석사 학위논문
Master’s Thesis

알츠하이머 질병 휴식상태 뇌기능자기공명영상
분석을 위한 그룹 회소사전학습 및 추론

Group Sparse Dictionary Learning and Inference for Resting-state fMRI
Analysis of Alzheimer’s Disease

이 정현 (李 政 炫  Lee, Jeonghyeon)
바이오및뇌공학과
Department of Bio and Brain Engineering

KAIST

2013
Group Sparse Dictionary Learning and Inference for Resting-state fMRI Analysis of Alzheimer’s Disease
Group Sparse Dictionary Learning and Inference for Resting-state fMRI Analysis of Alzheimer’s Disease

Major Advisor : Professor Ye, Jong Chul
Co-Advisor : Professor Jeong, Yong

by

Lee, Jeonghyeon

Department of Bio and Brain Engineering
KAIST

A thesis submitted to the faculty of KAIST in partial fulfillment of the requirements for the degree of Master of Science in Engineering in the Department of Bio and Brain Engineering. The study was conducted in accordance with Code of Research Ethics.


Approved by

Professor Ye, Jong Chul
[Major Advisor]

---

1Declaration of Ethical Conduct in Research: I, as a graduate student of KAIST, hereby declare that I have not committed any acts that may damage the credibility of my research. These include, but are not limited to: falsification, thesis written by someone else, distortion of research findings or plagiarism. I affirm that my thesis contains honest conclusions based on my own careful research under the guidance of my thesis advisor.
알츠하이머 질병 휴식상태 뇌기능저하기경영상
분석을 위한 그룹 회소사전학습 및 추론

이 정현

위 논문은 한국과학기술원 석사학위논문으로
학위논문심사위원회에서 심사 통과하였음.

KAIST

2012년 12월 14일

심사위원장  예 종철  (인)
심사위원  정용  (인)
심사위원  전대종  (인)
ABSTRACT

Recent research in neuroscience fields has revealed that the complicated experiments and the resting state responses are largely focused on by many researchers. In such tasks, however, we usually can not establish accurate hypothesis or even there is no experiment paradigm in the resting state measurement. And, it has been a core challenge to develop a universal method for squeezing out the meaningful information in the brain without any constraints of prior information or hypotheses on the functional magnetic resonance imaging (fMRI) data we measured. In contrast to the general linear model (GLM) that requires the hypothesis of expected response to certain tasks which is widely known and used, data-driven analysis methods have attracted attention thanks to their nature capable of blind separation of sources. The representative methods of these are principal component analysis (PCA) and independent component analysis (ICA). Although these classical data-driven analysis methods, PCA and ICA, have long been attractive methods in terms of providing analysis tools with no prior information about tasks of the fMRI data, which both have constraints that sources are assumed to be orthogonal and independent, respectively. These constrains often lead the restriction on explaining the underlying activities of brain in fMRI data. Recently, there has been increased interest in the use of neuroimaging techniques to investigate what happens in a brain at rest. Functional imaging studies have revealed that the default-mode network activity is disrupted in diseases such as Alzheimer’s (AD). However, there is no consensus, as yet, on the choice of analysis method for the application of resting-state analysis. A novel group analysis tool for data-driven resting state fMRI analysis using group sparse dictionary learning and mixed model is presented along with the promising indications of Alzheimer’s disease progression. Instead of using independency assumption as in popular ICA approaches, the proposed approach is based on the sparse graph assumption such that a temporal dynamics at each voxel position is a sparse combination of global brain dynamics. In estimating the unknown global dynamics and local network structures, we perform sparse dictionary learning for the concatenated temporal data across the subjects by constraining that the network structures within a group are similar. Under the homoscedasticity variance assumption across subjects and groups, we show that the mixed model group inference can be easily performed using second level GLM with summary statistics. Using extensive resting fMRI data set obtained from normal, Mild Cognitive Impairment (MCI), Clinical Dementia Rating scale (CDR) 0.5, CDR 1.0, and CDR 2.0 of Alzheimer’s disease patients groups, we demonstrated that the changes of default mode network extracted by the proposed method is more closely correlated with the progression of Alzheimer’s disease.
Contents

Abstract .................................................. i
Contents .................................................... ii
List of Tables ............................................ iv
List of Figures .......................................... v

Chapter 1. Introduction .................................. 1

Chapter 2. Theory ........................................ 3
  2.0.1 Sparse Graph Model for functional Connectivity Analysis .......... 3
  2.0.2 ReML Estimation of Group-wise Sparse Graph ..................... 5
  2.0.3 Inference on Group Differences ................................. 8

Chapter 3. Method ........................................ 11
  3.1 Data Acquisition ...................................... 11
  3.2 Data Analysis using Conventional Methods ........................... 11
    3.2.1 Independent Component Analysis ............................. 11
    3.2.2 Seed Based Analysis ....................................... 11
  3.3 Data Analysis using Sparse SPM ................................ 12
    3.3.1 Preprocessing ............................................. 12
    3.3.2 Group Sparse Dictionary Learning .......................... 12
    3.3.3 Group Estimation and Inference ............................. 15

Chapter 4. SPARSE-SPM: A new toolbox for a data-driven group fMRI analysis ........................................ 16

Chapter 5. Results ....................................... 21
  5.1 Evaluation of the Proposed Method ............................... 21
    5.1.1 Training Data set in Sparse Dictionary Learning .............. 21
    5.1.2 Initial Dictionary in Sparse Dictionary Learning ............ 21
    5.1.3 Number of iterations in Sparse Dictionary Learning ....... 21
  5.2 Comparative Studies ..................................... 25
    5.2.1 Extracted Default Mode Network ............................ 25
    5.2.2 Group Comparisons ....................................... 26
    5.2.3 Regression Analysis ....................................... 26
  5.3 Reproducibility of the results ................................ 27
List of Tables

3.1 Sparsity level estimation using MDL criterion for each group data. . . . . . . . . . . . . . 14
4.1 System environment in developing and testing of Sparse SPM . . . . . . . . . . . . . . . . 17
5.1 Residual errors for different training dataset . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
List of Figures

2.1 A graph theoretical model for Sparse SPM. ................................. 4

2.2 By sparse coding step, for each dictionary atom, its own voxel map is obtained. And then, each own voxel map can be thought of same community sharing mostly same information (the dictionary atom) flow in a brain. ........................................ 8

3.1 Training data extraction for sparse dictionary learning from individual data sharing a same a group gray matter mask. .................................................. 12

3.2 By sparse coding step, for each dictionary atom, its own voxel map is obtained. And then, each own voxel map can be thought of same community sharing mostly same information (the dictionary atom) flow in a brain. ........................................ 13

3.3 Signal at each voxel is sparse combination of global dictionary atoms which can be thought that a voxel belongs to few community not all in brain network structures, and common dictionary atom across different voxels represents that those voxels are bound with the community by sharing the common dictionary atom. ........................................ 13

3.4 The choice of the number of dictionary atoms using eigenspectrum of covariance matrix of normal group data. (DF: $p_1, q - k$ for the first row / DF: $p_1, M - k$ for the second row of the insets) .................................................. 14

3.5 Group activation detection using a learned group dictionary. .................. 15

4.1 Main panel of the SPARSE-SPM .................................................. 17

4.2 Data Load panel of the SPARSE-SPM ........................................ 17

4.3 Results viewer of the SPARSE-SPM ........................................ 18

4.4 Block diagram of the proposed method ....................................... 19

5.1 The DMN patterns from different training dataset for sparse dictionary learning step. ($p < 0.001$, DF: $p_1, q - k$ for the first three rows / DF: $p_1, M - k$ for the last three rows) 22

5.2 Although overall converged error was slightly higher in the most variable components case as in Figure. 5.2, the DMN pattern was observed clearer in the case of initial dictionary of the most variable components than in the case of the first few components from the data itself. ($p < 0.001$, DF: $p_1, q - k$ for the first two rows / DF: $p_1, M - k$ for the last two rows) 23

5.3 (a) Residual error after K-SVD training. (b) The residual errors after K-SVD training reflect the progression of AD disease ( **: $p < 0.00001$, *: $p < 0.0001$) ........................................ 24

5.4 The results of extracted DMN map using seed based analysis (first row), ICA (second row), the proposed method widh df($p_1, q - k$) (third row), and the proposed method widh df($p_1, M - k$) (fourth row). $p < 0.001$ for the proposed methods. ................................. 25

5.5 The 3D rendering images of the results using proposed method with $p$-value < 0.001. Each column of the images represents the group, normal, MCI, CDR 0.5, CDR 1.0, and CDR 2.0 from left to right. The images were created by using Amira(R). (DF: $p_1, q - k$ for the first row / DF: $p_1, M - k$ for the second row) ................................. 26
5.6 1 × 2 ANOVA maps with $p < 0.05$: (a) normal vs. MCI (b) MCI vs. CDR 0.5, (c) CDR 0.5 vs. CDR 1.0, (d) CDR 1.0 vs. CDR 2.0. The top rows represent slice −6, and bottom one is slice 30. (DF: $p_1$, $q - k$ for the left image / DF: $p_1$, $M - k$ for the right image for each comparison) ............................................ 26

5.7 Linear regression analyses. (a) DMN ($R^2 = 0.7740, p = 0.0491$), (b) MPFC ($R^2 = 0.7021, p = 0.0764$), (c) PCC ($R^2 = 0.6869, p = 0.0828$), (d) RLP ($R^2 = 0.7951, p = 0.0421$), (e) LLP ($R^2 = 0.7906, p = 0.0436$), (f) R/LLP ($R^2 = 0.7953, p = 0.0420$), (g) DMN ($R^2 = 0.6152, p = 0.1163$), (h) MPFC ($R^2 = 0.6876, p = 0.0825$), (i) PCC ($R^2 = 0.5763, p = 0.1367$), (j) RLP ($R^2 = 0.5906, p = 0.1290$), (k) LLP ($R^2 = 0.5814, p = 0.1339$), and (l) R/LLP ($R^2 = 0.5892, p = 0.1297$) (DF: $p_1$, $q - k$ for the first two rows / DF: $p_1$, $M - k$ for the last two rows) ......................................................... 28

5.8 The DMN patterns are observed in the case of random division of normal group data in the case of DF ($p_1, M - k$). ($p < 0.001$, DF: $p_1$, $q - k$ for the left image / DF: $p_1$, $M - k$ for the right image for each comparison) ......................................................... 29
Chapter 1. Introduction

Resting state brain activity has been consistently observed in functional connectivity MRI (fcMRI), as the temporal correlation of spontaneous low-frequency fluctuations ($< 0.1 \text{Hz}$) arising from blood oxygen level-dependent (BOLD) signal in cortical areas have demonstrated functional connectivities in these areas of brain [1, 2]. BOLD signal oscillations during rest reflect baseline neural activity representing underlying relevance of cognitive functions and neural physiology [3]. Spatiotemporally distinct resting state networks has been consistently identified in the following areas; 1) a posterior network including occipital cortex, 2) a posterior-lateral and midline network involving primarily the precuneus and anterior pole of the prefrontal lobe, 3) a lateral and midline network comprised of the pre- and post-central gyri, thalamus and hippocampus, 4) a network involving dorsal parietal and predominantly lateral prefrontal cortex, and 5) a ventral network which covers inferior occipital parietal, temporal and inferior prefrontal cortices, and etc [4].

Recently, a network in a set of brain regions which significantly co-deactivated during cognitive task-related experiments (often referred as ‘default-mode network (DMN)’) has been studied extensively in functional connectivity analysis. The DMN is typically detected in the posterior cingulate cortex (PCC), the bilateral inferior parietal cortex (IPC), the medial prefrontal cortex (MPFC), the left parahippocampal gyrus (PHG), and the left inferolateral temporal cortex (ITC) [5], and it has been shown that the default-mode network is closely involved with episodic memory processing [6, 7]. Furthermore, previous works have provided evidences that PCC which shows a neural deactivation in early Alzheimer’s disease (AD) is the earliest brain region to exhibit decreased metabolism in AD [8].

Seed-based approaches [9, 10, 11, 12] and independent component analysis (ICA)-based approaches [13, 14] are the most commonly used analysis methods in functional connectivity study. Seed-based approach extracts BOLD signal time courses from a region-of-interest (ROI) called as “seed” region, and computes cross-correlation between time course signals from the ROI and all other voxels in a brain to obtain a map of neuronal connectivity [15]. Despite of its popularities, it has a limitation that a priori definition of seed region significantly affects results. On the other hand, ICA automatically decomposes whole BOLD dataset into maximally independent components, which are relevant to each spatial maps associated with neural signal sources or noise. However, it has been demonstrated that the success of ICA for resting state data analysis is due to their ability to handle sparse components rather than independent components [16]. The observation that sparsity is more important factor for data-driven analysis is also supported by biological findings of sparse coding in many areas of a brain such as in simple-cells in the primary visual cortex (V1) [17], the medial temporal lobe (MTL) neurons [18] and single neurons in the left posterior hippocampus [19].

On the other hand, graph theory-based quantitative analyses of brain connectivity have been developed to study structural and functional brain network and their interactions [20]. Since the small-world properties of functional voxel-level networks in a low-frequency whole-brain network have been reported in resting state fMRI studies [21, 22], other network measures such as node degree or hubs, clustering coefficient, path length and modularity have been explored to quantify brain functional networks. Furthermore, economical small-world concept was introduced using several metrics by representing that the structural network of the macaque brain was shown to have high global efficiency of parallel information.
transfer and high efficiency, and they are sparsely connected [23]. However, graph theory-based analysis is dependent upon pre-defined parcellations, which are often obtained by ICA or coordinate based segmentations. Therefore, parcellation independent graph theory analysis are required.

Unlike the convention approaches, we propose a parcellation-free functional connectivity analysis that are derived solely based on the graph theoretical approach for brain network. More specifically, our signal decomposition is based on a sparse graph model that regards a temporal dynamics at each voxel as a sparse combination of unknown global information flow. Interestingly, we show that the sparse dictionary learning in [24] is an intuitive learning rule for such model, where resulting spatially adaptive design matrix represents local connectivities. As the mathematical framework for inference turns out similar to that of a standard SPM type analysis with only exception of spatially adaptive design matrix, rich statistical analysis tools such as p-value correction using random field theory as well as hypothesis driven inference can be used.

However, one of the remaining technical difficulties in the proposed model is that the extracted temporal dynamics corresponding to each network highly depends on each individual. Moreover, individual dependent regressors should be estimated at the same time while the group level statistical inferences should be performed using the subject specific regressors. This makes the group sparse learning and statistical inference complicated. In this paper, to overcome such technical difficulties in group analysis, we propose a unified mixed model framework where a group level sparse dictionary learning and group inference can be performed in a unified linear mixed model framework. More specifically, in estimating the unknown global dynamics and local network structures in a group level, we propose a group sparse dictionary learning for the concatenated temporal data across the subjects by constraining that the network structures within a group are similar. Under the homoscedasticity variance assumption across subjects and groups, which are often used in SPM, we show that such group sparse dictionary learning and the group inference can be derived under restricted maximum likelihood (ReML) framework.

To confirm the validity of the proposed method, we provide extensive comparison using group data from normal, MCI, CDR 0.5, CDR 1.0 and CDR 2.0 scale Alzheimer subjects. Results indicate that extracted DMNs using the proposed method exhibit excellent correlation with disease progression, whereas other existing methods such as seed-based, or ICA approach do not exhibit consistent changes of DMN patterns. Considering clinical findings of decreased metabolism in DMN with disease progression, we believe that the results in this paper provides a strong evidence that the graph theoretical approach is a powerful tool for resting state fMRI analysis.
Chapter 2. Theory

Throughout the paper, $x_i$ and $x_j$ correspond to the $i$-th row and the $j$-th column of matrix $X$, respectively. When $S$ is an index set, $X^S$ and $A_S$ correspond to a submatrix collecting corresponding rows of $X$ and columns of $A$, respectively; $x_S$ denotes a subvector collecting the corresponding elements of $X$. The superscripts $'$ and $\dagger$ denote the adjoint operator and pseudo-inverse, respectively. A vector $1_L$ denotes a $L$-dimensional vector with elements of ones, and $I_{k \times k}$ is $k \times k$ identity matrix.

2.0.1 Sparse Graph Model for functional Connectivity Analysis

In the proposed method, the interactions of neural signal between brain’s functional systems are modeled by a set of nodes (voxels) linked by connections as shown in Figure. 2.1. Each circle denotes voxel or node where a temporal dynamics of BOLD signal is measured continuously. A set of nodes in functional brain network form a community sharing same information flows in addition to long-range connections from different communities. For example, in Figure. 2.1, time series at the node 1 and 2, denoted by $y_1 \in \mathbb{R}^m$ and $y_2 \in \mathbb{R}^m$, is given by

$$y_1 = 3d_1 + d_3 + d_4,$$

$$y_2 = 4d_2 + d_4 + d_5 ,$$

where $m$ denotes the length of the temporal trace. If we define a global dictionary $D$ by collecting all local or long-range information flows as

$$D = [d_1, d_2, \ldots, d_5],$$

then we can easily see that the temporal dynamics at the nodes 1 and 2 are described as a sparse linear combination of the atoms in the global design matrix. In general, we have

$$y_n = D_{I_n} \alpha_n + \epsilon_n, \quad n = 1, \ldots, N$$

(2.1)

where $N$ is the number of voxel, $\epsilon$ denotes noise, $D_{I_n} \in \mathbb{R}^{m \times k}$ is a submatrix of $D \in \mathbb{R}^{m \times Q}$ composed of elements in the index set $I_n$, and $\alpha_{I_n} \in \mathbb{R}^k$ denotes the corresponding weight vectors. Note that a local subset index $I_n$ represents a local network structure at the $n$-th voxel, and a regressor is a representative dynamics in a network module or community that shares the same information flow.

In order to extract a statistically significant network module that shares the same information flow, a binary hypothesis test is conducted to test whether a specific information flow of interest ($z \in D$) is presented in each voxel. This test can be formulated as following: a null hypothesis $H_0 : \theta_n = 0$ is tested against an alternative hypothesis $H_1 : \theta_n \neq 0$, where $\theta_n$ is a regression coefficient for $z$:

$$y_n = z\theta_n + D_{I_n} \backslash z \alpha_{n \backslash z} + \epsilon_n$$

(2.2)

where $\backslash z$ denotes a reduced size matrix or vector made by removing the elements corresponding to the
atom $z$. When $z$ belong to the set of atoms from $D_{In}$, we have the following $F$-statistics:

$$F_n = \frac{y_n'(P_{D_{In}\backslash z}^\perp - P_{D_{In}}^\perp)y_n}{m-k}$$

(2.3)

where $P_{D_{In}}^\perp, P_{D_{In}\backslash z}^\perp$ denotes the projection on the orthogonal complement on the range space of $D_n$ and $D_n\backslash z$, respectively. If $z$ is not presented in the local dictionary $D_{In}$, then $\theta_n = 0$ and the signal at the voxel is irrelevant to the network associated with the atom $z$; so, $H_0$ holds and $F_n = 0$. For example, in Figure 2.1, if we test the presence of $z = d_1$ (or $z = d_2$), we can obtain the nodes in community 1 (or community 2). By testing $z = d_3$ or $z = d_4$, the connector hubs between two community can be obtained. Since the information flow $d_5$ is isolated with regression coefficient 1, it is likely that the test with $z = d_5$ are not statistically significant, so the isolated edges can be removed during the test.

For a group of $L$ subjects, we perform similar connectivity analysis using the concatenated times series from multiple subjects. More specifically, suppose $Y^{(i)} = \{y^{(i)}_1, \cdots, y^{(i)}_N\}$ denotes a collection of time trace across all $N$ voxels, then we construct a concatenated temporal time trace from $L$ subjects across all voxels $Y = [Y^{(1)'}, \cdots, Y^{(L)'})'$. Suppose, furthermore, we have the following decomposition of $Y$

$$Y = DW = \begin{bmatrix} Y^{(1)} \\ \vdots \\ Y^{(L)} \end{bmatrix}$$

(2.4)

where $D$ denotes the concatenated global sparse dictionary and $D^{(i)}$ denotes the corresponding $l$-th subject individual sparse dictionary, and $W$ is the corresponding coefficients, whose non-zero elements are sparse. In our model, the individual dictionary $D^{(i)}$ is assumed to be normalized such that each column has unit norm. Now, note that the $i$-th column of $D$ represents the representative temporal dynamics at the voxel locations that have non-zero coefficients in the $i$-th row of $W$. Hence, the $i$-th column of resulting individual dictionary $\{D^{(i)}\}_{i=1}^L$ share the same geometric connectivity. Hence, our assumption that the local design matrix index $I_n$ is same across all subjects imposes the constraint that local network structure within a group is same. This learning rule is very advantageous to identify the group differences since the network connectivity changes between groups are one of the main biomarkers in resting-state fMRI analysis.
2.0.2 ReML Estimation of Group-wise Sparse Graph

Mathematically, the temporal dynamics at the $n$ voxel of the subject $l$ can be modeled as

$$y_n^{(l)} = D_{I_n}^{(l)} w_n^{(l)} + \epsilon_n^{(l)}, \quad \epsilon_n^{(l)} \sim \mathcal{N}(0, R_n^{(l)}) \quad (2.5)$$

where $l = 1, \cdots, L$ denote the subject index, and $R_n^{(l)}$ denotes the $l$-th subject temporal correlation matrix at the $n$ voxel. Now, the subject differences within a group can be modeled as random effects:

$$w_n^{(l)} = \alpha_n + \beta_n^{(l)}, \quad \beta_n^{(l)} \sim \mathcal{N}(0, G_n), \quad (2.6)$$

where $\alpha_n$ denote a group mean, and $G_n$ is random effect covariance matrix. Then, for the concatenated individual data, we can obtain the following mixed model

$$y_n = X_n \alpha_n + Z_n \beta_n + \epsilon_n,$$  \hfill (2.7)

where

$$y_n = [y_n^{(1)'}, \cdots, y_n^{(L)'}]' \in \mathbb{R}^M$$

$$\beta_n = [\beta_n^{(1)'}, \cdots, \beta_n^{(L)'}]' \in \mathbb{R}^q$$

where $M = mL$ and $q = kL$; and the fixed effect and random effect matrix $X_n$ and $Z_n$ are given by

$$X_n = \begin{bmatrix} D_{I_n}^{(1)} \\ D_{I_n}^{(2)} \\ \vdots \\ D_{I_n}^{(L)} \end{bmatrix}, \quad Z_n = \begin{bmatrix} D_{I_n}^{(1)} & 0 & \cdots & 0 \\ 0 & D_{I_n}^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_{I_n}^{(L)} \end{bmatrix},$$

or equivalently,

$$X_n = Z_n X_G, \quad X_G = 1_L \otimes I_k.$$

Note that in our sparse graph estimation problem, there are several unknown that needs to be estimate, which include the global dictionary $D$, local network connectivity $\{I_n\}_{n=1}^N$, the fixed and random effect regression coefficients $\{\alpha_n, \beta_n\}_{n=1}^N$ as well as the covariance matrices $\{R_n^{(l)}\}_{n,l=1}^{N,L}$ and $\{G_n\}_{n=1}^N$. Maximum likelihood (ML) principle is perhaps the most popular parameter estimation procedure in statistics owing to many optimality properties, such as consistency [25]. However, one of the main problems of ML approach for covariance estimation is that it does not take into account of the loss in degrees of freedom resulting from the estimation of the model’s fixed effects [25, 26]. To address this issue, Patterson and Thomson (1971) [26] proposed a method which takes into account the loss in degrees of freedom resulting from estimating fixed effects. More specifically, they restricted their attention to a set of such contrast that is invariant to the fixed effect parameter and then estimate the covariance for such restricted cases of maximum likelihood. This idea is so-called restricted maximum likelihood (ReML) method [25, 27, 28, 29, 30]. More specifically, if we are interested in finding the restricted maximum likelihood that are invariant to the fixed effect parameter, we need to minimize the following form of the ReML equation [25, 31]:

$$C_{ReML} \left( \{R_n^{(l)}\}, \{G_n\}, \{I_n\}, \{\alpha_n\} \right) \quad (2.8)$$
\[
\begin{align*}
&= -\sum_n \left( \frac{1}{2} \log |V_n| + \frac{1}{2} \log |X_n^{-1}V_n^{-1}X_n| + \frac{1}{2} (y_n - X_n\alpha_n) V_n^{-1} (y_n - X_n\alpha_n) \right), \\
\text{where } V_n \text{ is given by} \\
V_n &= Z_n G_n Z_n' + R_n, \quad (2.9) \\
R_n &= \begin{bmatrix}
R_1^{(l)} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & R_L^{(l)}
\end{bmatrix}, \quad (2.10)
\end{align*}
\]

However, ReML is usually computationally expensive. To simplify the covariance estimation and the resulting inference, we first approximate that \(D(I_n)'D(I_n) \simeq I\). Since each individual dictionary is normalized to have unit norm, this implies that each regressors are nearly orthogonal, which is often used in SPM analysis. Next, we assume that the noise in individual temporal dynamics is white, i.e. \(R^{(l)}_n = \sigma_{n}^2 I\), and the random effect variance is spherical, i.e. \(G_n = g_n^2 I\). This can be easily satisfied using prewhitening procedure. Finally, we employ the homoscedasticity variance assumption across subjects and groups, i.e. \(\sigma_{n}^2 = \sigma_n^2\), which has been also often used in SPM. Then, the resulting cost function becomes

\[
C_{ReML} \left( \{\sigma_n^2\}, \{g_n^2\}, \{I_n\}, \{\alpha_n\} \right) = -\sum_n \left( \frac{L(M-k)}{2} \ln(\sigma_n^2 + \sigma^2) + \frac{1}{2(\sigma_n^2 + \sigma^2)} \|y_n - X_n\alpha_n\|^2 \right).
\]

Now, the problem can be solved using an alternating minimization. More specifically, for the estimated local connectivity \(I_n\) and the fixed effect parameter estimation \(\hat{\alpha}_n\), we have

\[
\hat{g}_n^2 + \hat{\sigma}_n^2 = \frac{1}{L(M-k)} \|y_n - X_n\hat{\alpha}_n\|^2 = \frac{1}{L(M-k)} \sum_{l=1}^{L} \|y_n^{(l)} - D(I_n)^{l}\hat{\alpha}_n\|^2 \quad (2.11)
\]

Next, for a given sparsity level \(k\) for the maximum number of local connection, we claim that the estimation problem of local connectivity \(I_n\), global dictionary \(D\), and the fixed effect parameter estimation \(\hat{\alpha}_n\) can be equivalently written as a minimization problem with respect to \(D, W\):

\[
\min_{D, W} \|Y - DW\|_F, \quad \text{subject to } \|w_n\|_0 \leq k. \quad (2.12)
\]

To see the equivalence, let us consider a sparse regression problem for a given global dictionary \(D\):

\[
\min_{w_n} \|y_n - Dw_n\|^2, \quad \text{subject to } \|w_n\|_0 \leq k. \quad (2.13)
\]

which is equivalent to

\[
\min_{I_n, \alpha_n} \|y_n - DL_n\alpha_n\|^2, \quad \text{subject to } |I_n| \leq k. \quad (2.14)
\]

Therefore, solving (2.13) and obtain the nonzero index of \(\hat{w}_n\) as \(I_n\), we can obtain \(D_{I_n}\) and \(\hat{\alpha}_n\).

The simultaneous estimation of \(D\) and \(W\) in (2.12) can be addressed by using a sparse dictionary learning algorithm called K-SVD [32], which attempts to find the best possible dictionary \(D\) for a sparse representation of \(Y\) in a greedy manner using sparse coding step and codebook update step. More
specifically, for a given dictionary $D$, the sparse coding solves (2.13) for $n = 1, \ldots, N$. The active index set $I_n$ is then estimated by collecting indices corresponding to the $k$-largest correlation between $y_n$ and $d_j$ as:

$$C_{y_n}(j) = \frac{||y_n^T d_j||^2}{||d_j||^2}, \quad j = 1, \ldots, Q,$$

where $Q$ denotes the number of atoms in a global dictionary $D$. The fixed effect parameter estimate $\hat{\alpha}_n$ is then given by

$$\hat{\alpha}_n = (D_{I_n}^T D_{I_n})^{-1} D_{I_n}^T y_n.$$

Then, dictionary update step refines the each column vector $d_j (j = 1, \ldots, Q)$ and the corresponding coefficient row vector $w^j$ sequentially by a rank-1 approximation using singular value decomposition. More specifically, with estimated $W$, K-SVD puts in question only one column in the dictionary, $d_j$, and the corresponding coefficient $\tilde{w}^j$, the $j$-th row of $\tilde{W}$. This can be solved using singular value decomposition (SVD) with sparsity constraint. Then, for each $j = 1, 2, \ldots, Q$, the K-SVD does the following: (i) define the index set $\omega^j$ corresponding to non-zero indices of $\tilde{w}^j$, (ii) compute $E_j = Y - \sum_{p \neq j} d_j \tilde{w}^j$, (iii) define $\Omega_j$ as a diagonal matrix with ones for the indices corresponding to $\omega^j$ and zeros elsewhere, (iv) choose a subset $E_j^R = E_j \Omega_j$, (v) take SVD to the restricted $E_j^R$,

$$E_j^R = U \Lambda V^T = \sum_{p=1}^{P} \sigma_p u_p v_p$$

and (vi) update $\hat{d}_j = u_1$, $\tilde{w}^j_R = \sigma_1 v_1$.

Note that the K-SVD procedure has very interesting interpretations for data-driven fMRI analysis. First, as shown in Figure 2.2, an update stage for an atom $d_j$ in dictionary update stage of K-SVD is equivalent to finding the largest principle component from a set of the raw data $\{y_n\}_{n \in \omega_j}$, where $\omega_j$ denotes the voxel indices that has non-zero coefficients from $\tilde{w}^j$. More specifically, $\omega_j$ denotes the voxel indices for a community or a long range connections that share the dynamics $\hat{d}_j$ as shown in Figure 2.2. Hence, the dictionary update stage is basically to update a representative temporal dynamics that constitute a specific network structures by removing noise using principle component analysis. Hence, this corresponds to filtering procedure.

Second, recall that a sparse coding step estimates the non-zero coefficients of the regression coefficients; so it updates the local network structures by which each voxel is connected to other voxels in brains. For example, as shown in Figure 2.2, the resulting non-zero voxels that corresponds to the $d_1$ constitutes a local community that shares the same temporal dynamics $\hat{d}_1$. Therefore, this procedure corresponds to a clustering procedure that identifies the voxels with similar dynamics. In particular, for a sparse coding stage, if we use the correlation (2.15) to update $I_n$, then the clustering is based on the correlation with respect to raw temporal dynamics $y_n$ and the updated local information flow $d_j$, which is very similar to the seed-based analysis. Hence, K-SVD sparse coding with (2.15) can be considered as seed-based clustering. However, the main difference between the conventional seed-based connectivity analysis and the proposed sparse coding stage is that in the proposed method the seed $d_j$ is adaptively updated during the dictionary estimation stage, which is not the case in the conventional approach.

Therefore, such repeated application of filtering and adaptive seed-based clustering procedure allows our method to exploit both the advantages of the conventional seed-based approach and the ICA type of clustering approaches, which makes the proposed algorithm very powerful.
Figure 2.2: By sparse coding step, for each dictionary atom, its own voxel map is obtained. And then, each own voxel map can be thought of same community sharing mostly same information (the dictionary atom) flow in a brain.

2.0.3 Inference on Group Differences

Suppose we are interested in comparing two different groups. Once the group sparse dictionary learning is performed for each group data, the temporal dynamics at the $n$ voxel of the subject $l$ in the $i$-th group model is modeled as

$$y_{n}^{(il)} = D_{n}^{(il)} x_{n}^{(il)} + e_{n}^{(il)}, \quad e_{n}^{(il)} \sim \mathcal{N}(0, R_{n}^{(il)})$$

(2.18)

where $i = 1, 2$, $l = 1, \cdots, L_i$ denote the indices for group and the subject index for each group, respectively, and $L_i$ is the number of subjects for the group $i$. Now, the subject differences within a group can be modeled as random effects:

$$x_{l}^{(il)} = \alpha_{n}^{(i)} + \beta_{n}^{(il)}, \quad \beta_{n}^{(il)} \sim \mathcal{N}(0, G_{n}),$$

(2.19)

where $\alpha_{n}^{(i)}$ denote a group mean. Hence, if we stack the data together, (after ignoring the voxel dependent index $n$), we have the following mixed model

$$y = X\alpha + Z\beta + \epsilon,$$

(2.20)

where

$$y = [y_{n}^{(11)'}, y_{n}^{(12)'}, \cdots, y_{n}^{(1L_1)',} \cdots, y_{n}^{(2L_2)'}]' \in \mathbb{R}^{M}$$

$$\alpha = [\alpha^{(1)'}, \alpha^{(2)'}]' \in \mathbb{R}^{p}$$

$$\beta = [\beta_{n}^{(11)'}, \beta_{n}^{(12)'}, \cdots, \beta_{n}^{(1L_1)',} \cdots, \beta_{n}^{(2L_2)'}]' \in \mathbb{R}^{q}$$

for $M = m \sum_i L_i$ and $q = k \sum_i L_i$; and the random effect matrix $Z$ is given by
\[ Z = \begin{bmatrix}
 D^{(11)} I^{(1)}_n & 0 & \cdots & 0 \\
 0 & D^{(12)} I^{(1)}_n & \cdots & 0 \\
 \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & \cdots & D^{(2L_z)} I^{(1)}_n 
\end{bmatrix}, \]

and the fixed effect matrix is

\[ X = ZX_G, \quad X_G = \begin{bmatrix}
 1_{L_1} \otimes I & 0 \\
 0 & 1_{L_2} \otimes I 
\end{bmatrix}. \]

Then, using the standard results for the solution of mixed model [33], the fixed effect parameter estimate \( \hat{\alpha} \) from the mixed model is equivalent to the second level GLM using the summary statistics

\[ \hat{\chi} = X_G \alpha + \eta, \quad \eta \sim \mathcal{N}(0, V_G), \quad (2.21) \]

where \( \hat{\chi} = [\hat{w}^{(11)}', \hat{w}^{(12)}', \ldots, \hat{w}^{(L_1)}', \ldots, \hat{x}^{(2L_2)}'] \) and

\[ \hat{w}^{(il)} = \left( D^{(il)} I^{(il)}_n \right)^{-1} D^{(il)} Y^{(il)}_n \quad (2.22) \]

Here, for \( G_n = g_2^2 I \), \( V_G \) is a block diagonal matrix whose \((il)\)-th block is composed of

\[ \left( D^{(il)} I^{(il)}_n \right)^{-1} + g_2^2 I. \quad (2.23) \]

For group inference on group mean activation or differences, we are interested in testing the following null hypothesis:

\[ H_0 : C \alpha = 0, \]

where \( C \in \mathbb{R}^{p_1 \times k} \) denotes the contrast matrix. Since \( (X' V^{-1} X)^{-1} = (X_G' V_G^{-1} X_G)^{-1} \) as shown in Appendix A, the test statistics for the mixed model is equivalent to the second level inference statistics:

\[ S = \frac{\hat{\chi}^T C' \left( C (X' V^{-1} X)^{-1} C' \right)^{-1} C \hat{\chi}}{p_1} = \frac{\hat{\chi}^T C' \left( C (X_G' V_G^{-1} X_G)^{-1} C' \right)^{-1} C \hat{\chi}}{p_1}. \]

Now, under the aforementioned variance assumptions and orthogonal regressors, we can show that \( \hat{V}_G \) is diagonal matrix, hence using the equivalence relationship in Appendix B, we can show that

\[ S = \frac{\hat{\chi}^T (P_{X_G,0} - P_{X_G} \hat{\chi}) \hat{\chi} M - k}{p_1} \sim F_{p_1, u} \quad (2.24) \]

where \( X_{G,0} \) denotes the reduced model by excluding the effect estimated by contrast \( C \) and \( u = m \sum_{il} L_{il} - k \) and \( \hat{\chi} \) is summary statistics. Since this \( F \)-statistics are standard statistics for ANOVA analysis, the result indicates that we can perform classical ANOVA analysis using the summary statistics, and such analysis is equivalent to the inference in mixed model as long as our assumption holds. Moreover, we do not need to perform computationally expensive ReML covariance estimation since the
ReML variance estimation parts are already built-in within the resulting $F$-statistics.

Since the degree of freedom is same across all voxels, it is easy to apply random field based $p$-value correction [30] for the resulting $F$-maps, which is another advantage of the proposed method compared to ICA approach that has heuristic way of addressing family-wise error correction.
Chapter 3. Method

3.1 Data Acquisition

We collected five groups of resting-state fMRI data: 1) 22 normal subjects (8 male, mean age 70 years), 2) 37 MCI patients (21 male, mean age 72 years), 3) 20 AD patients with CDR 0.5 (5 male, mean age 72 years), 4) 27 AD patients with CDR 1.0 (6 male, mean age 73.5 years), and 5) 13 AD patients with CDR 2.0 (6 male, mean age 73.6 years). During the task period, subjects were instructed to awake and alert, but not actively involved in a task with eye closed. A 3.0T fMRI system (Philips, Netherland) was used to measure the BOLD response. The echo planar imaging (EPI) sequence was used with TR/TE = 3000/35 ms, flip angle = 90°. Each acquisition consisted of 35 continuous slices, and FOV (RL, AP, FH) = 220 mm x 140 mm x 220 mm, and a voxel size of (RL, AP) = 2.875 mm x 2.875 mm. In the subsequent anatomical scanning session, T1-weighted structural images were acquired. A total of 100 acquisitions are obtained for each subject thus the total recording time was 300 sec. The experiments have been approved by the Institutional Review Board of the Samsung Medical Center in South Korea.

3.2 Data Analysis using Conventional Methods

We used two conventional methods for resting state fMRI analysis to provide the evidence that our proposed method shows more distinct difference in DMN spatial map along with the course of AD.

3.2.1 Independent Component Analysis

First, Multi-session temporal concatenation of Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC v3.0) within FMRIB’s Software Library (FSL) [34, 14] is used as an ICA methods. We began with Brain Extraction Tool (BET v2.1) for the anatomy data with the fractional intensity threshold of 0.5 and the option ‘Robust brain centre estimation (iterates BET2 several times)’ to obtain brain image out of whole anatomy image. The following parameters were applied in MELODIC analysis: 128 sec for the high pass filter cutoff, motion correction, smoothing using a Gaussian kernel of FWHM 8 mm, normalization into MNI coordinates with resampling resolution 2 mm, variance-normalise timecourses, and 20 independent components for the dictionary regressor to make the analysis condition as same as the one of our proposed method.

3.2.2 Seed Based Analysis

Second, we used Functional connectivity toolbox (conn) based on Statistical Parametric Mapping (SPM) for seed based analysis [35]. The region PCC was used for ROI of this analysis and following procedures were also applied: realignment, segmentation, normalization, smoothing using a Gaussian kernel of FWHM 8 mm, band pass filter with cutoff frequency of 0.008 – 0.09 Hz, and threshold p-value was 0.001.
3.3 Data Analysis using Sparse SPM

3.3.1 Preprocessing

The images were first spatially realigned to remove movement artifacts in fMRI time series. The images were then spatially normalized to a standard space, Montreal Neurological Institute (MNI) space, and resampled with voxel size 2 mm x 2 mm x 2 mm. Spatial smoothing was then applied with full-width at half-maximum (FWHM) Gaussian kernel size 8 mm x 8 mm x 8 mm. The brain region of functional data was extracted using a segmented anatomy data as a mask image with respect to gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). To filter the temporal time series at each voxel, we used a discrete cosine transform (DCT) filter with a cutoff frequency of 1/128 Hz, which is an appropriate range of frequency for resting-state data, which show low frequency oscillations with the range of 0.0 – 0.1 Hz, in general. From each time series, temporal DC components have been removed. These preprocessing steps were conducted using SPM8.

3.3.2 Group Sparse Dictionary Learning

For group analysis, 3-D coordinates for the segmented GM, WM, and CSF from each individual data are compared, and the corresponding voxel coordinates common across all subjects in a group are extracted to build reference masks for group maps of GM, WM, and CSF, respectively (See Figure 3.1 for the case of training data from GM data). Then, time series that correspond to the group GM, WM, and CSF maps are collected from each individual data and then concatenated together to build a group training data set for sparse dictionary learning. To reduce the computational burden of a sparse dictionary learning algorithm, rather than initializing the dictionary using the whole data, we choose a subset of the voxels by downsampling by a factor of 4. In this case, a good choice of initial dictionary can make a sparse dictionary learning algorithm converge fast and avoid local minimizer. Since the variation due to background noise should be smaller than those from true activations, we conjecture that the time series with large variances can constitute a good initial dictionary. Hence, we sorted the time traces from all voxels corresponding a group mask with respect to their $l_1$ norm, and then choose the time traces with largest $l_1$ norms as an initial dictionary. Note that the reason to use $l_1$ norm rather than $l_2$ is to deal with outliers that are often encountered using real data sets.

Figure 3.1: Training data extraction for sparse dictionary learning from individual data sharing a same a group gray matter mask.
As the sparse coding step makes each dictionary atom has its own voxel map, this step can also be considered as clustering procedure in brain functional network perspective as in Figure 3.2. Each voxel map can be thought of same community sharing mostly same information ($d_n$) flow in a brain. It is also interpreted in a voxel based point of view such that signal at each voxel is sparse combination of global dictionary atoms which can be thought that a voxel belongs to few community not all in brain network structures, and common dictionary atom across different voxels represents that those voxels are bound with the community by sharing the common dictionary atom as in Figure 3.3.

Figure 3.2: By sparse coding step, for each dictionary atom, its own voxel map is obtained. And then, each own voxel map can be thought of same community sharing mostly same information (the dictionary atom) flow in a brain.

Figure 3.3: Signal at each voxel is sparse combination of global dictionary atoms which can be thought that a voxel belongs to few community not all in brain network structures, and common dictionary atom across different voxels represents that those voxels are bound with the community by sharing the common dictionary atom.

Sparse dictionary learning is then performed to update the dictionary and the coefficient by using the training data set and the initial dictionary mentioned above. We used the K-SVD algorithm [32] for the dictionary learning step with following optimized parameters affecting a lot to the results.

\[
y_m = 3d_1 + d_6 + d_{19} \quad D_{l_2} = [d_1, d_6, d_{19}]
\]

\[
y_n = 4d_2 + d_6 + d_{16} \quad D_{l_2} = [d_2, d_6, d_{16}]
\]

$d_1, d_2$: Information flow in community 1, 2

$d_6, d_{16}, d_{19}$: Long range information flow
**Number of dictionary atoms selection**

Similar to the blind source separation, the determination of the number of atoms in a dictionary is an important issue. Since it represents the number of linearly independent temporal dynamics across whole brain, we use an eigen-spectrum based approach to identify the number of components. More specifically, from the concatenated data set $Y$ across voxels that corresponds to a group GM map, its singular value decomposition is plotted, a signal subspace dimension is calculated and used as the number of atoms in the dictionary. More specifically, as shown in Figure. 3.4, the singular values rapidly decrease and after 20 components the values are nearly the same. Hence, 20 components should be the signal subspace dimension. The insets in Figure. 3.4 illustrate estimated DMN patterns using the proposed method with the corresponding number of dictionary atoms. Based on the DMN patterns as well as the eigen-spectrum in Figure. 3.4, the case with the dictionary atom of 20 depicted the best DMN patterns. Therefore, we choose the number of dictionary atoms as 20 for our resting-state fMRI data.

![Figure 3.4: The choice of the number of dictionary atoms using eigenspectrum of covariance matrix of normal group data.](image)

**Sparsity level selection**

Another important hyper-parameter that we need to estimate for our method is the sparsity level $k = |I_n|$. In our previous work for individual analysis [24], we proposed the minimum description length (MDL) as a criteria in deciding the sparsity level. Hence, we conducted the same procedure for the group data, which results are shown in Table 3.1. In general, a bigger sparsity level provides more fragmented DMN structures, so the different sparsity level for each group may provide different topological structures of the DMN network. Since one of our goals is to provide group-wise comparison, we therefore use the sparsity level $k = 3$ across all groups since it is the minimum number of sparsity among different groups and provide least fragmented DMN across all groups.

**Table 3.1: Sparsity level estimation using MDL criterion for each group data.**

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>MCI</th>
<th>CDR 0.5</th>
<th>CDR 1.0</th>
<th>CDR 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated sparsity</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
3.3.3 Group Estimation and Inference

Once the group sparse dictionary learning has been completed for each group, we use the estimated global dictionaries to identify the local network structures across the whole voxels in GM, WM, and CSF areas. Recall that this is an additional necessary step since the group sparse dictionary learning was performed using the downsampled data set to reduce the computational complexity, but we need to identify the network structures across all voxels in GM, WM, and CSF. This step can be performed as follows. First, we identify the group-wise coordinates for GM, WM, and CSF masks, then apply one more sparse coding using the global dictionary that has been estimated using down-sampled data as illustrated in Figure 3.1. This provides a spatially varying design matrix \( \{D_{in}\}_{n=1}^{N} \) for every voxels in GM, WM, and CSF areas. Then, the spatially varying design matrix \( \{D_{in}\} \) are decomposed into individual design matrices \( \{D_{in}^{(il)}\}_{n,i,l} \), whose regressors are normalized to have the same magnitude value 1. Then, the final regression coefficients are estimated by using least square method with respect to the design matrices. Then, each individual summary statistics are combined to obtain a group activation map. These procedures are summarized in Figure 3.5.

Figure 3.5: Group activation detection using a learned group dictionary.
Chapter 4. SPARSE-SPM: A new toolbox for a data-driven group fMRI analysis

A new “SPARSE-SPM Toolbox” software package is developed for the proposed data-driven group fMRI analysis. It uses few functions from SPM 8 package (Wellcome Department of Cognitive Neurology, London, UK, http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). It runs under MATLAB (Mathworks, Natick, MA, http://www.mathworks.com) environment. Based on the data-driven sparse general linear model (GLM) and statistical analysis, SPARSE-SPM provides statistical analysis to extract activation maps as well as an estimated spatially adaptive design matrix. Figure 4.1 shows a main panel of the toolbox software. It is composed of three main parts which are “Preprocessing and Data load”, “Dictionary Learning and Estimate” and “Inference and Results”.

In preprocessing step, SPARSE-SPM adopts SPM’s canonical preprocessing steps and provides the same function of them for users’ convenience. It includes spatial preprocessing, such as realignment, segmentation, normalization and smoothing, and temporal filtering to remove noise lying in low frequency bands, such as scanner drifts, cardiac and respiratory artifacts, from the data. The highpass filter is implemented using a discrete cosine transform (DCT) basis set.

In data load step as in Figure 4.2, SPARSE-SPM firstly collects required values to be analyzed and finds common voxels among multiple subjects in a group and then concatenates each subject data including normalization and DC removal steps in temporal way along with the common voxels. Note that DC component in each subject data is removed to avoid DC-bias effects. In sparse dictionary learning step, using user-defined parameters and conditions, K-SVD algorithm is used to train a global dictionary from the measurement data. Particularly, sparsity level can be determined by both MDL criterion and user-definition.

In estimate step, SPARSE-SPM estimates a spatially adaptive design matrix for each voxel by back-sparse-coding using the trained global dictionary. And then, it computes each summary statistics subject by subject with the individual dictionary from the trained global concatenated dictionary.

In inference step, SPARSE-SPM works to find value of test statistics for each voxel to convince of certain activation area statistically. We applied a mixed model to our methods and it turned out that various test statistics can be explained and performed in the unified mixed model concept together with ANOVA. And then, it computes $F$-statistics for every voxels with respect to each dictionary component among sparsely learned global dictionary and performs ANOVA to reveal the difference among groups. It also calculates threshold values of $F$-statistics with a given $p$-value by using uncorrected method or random field theory as suggested in [30].

In result step, the resulting $F$-map or ANOVA map is loaded to overlay them on the T1 template images as in Figure 4.3.

Sparse SPM has been developed and tested with the system environment shown in Table 4.1. Mostly, sparse SPM will work on any computer with MATLAB 7 or higher with approximately 2.0 GB RAM. Note that the computational time for our analysis took approximately 50 minutes per our current data for each subject with parameter settings above.
Figure 4.1: Main panel of the SPARSE-SPM

Figure 4.2: Data Load panel of the SPARSE-SPM

Table 4.1: System environment in developing and testing of Sparse SPM

<table>
<thead>
<tr>
<th></th>
<th>Pentium(R) Dual-Core CPU 2.50 GHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>4 GB RAM</td>
</tr>
<tr>
<td>Operating System</td>
<td>Windows(R) 7 Enterprise</td>
</tr>
<tr>
<td>Software</td>
<td>Matlab(R) 7 and SPM 8</td>
</tr>
</tbody>
</table>
Figure 4.3: Results viewer of the SPARSE-SPM
Figure 4.4: Block diagram of the proposed method
The overall flowchart of sparse SPM is given in Figure. 4.4.
The Sparse SPM toolbox will be available for download in authors's homepage
(http://bisp.kaist.ac.kr/SparseSPM/).
Chapter 5. Results

5.1 Evaluation of the Proposed Method

5.1.1 Training Data set in Sparse Dictionary Learning

We first conducted the sparse dictionary learning by using three different training data sets obtained from GM, GM+WM, and GM+WM+CSF, respectively. As shown in Figure. 5.1, the results using training data from GM show most intact DMN pattern for the normal subject group. Considering the fact that neural activation occurs mostly on gray matters that contains neural cell bodies, this results confirms that training using GM can extract the real neural dynamics, rather than other physiological noises. The residual errors after sparse dictionary learning for every case and group are shown in Table 5.1. This also indicates that the smallest residual errors are shown in the case of GM training dataset.

<table>
<thead>
<tr>
<th></th>
<th>NL</th>
<th>MCI</th>
<th>CDR 0.5</th>
<th>CDR 1.0</th>
<th>CDR 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>0.7211</td>
<td>0.7340</td>
<td>0.7404</td>
<td>0.7418</td>
<td>0.7573</td>
</tr>
<tr>
<td>GM + WM</td>
<td>0.7556</td>
<td>0.7679</td>
<td>0.7762</td>
<td>0.7851</td>
<td>0.7669</td>
</tr>
<tr>
<td>GM + WM + CSF</td>
<td>0.7636</td>
<td>0.7787</td>
<td>0.7843</td>
<td>0.7920</td>
<td>0.7838</td>
</tr>
</tbody>
</table>

5.1.2 Initial Dictionary in Sparse Dictionary Learning

Next, we investigate the choice of initial dictionary in sparse dictionary learning stage. In our previous study [24], the initial dictionary was chosen by taking data itself in sequential order. However, a group DMN pattern for normal subject group extracted using an initial dictionary proposed by Lee et al. (2011) [24] often fails to identify the activation in medial prefrontal cortex (MPFC) area, whereas our new initialization with the data with largest $l_1$ magnitude results in clear DMN patterns as shown in Figure. 5.2.

5.1.3 Number of iterations in Sparse Dictionary Learning

We next investigated the convergence of the K-SVD sparse dictionary learning algorithm. As shown in Figure. 5.3, after 5 iterations of K-SVD, the residual error during sparse dictionary learning is no more decreased, which indicates the convergence. Another interesting observation from Figure. 5.3 is that the resulting residual error is smallest for the case of normal group, whereas they increases with the progression of the AD disease. As we fixed the number of atoms in a dictionary across all group, the results implies that with the progression of the AD, there exists more irregular temporal dynamics within a brain that cannot be accounted for by the fixed number of atoms in a dictionary. As DMN turns out the most influential network in learned dictionary, this also indicates that temporal dynamics in the raw data explained by DMN becomes weaker as AD progression, which may reflect the weakening of the DMN.
Figure 5.1: The DMN patterns from different training dataset for sparse dictionary learning step.
\( p < 0.001, \text{ DF: } p_1, q - k \text{ for the first three rows } / \text{ DF: } p_1, M - k \text{ for the last three rows} \)
Figure 5.2: Although overall converged error was slightly higher in the most variable components case as in Figure 5.2, the DMN pattern was observed clearer in the case of initial dictionary of the most variable components than in the case of the first few components from the data itself. ($p < 0.001$, DF: \(p_1, q - k\) for the first two rows / DF: \(p_1, M - k\) for the last two rows)
Figure 5.3: (a) Residual error after K-SVD training. (b) The residual errors after K-SVD training reflect the progression of AD disease (**: p < 0.00001, *: p < 0.0001)
5.2 Comparative Studies

5.2.1 Extracted Default Mode Network

Comparative analyses of seed based analysis, ICA, and the proposed method, were conducted among normal, MCI, CDR 0.5, CDR 1.0, and CDR 2.0. As in Figure 5.4, the DMN patterns can be extracted by any method, however the noticeable diminishing of DMN pattern along with AD progression can be clearly seen by the proposed method while hardly distinguishable changes were acquired with other methods. Especially, the decreasing tendency of area PCC from MCI to CDR 2.0 are distinctly observed in 3D visualization of the DMNs (Figure 5.5).

Figure 5.4: The results of extracted DMN map using seed based analysis (first row), ICA (second row), the proposed method with $df(p_1, q-k)$ (third row), and the proposed method with $df(p_1, M-k)$ (fourth row). $p < 0.001$ for the proposed methods.
5.2.2 Group Comparisons

After performing omnibus $F$-test to find statistically significant group effect on DMN, we performed $1 \times 2$ ANOVA to find the differences between adjacent groups. Figure 5.5 illustrates two representative slices. As shown in the figure, drastic changes in DMN and additional superior temporal and inferior frontal gyri are observed as AD progression.

5.2.3 Regression Analysis

To quantitatively analyze the correlation of the extracted DMN patterns with the AD progression, linear regression analysis was performed. A template of DMN was directly extracted from the results of our normal group data. From this, separate templates for MPFC, PCC, and R/LLP templates were extracted based on the MNI coordinates (MPFC: $Y > 12$ mm, PCC: $Y \leq 12$ mm and $-28$ mm $\leq X \leq 32$ mm, RLP: $Y \leq 12$ mm and $X < 32$ mm, LLP: $Y \leq 12$ mm and $X > -28$ mm, and R/LLP: $Y \leq 12$ mm and $X > -28$ mm and $X < 32$ mm) considering our data results and a priori region of interests of
DMN [36]. Our regression analysis clearly shows statistically significant decreases of activation for all areas of DMN as the disease progression as in Figure. 5.7.

5.3 Reproducibility of the results

To confirm the reproducibility of our proposed method, we conducted the analysis with the randomly divided two groups within normal group. The group results of several times of such random division of normal group consistently shows DMN pattern in the case of DF \((p_1, M - k)\) and one of them is shown in Figure. 5.8. Hence, the reproducibility check of our proposed method has been made by obtaining DMN pattern in this random division case.
Figure 5.7: Linear regression analyses. (a) DMN ($R^2 = 0.7740, p = 0.0491$), (b) MPFC ($R^2 = 0.7021, p = 0.0764$), (c) PCC ($R^2 = 0.6869, p = 0.0828$), (d) RLP ($R^2 = 0.7951, p = 0.0421$), (e) LLP ($R^2 = 0.7906, p = 0.0436$), (f) R/LLP ($R^2 = 0.7953, p = 0.0420$), (g) DMN ($R^2 = 0.6152, p = 0.1163$), (h) MPFC ($R^2 = 0.6876, p = 0.0825$), (i) PCC ($R^2 = 0.5763, p = 0.1367$), (j) RLP ($R^2 = 0.5906, p = 0.1290$), (k) LLP ($R^2 = 0.5814, p = 0.1339$), and (l) R/LLP ($R^2 = 0.5892, p = 0.1297$) (DF: $p_1, M - k$ for the first two rows / DF: $p_1, M - k$ for the last two rows)
Figure 5.8: The DMN patterns are observed in the case of random division of normal group data in the case of DF \((p_1, M - k)\). \((p < 0.001, \text{ DF: } p_1, q - k \text{ for the left image / DF: } p_1, M - k \text{ for the right image for each comparison})\)
Chapter 6. Conclusion

In this paper, we developed a unified mixed-model for group sparse dictionary learning and inference for resting state fMRI analysis. Unlike the ICA methods, the new algorithm exploits that a temporal dynamics at each voxel can be represented as a sparse combination of global dynamics thanks to the property of small-worldness of brain network\([37, 38, 39]\). Especially, the sparse coding step in sparse dictionary learning step of our proposed method makes this method available for pacellation-free brain functional connectivity analysis solely based on the graph theoretical approach. And then, the simple group dictionary learning was made upon the reasonable assumption that the network structure in a group are similar which can be accepted without loss of generality. Using mixed model with ANOVA framework, we developed an unified framework to perform various statistical tests in a simple manner. We compared and validated our tools with the existing seed-based and ICA approaches for normal, MCI and Alzheimer’s disease with different disease stage. The results indicated that DMN network extracted using our method show clear diminishing patterns in DMN activation maps with the progression of disease which closely correlate with the recent clinical finding as in \([40]\). This indicates that the proposed tool has its strength not only in avoiding independence assumption which is paradox in brain’s functional connectivity research based on certain dependency among signals, but also in having great potential for resting state fMRI analysis.
Appendices
Using the matrix inversion lemma, we have

\[ V^{-1} = (ZGZ' + R)^{-1} = R^{-1} - R^{-1}Z(G^{-1} + Z'R^{-1}Z)^{-1}Z'R^{-1}. \]

Hence,

\[
Z'V^{-1} = Z'(R^{-1} - R^{-1}Z(G^{-1} + Z'R^{-1}Z)^{-1}Z'R^{-1}) \\
= (I - Z'R^{-1}Z(G^{-1} + Z'R^{-1}Z)^{-1})Z'R^{-1} \\
= (I - (Z'R^{-1}Z + G^{-1} - G^{-1})(G^{-1} + Z'R^{-1}Z)^{-1})Z'R^{-1} \\
= G^{-1}(G^{-1} + Z'R^{-1}Z)^{-1}Z'R^{-1} \\
= ((Z'R^{-1}Z)^{-1} + G)^{-1} (Z'R^{-1}Z)^{-1}Z'R^{-1},
\]

which leads us to

\[ Z'V^{-1}Z = ((Z'R^{-1}Z)^{-1} + G)^{-1}. \]

Therefore, for \( X = ZX_0 \), we have

\[ X'V^{-1}X = X'V^{-1}V^{-1}X_0 = X'V^{-1}X_0 \]

where

\[ V_G = (Z'V^{-1}Z)^{-1} = (Z'R^{-1}Z)^{-1} + G \]
Chapter B. Equivalence

Consider the following form of covariance matrices:

\[
V = ZGZ' + R, \quad (B.1)
\]
\[
\Omega = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} \quad (B.2)
\]

where \(R, G\) and \(V\) are assumed invertible. Using the definition of \(V\) and \(\Omega\) in Eqs. (B.1) and (B.2), it is easy to show

\[
\Omega X = 0, \quad X'\Omega = 0, \quad P_X\Omega = \Omega, \quad P_X V\Omega = I \quad (B.3)
\]

Now, we show that \(\Omega = (P_XV P_X\dagger)\dagger\). To this, we need to show the following

\[
P_XV P_X\Omega P_XV P_X = P_XV P_X\dagger
\]
\[
\Omega P_XV P_X\Omega = \Omega
\]
\[
(P_XV P_X\dagger)\dagger = P_XV P_X\dagger
\]
\[
(\Omega P_XV P_X\dagger)\dagger = \Omega P_XV P_X\dagger
\]

These are straightforward results using the properties in Eq. (B.3). Now, as \(P_X = P_Q\), we have \((P_XV P_X\dagger)\dagger = (P_QV P_Q)\dagger = Q(Q'VQ)^{-1}Q'\). Therefore, we have

\[
Q(Q'VQ)^{-1}Q' = (P_XV P_X\dagger)\dagger
\]

Now, we are ready to prove the main result on the equivalence. As we are only interested in estimable function, we assume \(L' \subset R(X')\) [41]. Then, the test statistics is

\[
S = \frac{\alpha' C' (C' (X'X)^{-1} C) \alpha}{\hat{\sigma}^2 \text{rank}(C)}
\]
\[
= \frac{y'C' (X'X)^{-1} C (X'X)^{-1} C (X'X)^{-1} X'y}{\hat{\sigma}^2 p_1}
\]

From the definition of generalized inverse, we can find the full rank matrix \(X^*\) such that \(R(X^*) = R(X)\) and \((X'X)^{-1} = (X'X^*)^{-1}\). Therefore, using the definition Eq. (B.2), we can define \(\Omega\) such that

\[
\Omega = (X'^* X^*)^{-1} - (X'^* X^*)^{-1} C' (C' (X'X)^{-1} C) (X'X)^{-1} C (X'^* X^*)^{-1}
\]
\[
= (P_{X^*} (X'^* X^*)^{-1} P_{X^*})\dagger
\]
Then, we have

\[
S = \frac{y'X' \left( (X'X)^{-1} - \Omega X'X \right) X'y}{\sigma^2 p_1}
\]

\[
= \frac{y'X' \left( (X'X)^{-1} - (P^\perp C'X^2P^\perp C' )^\dagger \right) X'y}{\sigma^2 p_1}
\]

\[
= \frac{y'X' \left( (X'X)^{-1} - Q(QX'XQ)^{-1}Q \right) X'y}{\sigma^2 p_1}
\]

\[
= \frac{y'(P_X - P_{X_0})y}{\sigma^2 p_1}
\]

\[
= \frac{y'(P^\perp_{X_0} - P^\perp_X)y}{\sigma^2 p_1} N - p
\]

\[
= \frac{y'P^\perp_X y}{\sigma^2 p_1} N - p
\]

where \( X_0 = XQ \) denotes the reduced model by excluding the effect of the contrast matrix \( C \). This concludes the proof.
References


Summary

Group Sparse Dictionary Learning and Inference for Resting-state fMRI Analysis of Alzheimer’s Disease

최근 신경과학 분야에서 주목받고 있는 뇌의 기능적 연결성 연구와 알츠하이머 질병 (Alzheimer’s Disease) 상태와 밀접한 연관성을 보여주는 휴식상태 (Resting-state) 뇌기능자 가능 영상 (fMRI) 데이터에서의 능성능영역간 기분상태네트워크 (Default Mode Network) 분석에 있어서 기존에 널리 쓰이고 있는 독립성분석법 (Independent Component Analysis)은 신호간의 독립성 (Independence)에 기반하고 있어 뇌 내부 신호간 연관성을 가정할 수 밖에 없는 뇌의 기능적 연결성 연구에 적응함에 있어 논리적인 모순을 가질 수 밖에 없었다. 더욱이, 최근 연구에 의하면 독립성분석법이 동차발생적인 다양한 능성 패턴들의 독립성을 보장하지 않으며, 그 알고리즘의 성공률이 신호의 독립성보다 희소성 (Sparsity)에 더욱 의존하고 있음이 밝혀졌다. 이는 희소코딩 (Sparse Coding)을 통한 V1 세포들의 시각수용영역 표현, 중간관자엽 (Medical temporal lobe) 전기생리학적 실험결과 등의 일련의 생물학적 발견과도 일치한다. 또한, 그룹 데이터 분석에 있어 기본상태네트워크와 같은 관심있는 해당시간축 신호 변화 모델이 개별별로 다르기 때문에 개인 종속적인 시간축 신호 변화 모델을 구하려고 동시에 그 개인별 특성 변화 모델을 이용한 그룹 추론을 하는 것이 매우 복잡하였다. 본 논문은 이런 기존의 데이터기반 분석방법들의 단점을 극복하고 그룹 차원에도 유연하게 적용가능한 신호의 희소성에 기반한 통합된 혼합 모델 (Mixed Model)을 이용한 그룹 희소사전학습법 (Group Sparse Dictionary Learning)을 제시하였으며, 제안 방법의 검증을 위하여 다수의 실제 뇌기능자 가능영상 휴식상태 데이터를 정상 집단부터 경도 인지 장애 (Mild Cognitive Impairment), 그리고 임상적 및 정신 신경 시스템 (Clinical Dementia Rating) 0.5, 1.0, 2.0 단계까지 확보하여 분석한 그룹별 기분상태네트워크 활성화 정도가 실제 질병의 증상을 정확히 반영할 수 있도록 기분상태네트워크의 활성화 정도가 약간은 기존의 의학계의 연구 결과와 잘 일치함을 보였다. 이는 본 제안 방법이 향후 알츠하이머 질병 증상의 바이오 마커로서의 역할을 기대해 볼 수 있는 대목이고, 기존 데이터기반 대비 개념적 모순없이 생물학 및 의학계의 발견과도 잘 일치하는 새로운 차원의 분석기법의 초석이 될 것임을 기대할 수 있다.