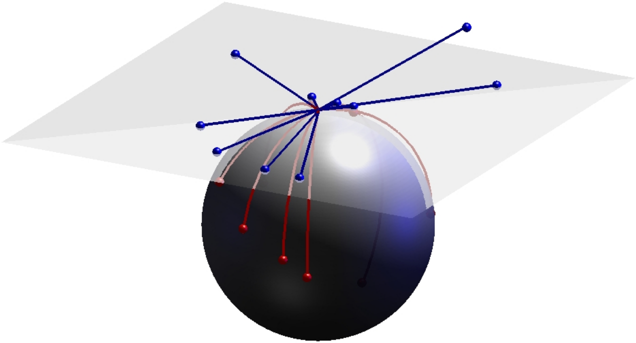
1. $d(\mathbf{x},\mathbf{y}) \geq 0$ (non-negativity)
2. $d(\mathbf{x},\mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$ (identity of indiscernibles)
3. $d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$ (symmetry)
4. $d(\mathbf{x},\mathbf{z}) \leq d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z})$ (triangle inequality).

Using metric distances

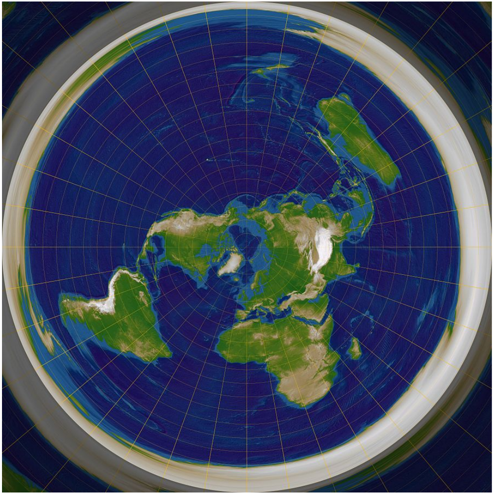


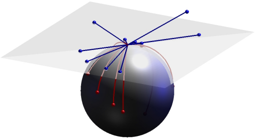
Miller et al. "Collaborative computational anatomy: an MRI morphometry study of the human brain via diffeomorphic metric mapping." Human Brain Mapping 30(7):2132–2141 (2009).

Linear approximations to nonlinear problems



Linear approximations to nonlinear problems





"Azimuthal Equidistant N90" by RokerHRO - Own work. Licensed under CC BY-SA 3.0 via Wikimedia Commons - `http://commons.wikimedia.org/wiki/File:Azimuthal_Equidistant_N90.jpg`

Thank you.