Electromagnetic Brain Imaging: A Bayesian Perspective

東京医科歯科大学　先端技術医療応用学講座

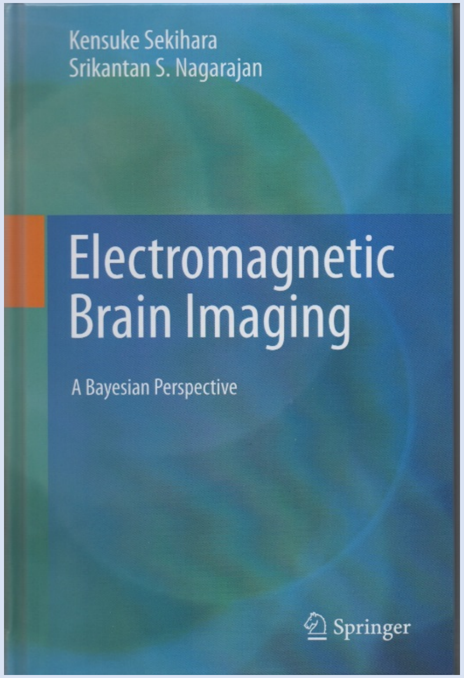
株式会社　シグナルアナリシス

関原謙介

生体磁気データの処理に有用な2つのベイズ推定手法を紹介する．

1）スパースベイズ推定（Sparse Bayes Learning）を紹介し，スパースベイズの仮定から（一見無関係に思える）他の逆問題解法が導けることを示す．

2）ベイズ因子分析を紹介し，センサーノイズ・干渉信号除去，信号源推定への応用を議論する．



生体磁気逆問題

線型方程式: $\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{x}(t) + \boldsymbol{\varepsilon}$ を解く

$\boldsymbol{y}(t)$: センサーデータ

$\boldsymbol{x}(t)$: ボクセルデータ

$\boldsymbol{H}$:リードフィールド行列

$\boldsymbol{\varepsilon}$: ノイズ

最尤推定解: $\boldsymbol{\varepsilon} \sim N(\boldsymbol{\varepsilon} \mid 0, \sigma^2 \boldsymbol{I})$ の仮定のもとで:

コスト関数: $F = \left\| \boldsymbol{y}(t) - \boldsymbol{H}\boldsymbol{x}(t) \right\|^2 + \gamma \left\| \boldsymbol{x}(t) \right\|^2$ を最小とする$\boldsymbol{x}(t)$として求まる.

ミニマムノルムの解

$$\hat{\boldsymbol{x}}(t) = \boldsymbol{H}^T (\boldsymbol{H}\boldsymbol{H}^T + \gamma \boldsymbol{I})^{-1} \boldsymbol{y}(t)$$

正則化(regularization)

L_2-正則化 ミニマムノルム解とも呼ばれる.

Sparse Bayesian Learning

確率モデル

事前分布: $p(x_j \mid \alpha_j) = N(x_j \mid 0, \alpha_j^{-1})$

ノイズ: $p(\boldsymbol{\varepsilon}) = N(\boldsymbol{\varepsilon} \mid 0, \beta^{-1}\boldsymbol{I})$

ベイズ推定解

$$\overline{\boldsymbol{x}} = \beta(\boldsymbol{\Phi} + \beta\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{y}$$

ここで $\boldsymbol{\Phi} = diag(\boldsymbol{\alpha}) = diag(\alpha_1, \ldots, \alpha_N)$

ハイパーパラメータ $\boldsymbol{\alpha}$ は以下の周辺尤度の最大化により求める

周辺尤度: $p(\boldsymbol{y} \mid \boldsymbol{\alpha}) = \int p(\boldsymbol{y} \mid \boldsymbol{x})p(\boldsymbol{x} \mid \boldsymbol{\alpha})d\boldsymbol{x}$

コスト関数: $-2\log p(\boldsymbol{y} \mid \boldsymbol{\alpha}) = \log|\boldsymbol{\Sigma}_y| + \boldsymbol{y}^T\boldsymbol{\Sigma}_y^{-1}\boldsymbol{y}$

ここで $\boldsymbol{\Sigma}_y = \beta^{-1}\boldsymbol{I} + \boldsymbol{H}\boldsymbol{\Phi}^{-1}\boldsymbol{H}^T$ である.

この最小化は，EMアルゴリズムやConvexity-basedアルゴリズムなどにより実行できる.

Bayesian L_2-regularized minimum-norm solution

事前分布をわずかに変更すれば:

事前分布: $p(x_j) = N(x_j \mid 0, \alpha_j^{-1}) \Rightarrow N(x_j \mid 0, \alpha^{-1})$

ノイズ: $p(\boldsymbol{\varepsilon}) = N(\boldsymbol{\varepsilon} \mid 0, \beta^{-1}\boldsymbol{I})$

ベイズ推定解

$$\overline{\boldsymbol{x}}(t) = \beta \boldsymbol{\Gamma}^{-1}\boldsymbol{H}^T\boldsymbol{y}(t) = \underbrace{\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{H}^T + \frac{\alpha}{\beta}\boldsymbol{I}\right)^{-1}\boldsymbol{y}(t)}_{L_2\text{-正則化 ミニマムノルム解}}$$

$$\boldsymbol{\Gamma} = \alpha\boldsymbol{I} + \beta\boldsymbol{H}^T\boldsymbol{H}$$

ベイズ的な導出では，最適な正則化パラメータ $\dfrac{\alpha}{\beta}$ を求める手法が組み込まれている

例：EMアルゴリズムによる
更新式

$$\hat{\alpha}^{-1} = \overline{\boldsymbol{x}}^T\overline{\boldsymbol{x}} + \mathrm{tr}(\boldsymbol{\Gamma}^{-1})$$

$$\hat{\beta}^{-1} = \left\|\boldsymbol{y} - \boldsymbol{H}\overline{\boldsymbol{x}}\right\|^2 + \mathrm{tr}(\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{\Gamma}^{-1})$$

L_1 regularized minimum-norm solution

L_1ノルム解

$$\hat{\mathbf{x}}(t) = \arg\min F : \ F = \left\| \boldsymbol{y}(t) - \boldsymbol{H}\boldsymbol{x}(t) \right\| + \xi \sum_{j=1}^{N} | x_j(t) |$$

L$_1$ノルム解もスパースベイズ推定の枠組みから導出できる.

事後分布を以下から導出する:

$$p(\boldsymbol{x} \mid \boldsymbol{y}) = \int p(\boldsymbol{y} \mid \boldsymbol{x}) p(\boldsymbol{x} \mid \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) d\boldsymbol{\alpha} = p(\boldsymbol{y} \mid \boldsymbol{x}) \int p(\boldsymbol{x} \mid \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) d\boldsymbol{\alpha}$$

無情報事前分布 $p(\boldsymbol{\alpha}) = const$ の仮定では上の積分は計算できないので,
$p(\boldsymbol{\alpha})$にGamma分布を仮定し, 積分の後でGamma分布のパラメータa,bをゼロとする.

このとき, $p(\boldsymbol{x}) = \int p(\boldsymbol{x} \mid \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \rightarrow \prod_{j=1}^{N} | x_j |$ となるので

コスト関数: $-\log p(\boldsymbol{x} \mid \boldsymbol{y}) \rightarrow \beta \left\| \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x} \right\|^2 + \sum_{j=1}^{N} \log | x_j |$

を得る.（この解はL_1ノルム解と厳密に同じでは無いがほぼ等しい特性を示す.）

Adaptive Beamformer

スパースベイズ推定解は等価な空間フィルター形式で書き直せる

$$\overline{x}_j(t) = \alpha_j^{-1} \boldsymbol{l}_j^T \boldsymbol{\Sigma}_y^{-1} \boldsymbol{y}(t) \qquad \text{ただし,}\quad \boldsymbol{H} = [\boldsymbol{l}_1, \ldots, \boldsymbol{l}_N]$$

未知量 α_j を求めるのに α_j を非確率変数のように取り扱う.

まず, $\overline{x}_j(t)$ のパワーを計算する

$$\left\langle \overline{x}_j(t)^2 \right\rangle = \alpha_j^{-2} \boldsymbol{l}_j^T \boldsymbol{\Sigma}_y^{-1} \left\langle \boldsymbol{y}(t) \boldsymbol{y}^{\mathrm{T}}(t) \right\rangle \boldsymbol{\Sigma}_y^{-1} \boldsymbol{l}_j$$

サンプルデータ共分散 $\boldsymbol{R} = \left\langle \boldsymbol{y}(t) \boldsymbol{y}^{\mathrm{T}}(t) \right\rangle$ をモデルデータ共分散 $\boldsymbol{\Sigma}_y$ に等しいとおき,

ユニットゲイン制約: $\left\langle \overline{x}_j(t)^2 \right\rangle = \alpha_j^{-1}$ を仮定すれば,

Adaptive Beamformer: $\overline{x}_j(t) = \boldsymbol{w}^T \boldsymbol{y}(t)$: $\quad \boldsymbol{w}^T = \dfrac{\boldsymbol{l}_j^T \boldsymbol{R}^{-1}}{\boldsymbol{l}_j^T \boldsymbol{R}^{-1} \boldsymbol{l}_j}$ を得る.

Recursive Null-Steering (RENS) Beamformer

スパースベイズ推定解

$$\bar{x}_j(t) = \alpha_j^{-1} \boldsymbol{l}_j^T \boldsymbol{\Sigma}_y^{-1} \boldsymbol{y}(t)$$

$$\boldsymbol{\Sigma}_y = \beta^{-1} \boldsymbol{I} + \boldsymbol{H}\boldsymbol{\Phi}^{-1}\boldsymbol{H}^T$$

ここで，ノイズパワー β^{-1} をゼロとおき，$\left\langle \bar{x}_j(t)^2 \right\rangle = \alpha_j^{-1}$ を仮定すれば

$$\bar{x}_j(t) = \frac{\boldsymbol{l}_j^T \boldsymbol{\Sigma}_y^{-1} \boldsymbol{y}(t)}{\boldsymbol{l}_j^T \boldsymbol{\Sigma}_y^{-1} \boldsymbol{l}_j}$$

$$\boldsymbol{\Sigma}_y = \boldsymbol{H}\boldsymbol{\Phi}^{-1}\boldsymbol{H}^T = \boldsymbol{H} \begin{bmatrix} x_1(t)^2 & & 0 \\ & \ddots & \\ 0 & & x_N(t)^2 \end{bmatrix} \boldsymbol{H}^T = \sum_{j=1}^{N} x_j(t)^2 \boldsymbol{l}_j \boldsymbol{l}_j^T$$

を得る．これはRENS Beamformerの更新式に等しい．

Kumihashi I, and Sekihara K (2010) IEEE Trans. Biomed Eng.57:1358–65.

Bayesian Factor Analysis (BFA)

観測データ : $\boldsymbol{y}(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_M(t) \end{bmatrix}$,　因子活動 : $\boldsymbol{u}(t) = \begin{bmatrix} u_1(t) \\ \vdots \\ u_L(t) \end{bmatrix}$　$(M \gg L)$

$\boldsymbol{y}_k = \boldsymbol{y}(t_k),\ \ \boldsymbol{u}_k = \boldsymbol{u}(t_k)$ と表記

BFAデータモデル : $\boldsymbol{y}_k = \boldsymbol{A}\boldsymbol{u}_k + \boldsymbol{\varepsilon}$

- Mチャンネルの観測データを，各チャンネルで共通な活動$\boldsymbol{A}\boldsymbol{u}_k$ と，チャンネルに固有な活動 に分離する.
- 多数のセンサーで同時計測されたデータでは, $\boldsymbol{A}\boldsymbol{u}_k$を信号成分, $\boldsymbol{\varepsilon}$ をノイズ成分と解釈することができる.
- 信号成分は，(データ数Mよりはるかに) 少数(L)の因子活動の和として表わされる.

ベイズ因子分析では $\boldsymbol{y}_k\ (k = 1, \ldots, K)$から，混合行列$\boldsymbol{A}$，因子活動 $\boldsymbol{u}_k (k = 1, \ldots K)$ と, ノイズ$\boldsymbol{\varepsilon}$ の共分散行列を推定する.

確率モデル

事前確率分布: $p(\boldsymbol{u}_k) = N(\boldsymbol{u}_k \mid 0, \boldsymbol{I})$ ← 因子は独立と事前分布で仮定

ノイズ確率分布: $p(\boldsymbol{\varepsilon}) = N(\boldsymbol{\varepsilon} \mid 0, \boldsymbol{\Lambda}^{-1})$

事後確率分布: $p(\boldsymbol{u}_k \mid \boldsymbol{y}_k) = N(\boldsymbol{u}_k \mid \bar{\boldsymbol{u}}_k, \boldsymbol{\Gamma}^{-1})$ ← 因子の相関は事後分布の共分散行列に反映される

EMアルゴリズムにより因子活動 $\bar{\boldsymbol{u}}_k$ と混合行列 $\hat{\mathbf{A}}$ ノイズ共分散 $\hat{\boldsymbol{\Lambda}}^{-1}$ を求める

Eステップ更新式:

$$\boldsymbol{\Gamma} = \boldsymbol{I} + \boldsymbol{A}^T \boldsymbol{\Lambda} \boldsymbol{A}$$

$$\bar{\boldsymbol{u}}_k = \boldsymbol{\Gamma}^{-1} \boldsymbol{A}^T \boldsymbol{\Lambda} \boldsymbol{y}_k = (\boldsymbol{I} + \boldsymbol{A}^T \boldsymbol{\Lambda} \boldsymbol{A})^{-1} \boldsymbol{A}^T \boldsymbol{\Lambda} \boldsymbol{y}_k$$

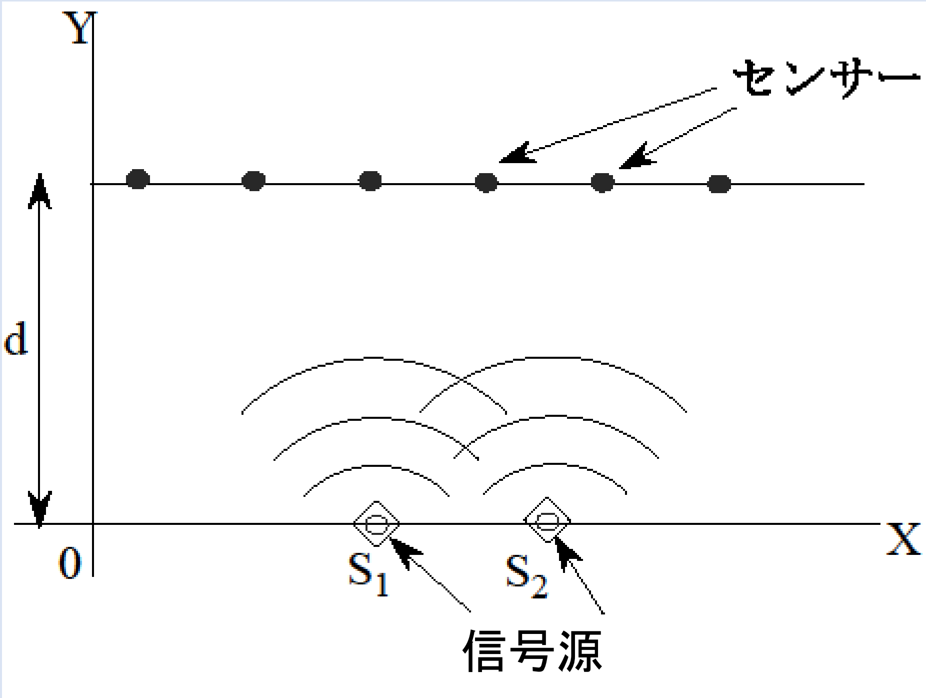
Mステップ更新式:

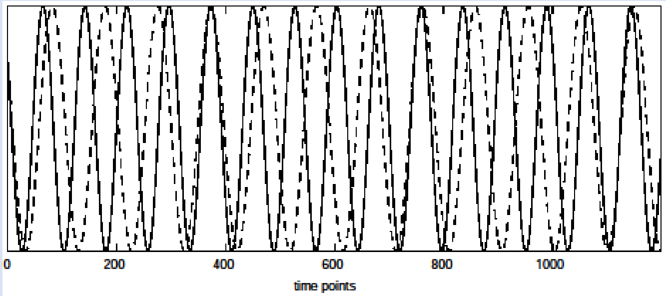
$$\hat{\mathbf{A}} = \boldsymbol{R}_{yu} \boldsymbol{R}_{uu}$$

$$\hat{\boldsymbol{\Lambda}}^{-1} = \frac{1}{K} \mathrm{diag}\left[\boldsymbol{R}_{yy} - \boldsymbol{A} \boldsymbol{R}_{uy} \right]$$

更新アルゴリズム終了時点での混合行列 $\hat{\mathbf{A}}$ と因子活動 $\bar{\boldsymbol{u}}_k$ を用いて計算した $\hat{\mathbf{A}}\bar{\boldsymbol{u}}_k$ が観測データに含まれる信号の推定解である.

コンピュータシミュレーション：1次元配列センサーアレイ





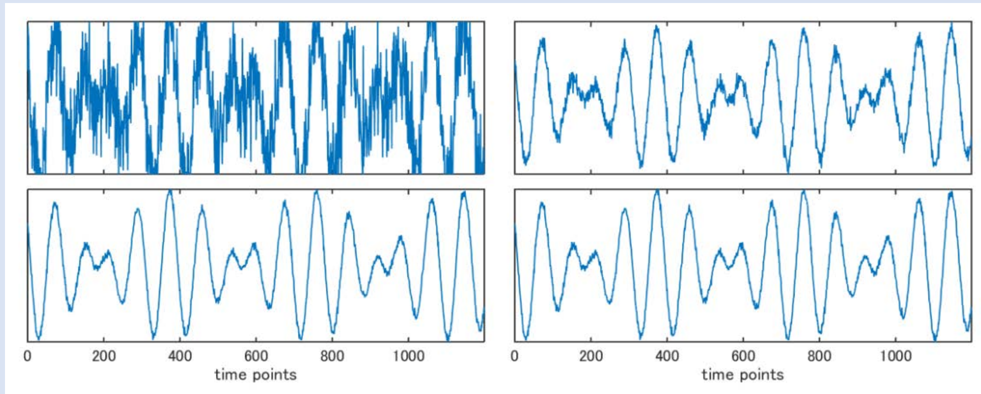
信号源タイムコース

2個の信号源が発する磁場信号を1次元配列のセンサーアレイ（300センサー）で計測

シミュレーションの結果

センサー出力

処理結果：
ファクター数2



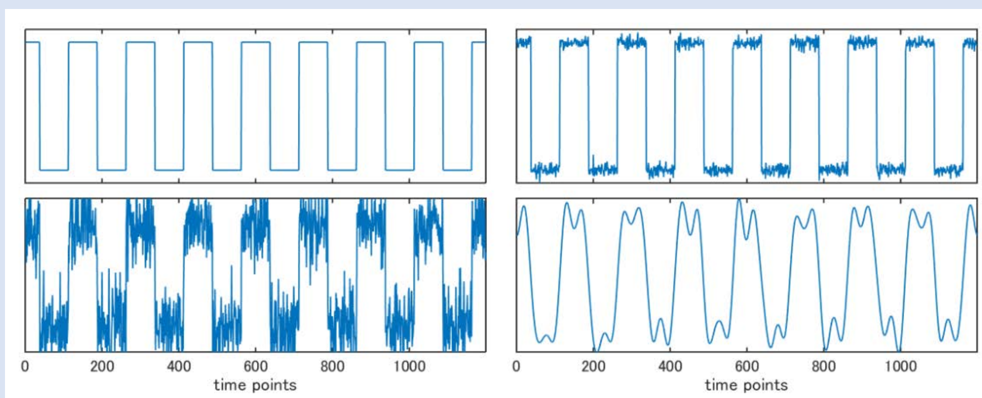
処理結果：
ファクター数20

変分ベイズ因子分析
ファクター数20

矩形波を用いたシミュレーションの結果

ソースに仮定した
波形

センサー出力



ベイズ因子分析結果

デジタルフィルター
による結果

干渉信号除去への拡張：Partitioned Factor Analysis (PFA)

PFAデータモデル：

$$\boldsymbol{y}_k = \boldsymbol{B}\boldsymbol{v}_k + \boldsymbol{\varepsilon}$$ コントロールデータ

$$\boldsymbol{y}_k = \boldsymbol{A}\boldsymbol{u}_k + \boldsymbol{B}\boldsymbol{v}_k + \boldsymbol{\varepsilon}$$ 計測データ

コントロールデータ（妨害信号のみを含むデータ領域）が入手できることが前提となる.

PFAアルゴリズム：

1) BFAアルゴリズムをコントロールデータに適用し，妨害信号に対する混合行列$\boldsymbol{B}$および を求める.

2) 計測データは

$$\boldsymbol{y}_k = \boldsymbol{A}\boldsymbol{u}_k + \boldsymbol{B}\boldsymbol{v}_k + \boldsymbol{\varepsilon} = [\boldsymbol{A}, \boldsymbol{B}]\begin{bmatrix}\boldsymbol{u}_k \\ \boldsymbol{v}_k\end{bmatrix} + \boldsymbol{\varepsilon} = \boldsymbol{C}\boldsymbol{z}_k + \boldsymbol{\varepsilon}$$

と表される．計測データに対しBFAを適用し，混合行列$\boldsymbol{A}$ とファクター活動 $\boldsymbol{u}_k, \boldsymbol{v}_k$を求める.

3) $\boldsymbol{A}\bar{\boldsymbol{u}}_k$ が信号成分推定結果である.

Nagarajan SS et al. (2007) Stat. Med. 26, 3886—3910

信号源推定への拡張：Saketini Algorithm

データモデル： $\boldsymbol{y}_k = \boldsymbol{L}(r)\boldsymbol{s}_k + \boldsymbol{B}\boldsymbol{v}_k + \boldsymbol{\varepsilon}$

$\boldsymbol{y}_k$ ： センサーデータ

$\boldsymbol{L}(\boldsymbol{r})$ ： 位置$\boldsymbol{r}$ におけるセンサーリードフィールド行列

$\boldsymbol{s}_k$ ： 位置$\boldsymbol{r}$ における信号源活動

$\boldsymbol{B}\boldsymbol{v}_k$ ： 位置$\boldsymbol{r}$ 以外の全ての位置における信号源活動のセンサーデータへの寄与

$\boldsymbol{\varepsilon}$ ： センサーノイズ

Saketini Algorithm （scanning algorithm)

$\boldsymbol{L}$ を与えて，以下のモデルに従い位置$\boldsymbol{r}$における信号源活動 $\boldsymbol{s}_k$ をBFAをセンサーデータ $\boldsymbol{y}_k$ に適用することにより求める.

$$\boldsymbol{y}_k = \boldsymbol{L}\boldsymbol{s}_k + \boldsymbol{B}\boldsymbol{v}_k + \boldsymbol{\varepsilon} = [\boldsymbol{L}, \boldsymbol{B}]\begin{bmatrix}\boldsymbol{s}_k \\ \boldsymbol{v}_k\end{bmatrix} + \boldsymbol{\varepsilon} = \boldsymbol{C}\boldsymbol{z}_k + \boldsymbol{\varepsilon}$$

Zumer JM *et al.* (2007) Neuroimage 37:102-15.

まとめ

- Sparse Bayesian Learningアルゴリズムを紹介し，Sparse Bayesian Learningの枠組みから，（一見無関係に見える）他の代表的な再構成手法が導かれることを示した．
- Bayesian Factor Analysisについて，このアルゴリズムがセンサーノイズ除去，アーチファクト除去，信号源再構成など多目的に用いることのできる有用なアルゴリズムであることを紹介した．


最後まで聞いていただき、ありがとうございました