ICA in FMRI: A Basic Introduction

FMRIB Technical Report TR01MJ2

Mark Jenkinson and Christian Beckmann

Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB),
Department of Clinical Neurology, University of Oxford, John Radcliffe Hospital,
Headley Way, Headington, Oxford, UK

Abstract

This report gives a basic introduction to the ideas and theory behind ICA as applied to FMRI. It also considers the initial PCA dimensionality reduction, and discusses the orthogonality of various signals.

1 Introduction

Independent Component Analysis in FMRI is (usually) used to find a set of statistically independent spatial maps together with associated time courses. This is known as *spatial* ICA, and is used when there are more voxels of interest (i.e. those in the brain/cortex) than time points. If, on the other hand, there are more time points then it becomes possible to do *temporal* ICA.

These decompositions are well determined as long as the data consists of the mixture of a sufficient number of non-Gaussian signals. Therefore it is important to initially estimate the number of non-Gaussian signals and not attempt to estimate too many components, as then the decomposition is underconstrained.

2 Basic ICA

Let the individual voxel time courses be arranged in a matrix, Y, such that each column represents a single voxel time course, and each row represents a spatial image (at a fixed time point). That is, Y is a $T \times N$ matrix where T is the number of time points, and N is the number of voxels, where we assume $T < N$ so that spatial ICA is being performed.

The matrix is then preprocessed to:

1. remove the mean spatial map (the average of all the rows of Y) from each row of Y;
2. (optional) normalise the variance of each individual time course (each column of Y set to have unit variance);
3. remove the mean time course (the average of all the columns of Y) from each column of Y.

With this data, a general ICA decomposition can be written as

$$Y = AS \tag{1}$$

where A is a $T \times T$ matrix of time courses and S is a $T \times N$ matrix of spatial maps which are pairwise *independent* in a statistical sense. Note that the nth row of S is a spatial map that is associated with the nth column of A (a time course).

2.1 Simple Problem Formulation

In practice ICA boils down to the following problem:

Find the matrix A which minimises the function $f(S)$, where $Y = AS$.

The function $f(\cdot)$ is chosen to be an empirical estimate of the statistical dependence between the rows of S (the spatial maps). For instance, the functions which have been used to perform ICA include: (1) Mutual Information; (2) Negentropy and (3) a summary of a finite number of higher order moments/cumulants.

3 The Adventures of PCA in the Decorrelation Manifold[1]

A PCA decomposition of the original data can be written as

$$Y = A_P S_P \tag{2}$$

where the spatial maps, S_P are uncorrelated and of unit variance. That is, the principle basis vectors (spatial maps) are *orthonormal* (i.e. orthogonal and normalised). In matrix notation this is written as $S_P S_P^T = I$.

Note that this is the same as the ICA decomposition, but uses a different function to minimise — in this case, one that measures correlation.

The PCA decomposition is easily found using SVD. That is $Y = UDV^T$, with U and V being orthogonal matrices (i.e. $UU^T = VV^T = I$). Hence the PCA decomposition is given by $S_P = V^T$ and $A_P = UD$.

3.1 Dimensionality Reduction

The variance (power) of the associated time courses is given by the diagonal matrix $A_P^T A_P = D^T U^T U D = D^2$ which are the squares of the singular values. It is then possible to threshold the power so that only components that represent significant amounts of signal (assumedly not noise) are retained.

Keeping the first M components (or sources) leaves a matrix A_1 of M time courses and S_1 of the corresponding M spatial maps. That is

$$A_P = [\, A_1 \;:\; A_{reject} \,] \qquad \text{and} \qquad S_P^T = [\, S_1^T \;:\; S_{reject}^T \,]\,.$$

The reconstructed data, which has an appropriately reduced number of sources, is

$$Y_1 = A_1 S_1. \tag{3}$$

Equivalently, the dimension of the data can now be effectively reduced by pre-multiplying by the pseudo-inverse of A_1. That is

$$Y_R = A_1^- Y_1 = S_1 \tag{4}$$

where $A_1^- = (A_1^T A_1)^{-1} A_1^T$. This data matrix, Y_R, is now $M \times N$ (it has fewer time points).

Note that in each case the spatial maps are still orthogonal: that is, $S_1 S_1^T = I$.

3.2 Decorrelation Manifold

Let the matrix S_1 be mixed by a matrix Q^{-1}. The reduced data can now be represented as

$$Y_R = Q Q^{-1} S_1 = Q S_2 \tag{5}$$

where $S_2 = Q^{-1} S_1$ is a new set of spatial maps.

The correlation of these spatial maps is given by $S_2 S_2^T = Q^{-1} S_1 S_1^T Q^{-T} = Q^{-1} Q^{-T}$. Therefore, in order to keep the maps uncorrelated it is necessary to impose the condition that $Q^T Q = I$, which states that Q must be an orthogonal matrix. The set of all such matrices Q represents a set called a manifold, and since it maintains decorrelation it is known as the decorrelation manifold.

[1] In the style of Didier.

4 ICA on the PCA Reduced Data

As any collection of independent spatial maps will also be uncorrelated, the ICA result will lie in the decorrelation manifold. That is, the ICA problem can be written as one of finding the matrix Q from the previous section such that $f(S_2)$ is minimised.

So the problem is to find the orthogonal matrix Q such that $f(S_2)$ is minimised where $Y_R = QS_2$.

Note that the term 'unmixing matrix' is often used in ICA and refers to $W = Q^{-1}$ for the reduced data or $W = (A_1 Q)^{-1}$ for the full data.

5 Summary

ICA can be summarised as follows:

1. Find the PCA decomposition of the data : $Y = A_P S_P$ where $A_P = UD; S_P = V^T$ with $Y = UDV^T$ being the SVD.
2. Reduce the dimensionality by selecting the largest M components of the PCA (thresholding the power – given by D^2), giving $Y_R = S_1$.
3. Find the orthogonal matrix Q such that $f(S_2)$ is minimised where $Y_R = QS_2$. The function $f(\cdot)$ needs to measure statistical dependence of the rows of S_2 (e.g. Negentropy).
4. The resulting rows of the matrix S_2 are the ICs (spatial maps) which are orthonormal and the columns of the matrix $A_2 = A_1 Q$ are the associated time courses in the original data space (which are not orthogonal in general, although in the reduced data space the columns of Q are orthogonal).

6 Disclaimer

Clearly many details of the implementation of optimisation and cost function definitions have been excluded from this document. While these are extremely important for a thorough understanding of ICA and its implementation, it is beyond the scope of this report to address these issues (see Christian Beckmann's first year report for a detailed account).

References