

卒業論文

R による ICA 入門

宇佐美 敬允
北里大学理学部物理学科非線形研究室

2014 年 2 月 28 日

概要

本論文は、独立成分分析の初学者でも直感的に独立成分分析（ICA）を理解し、プログラミング言語である R を使って、独立成分分析で主に使われる fastICA のプログラムを組めるようになることを目標としたものである。また、プログラム中で使われている R の関数については、R の専門書やインターネットにあるサイトを参照していただきたい。

目次

1	導入	2
2	最適化理論	3
2.1	最急降下法	3
2.2	ニュートン法	3
2.3	制約条件つき最適化問題	4
3	主成分分析	5
3.1	主成分	5
3.2	第一主成分の導出	5
3.3	第 m 主成分の導出	6
3.4	1 固有ベクトルアルゴリズム	7
3.5	直交化	7
3.6	m 個の固有ベクトルアルゴリズム	8
3.7	白色化	8
4	情報理論	11
4.1	エントロピー	11
4.2	微分エントロピー	11
4.3	ネゲエントロピー	12
4.4	最大エントロピー法	12
4.5	ネゲエントロピーの近似	14
5	独立成分分析	16
5.1	独立成分分析の原理	16
5.2	ネゲエントロピーを尺度とした ICA	17
5.3	確率的最急降下法を用いた ICA アルゴリズム	17
5.4	fastICA アルゴリズム	18
5.5	一成分における fastICA アルゴリズム	18
5.6	対称的直交化	20
5.7	複数成分における fastICA アルゴリズム	21
6	ICA の応用	24
6.1	ICA による脳波データのノイズ除去	24

1 導入

独立成分分析 (ICA) とは、多変量のデータから、信号源の独立性以外の仮定を全く必要とせずに独立成分を推定する一つの手法である。基本的な ICA モデルとして以下を仮定する。

$$\mathbf{x} = A\mathbf{s}$$

ここで、 $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T$, $\mathbf{s}(t) = (s_1(t), \dots, s_n(t))^T$, $A, t \in t(1, 2, \dots)$ をそれぞれ、観測ベクトル、信号源ベクトル、 n 行 n 列の混合行列、また大文字 T を転置、小文字 t を 1~の時間の変数として与えられる。我々が知っているのは観測値ベクトル \mathbf{x} だけであり、混合行列 A と信号源ベクトル \mathbf{s} に関しては何も知らない。しかし、未知の信号源 \mathbf{s} に対して独立性のみを仮定すれば、観測値ベクトル \mathbf{x} から混合行列 A の逆行列 A^{-1} および信号源ベクトル \mathbf{s} を推定することができる。

基本的な ICA の原理を説明するためには、白色化 (第 3 章)

$$\mathbf{z} = V\mathbf{x}$$

と非ガウス性 (第 4 章)

$$\mathbf{y} = W\mathbf{z}$$

の説明が不可欠であり、信号源ベクトル \mathbf{s} は

$$\mathbf{s} = \mathbf{y} = A^{-1}\mathbf{x} = WV\mathbf{x}$$

により求めることができる。

2 最適化理論

一般的な統計的な推定は、損失関数あるいは目的関数の最適化に基づくことが多い。最適化とは関数の最大値あるいは最小値を見つけることで、損失関数および目的関数は、最適化を行う対象のことを呼ぶ。ICA および PCA の解は目的関数の最大化によって得られる。この章では、古典的な最適化の手法である最急降下法とニュートン法について説明する。

2.1 最急降下法

最急降下法とは、関数の最小化もしくは最大化の問題に対する最も古典的な手法である。最急降下法では、反復的に目的関数 $J(\mathbf{w})$ を最小化する。時刻 $t = 1, 2, \dots$ における更新則は、

$$\mathbf{w}(t) = \mathbf{w}(t-1) - \alpha(t) \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} \Big|_{\mathbf{w}=\mathbf{w}(t-1)}$$

で与えられる。初期値 $\mathbf{w}(0)$ から始められ、 $J(\mathbf{w})$ の勾配はその点で計算し、勾配ベクトルとは逆方向、つまり最急降下する方向に移動する。またパラメータ $\alpha(t)$ は勾配ベクトルの反対方向へ動く方向を決め、学習係数もしくは刻み幅と呼ばれる。更新則は反復が収束するまで更新され、ユーグリッド距離 $\|\mathbf{w}(t) - \mathbf{w}(t-1)\|$ が小さい許容範囲以下になった時、収束したと判定する。この更新則を略記法を用いて記述すれば、

$$\mathbf{w}(t) \leftarrow \mathbf{w} - \alpha \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}}$$

となる。また矢印は代入という意味である。

また最急降下法の弱点として、最急降下法は関数全体の極小値ではなく初期値に一番近い局所的な極小を求めることである。従って、このアルゴリズムの適用にはよい初期値を選ぶことである。また、ICA および PCA のように目的関数が観測データに依存するような時、用いる最急勾配法は確率的勾配降下法と呼ばれる。

2.2 ニュートン法

多変数の目的関数 $J(\mathbf{w})$ の \mathbf{w} の周囲でのテイラー級数展開は

$$J(\mathbf{w}') = J(\mathbf{w}) + \left[\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} \right]^T (\mathbf{w}' - \mathbf{w}) + \frac{1}{2} (\mathbf{w}' - \mathbf{w})^T \frac{\partial^2 J(\mathbf{w})}{\partial \mathbf{w}^2} (\mathbf{w}' - \mathbf{w}) + \dots$$

関数 $J(\mathbf{w})$ の最小値を探すために $\Delta \mathbf{w} = \mathbf{w}' - \mathbf{w}$ とおいて

$$J(\mathbf{w}') - J(\mathbf{w}) = \left[\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} \right]^T \Delta \mathbf{w} + 1/2 \Delta \mathbf{w}^T \frac{\partial^2 J(\mathbf{w})}{\partial \mathbf{w}^2} \Delta \mathbf{w}$$

を $\Delta \mathbf{w}$ に関して最小化するとその勾配は、 $\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} + \frac{\partial^2 J(\mathbf{w})}{\partial \mathbf{w}^2} \delta \mathbf{w}$ となり、勾配を 0 とおけば

$$\Delta \mathbf{w} = -\left[\frac{\partial^2 J(\mathbf{w})}{\partial \mathbf{w}^2}\right]^{-1} \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}}$$

が得られる。これより反復法は

$$\mathbf{w} \leftarrow \mathbf{w} - \left[\frac{\partial^2 J(\mathbf{w})}{\partial \mathbf{w}^2}\right]^{-1} \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}}$$

になる。このアルゴリズムをニュートン法と呼ぶ。

2.3 制約条件つき最適化問題

制約条件を取り入れる最も用いられている方法はラグランジュの未定乗数法である。ラグランジュ関数 $L(\mathbf{w}, \lambda_1, \dots, \lambda_k)$ は目的関数 $J(\mathbf{w})$ 、制約方程式 $H_i(\mathbf{w}) = 0 (i, \dots, k)$ 、ラグランジュ乗数 $\lambda_1, \dots, \lambda_k$ によって定義される。

$$L(\mathbf{w}, \lambda_1, \dots, \lambda_k) = J(\mathbf{w}) + \sum_{i=1}^k \lambda_i H_i(\mathbf{w})$$

ラグランジュ乗数の個数 k は制約条件式のスカラー式としての個数と同じである。ラグランジュ関数の勾配が 0 となる点が制約条件つき最適化問題の解として与えられる。

$$\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} + \sum_{i=1}^k \lambda_i \frac{\partial H_i(\mathbf{w})}{\partial \mathbf{w}} = 0$$

よくある制約条件の問題として \mathbf{w} のノルムが 1 であるという制約条件つき最適化問題がある。その場合はラグランジュ関数 $L(\mathbf{w}, \lambda_1)$ は

$$L(\mathbf{w}, \lambda_1) = \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} + \lambda_1 (\|\mathbf{w}\|^2 - 1)$$

となり、 \mathbf{w} に関する勾配は

$$\frac{\partial L(\mathbf{w}, \lambda_1)}{\partial \mathbf{w}} = \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} + 2\lambda_1 \mathbf{w}$$

が得られる。よって学習則は

$$\mathbf{w} \leftarrow \mathbf{w} + \alpha \left(\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} + 2\lambda_1 \mathbf{w} \right)$$

となり、いい収束を得るには適当な学習係数 α と適当な λ_1 の値を選ぶ必要がある。

\mathbf{w} のノルムが 1 であるような簡単な制約条件には射影法がはるかに簡単である。射影法は、 \mathbf{w} を \mathbf{w} と同次元の単位球面に直交射影させることである。つまり、先ほどの学習則は

$$\mathbf{w} \leftarrow \mathbf{w} + \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}}$$

$$\mathbf{w} \leftarrow \mathbf{w} / \|\mathbf{w}\|$$

と書き換えることができる。

3 主成分分析

この章では、独立成分分析の前処理として非常に有効である主成分分析の原理および主成分分析の解を求めるための確率的勾配降下法を用いたオフライン学習則の両方を述べる。主成分分析とは、多変量のデータに対して線形変換をすることによって、新しい座標成分を互いに無相関にするための数学的な手続きである。

3.1 主成分

PCA モデルを n 個の要素を持つ確率ベクトル \mathbf{x} 、またこの確率ベクトルは標本 $\mathbf{x}(1) \sim \mathbf{x}(T)$ が与えられているとする。

$$\mathbf{x} = (x_1(t), x_2(t), \dots, x_n(t))^T$$

最初に PCA の計算を簡略化するために、 \mathbf{x} に対して中心化を行う。

$$\mathbf{x} \leftarrow \mathbf{x} - E[\mathbf{x}]$$

\mathbf{x} の確率密度関数は一般には分からないので、平均 $E[\mathbf{x}]$ は得られている標本 $\mathbf{x}(1) \sim \mathbf{x}(T)$ によって

$$E[\mathbf{x}] \simeq \sum_{t=1}^T \mathbf{x}(t) / T$$

と推定される。以後、 \mathbf{x} の平均は 0 とする。

3.2 第一主成分の導出

z_1 は、 \mathbf{x} の要素 x_1, \dots, x_n の線形結合

$$z_1 = \sum_{k=1}^n w_{k1} x_k = \mathbf{w}_1^T \mathbf{x}$$

で表し、 n 次元ベクトル \mathbf{w}_1 の要素 w_{11}, \dots, w_{n1} はスカラーの係数で荷重と呼ばれ、 \mathbf{w}_1^T は \mathbf{w}_1 の転置を表し、 z_1 の分散 C_{z_1} が最大である時、第一主成分と呼ぶ。従って、PCA の目的関数 $J_1^{pca}(\mathbf{w}_1)$ は

$$J_1^{pca}(\mathbf{w}_1) = E[z_1^2] = E[(\mathbf{w}_1^T \mathbf{x})^2] = \mathbf{w}_1^T E[\mathbf{x} \mathbf{x}^T] \mathbf{w}_1 = \mathbf{w}_1^T C_x \mathbf{w}_1$$

とし、独立成分分析では \mathbf{w}_1 のノルム $\|\mathbf{w}_1\|=1$ の制約の下で最大化する荷重ベクトル \mathbf{w}_1 を探すことになる。制約条件の下で、目的関数を最大化させるようなパラメータを推定するには、ラグラン

ジユの未定乗数法を用いて解くことができる。未定乗数を λ_1 とすると目的関数 $L_1^{pca}(\mathbf{w}, \lambda_1)$ は

$$L_1^{pca}(\mathbf{w}_1, \lambda_1) = -J_1^{pca}(\mathbf{w}_1) + \lambda_1(\|\mathbf{w}_1\|^2 - 1)$$

となる。 $J_1^{pca}(\mathbf{w}_1)$ に負号がついたのは最大値をもとめるためである。ラグランジュ関数 $L_1^{pca}(\mathbf{w}_1, \lambda_1)$ を w_1 で微分すると

$$\frac{\partial L_1^{pca}(\mathbf{w}_1, \lambda_1)}{\partial \mathbf{w}_1} = -2C_x \mathbf{w} + 2\lambda_1 \mathbf{w}_1 = 0$$

$$C_x \mathbf{w}_1 = \lambda_1 \mathbf{w}_1$$

となる。PCA 問題の解は行列 C_x の単位長の固有ベクトル $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ を用いて表すことができ、対応する固有値 d_1, d_2, \dots, d_n は $d_1 \geq d_2 \geq \dots \geq d_n$ を満たすようにする。 $L_1^{pca}(\mathbf{w}_1, \lambda_1)$ を最大化する解は

$$\mathbf{w}_1 = \mathbf{e}_1$$

で与えられ、 \mathbf{x} の第一主成分は $y = \mathbf{e}_1^t \mathbf{x}$ である。

3.3 第 m 主成分の導出

\mathbf{x} の第二主成分 $z_2 = \mathbf{w}_2^T \mathbf{x}$ は、 $\|\mathbf{w}_2\| = 1$ の制約に加えて、 $\mathbf{w}_1^t \mathbf{w}_2 = 0$ という制約の下で、 z_2 の分散 C_{z2} を最大化する。未定乗数を λ_1, λ_2 とすると目的関数は

$$L = -\mathbf{w}_2^t C_x \mathbf{w}_2 + \lambda_1(\mathbf{w}_1^t \mathbf{w}_2) + \lambda_2(\mathbf{w}_2^t \mathbf{w}_2 - 1) \quad (1)$$

となる。これを \mathbf{w}_2, λ_1 でそれぞれ偏微分すると

$$\frac{\partial L}{\partial \mathbf{w}_2} = 2C_x \mathbf{w}_2 + \lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 = 0 \quad (2)$$

$$\frac{\partial L}{\partial \lambda_1} = \mathbf{w}_1^T \mathbf{w}_2 = 0 \quad (3)$$

(3) より、(2) の両辺の左側から \mathbf{w}_2^t をかけると

$$\mathbf{w}_2^t C_x \mathbf{w}_2 = \lambda_2$$

ゆえに \mathbf{w}_2 は (1) を最大化する固有ベクトル \mathbf{e}_2 が与えられ、第二主成分 $z_2 = \mathbf{w}_2^t \mathbf{x}$ である。次にこれを m 番目の主成分 $z_m = \mathbf{w}_m^T \mathbf{x}$ に拡張でき、 $\|\mathbf{w}_m\| = 1$ と $\sum_{i=1}^{m-1} \mathbf{w}_m^t \mathbf{w}_i = 0$ の制約条件の下で z_m の分散を最大化する目的関数は

$$L = -\mathbf{w}_m^T C_x \mathbf{w}_m + \sum_{i=1}^{m-1} \lambda_i \mathbf{w}_m^T \mathbf{w}_i + \lambda_m (\mathbf{w}_m^t \mathbf{w}_m - 1)$$

であり、これを \mathbf{w}_m と $\lambda_1 \sim \lambda_m$ で偏微分し、連立方程式を解けば

$$\mathbf{w}_m^T C_x \mathbf{w}_m = \lambda_m$$

より、第 m 主成分は $z_m = \mathbf{e}_m^T \mathbf{x}$ で与えられる。

3.4 1 固有ベクトルアルゴリズム

目的関数 $J_1^{pca} = C_z = \mathbf{w}_1^T C_x \mathbf{w}_1 = \mathbf{w}_1^T E[\mathbf{x}\mathbf{x}^T] \mathbf{w}_1$ より、最急勾配法を用いた主成分分析のアルゴリズムを以下に記す。

1. \mathbf{w}_1 が $\|\mathbf{w}_1\| = 1$ となるように乱数で決定する。
2. $\mathbf{w}_1 \leftarrow \mathbf{w}_1 - \alpha E[\mathbf{x}\mathbf{x}^T] \mathbf{w}_1$ (α : 学習係数)
3. $\mathbf{w}_1 \leftarrow \mathbf{w}_1 / \|\mathbf{w}_1\|$

\mathbf{x} の確率密度関数は分からないので

$$E[\mathbf{x}\mathbf{x}^T] \simeq \sum_{i=1}^T \mathbf{x}(i) \mathbf{x}^T(i) / T$$

であるとする。

3.5 直交化

m 個の固有ベクトルのアルゴリズムの説明の前に、直行化について説明する。PCA の解は理論的には直交系であることがわかっているが、反復アルゴリズムはいつでも自動的に直交性を保証するものでないので、各反復後にベクトルを直交化する必要がある。ここでは、古典的な直交化法であるグラム＝シュミットの直交化法について説明する。今、 n 次元のベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が与えられているとし

$$\mathbf{e}_1 = \mathbf{a}_1 / \|\mathbf{a}_1\|$$

をもととして新しく \mathbf{a}_1 と直交するベクトル \mathbf{b}_2 を

$$\mathbf{b}_2 = \mathbf{a}_2 + \alpha \mathbf{e}_1$$

と設定し、 \mathbf{a}_1 と \mathbf{e}_1 は直交しているから

$$\alpha = -\mathbf{e}_1^T \mathbf{a}_2$$

より

$$\mathbf{b}_2 = \mathbf{a}_2 - (\mathbf{e}_1^T \mathbf{a}_2) \mathbf{e}_1$$

$$\mathbf{e}_2 = \mathbf{b}_2 - (\mathbf{e}_1^T \mathbf{b}_2) \mathbf{e}_1 / \|\mathbf{b}_2 - (\mathbf{e}_1^T \mathbf{b}_2) \mathbf{e}_1\|$$

\mathbf{e}_1 と直交する単位行列 \mathbf{e}_2 が求められる。これを \mathbf{a}_m について拡張する。

$$\mathbf{b}_m = \mathbf{a}_m + \sum_{i=1}^{m-1} \alpha_i \mathbf{e}_i$$

\mathbf{b}_m は $\mathbf{e}_i (1, \dots, m-1)$ とそれぞれ直交することを利用して, $\alpha_i = -\mathbf{e}_i^T \mathbf{a}_m$ より m 番目の単位ベクトルは

$$\begin{aligned} \mathbf{b}_m &= \mathbf{a}_m - \sum_{i=1}^{m-1} (\mathbf{e}_i^T \mathbf{a}_m) \mathbf{e}_i \\ \mathbf{e}_m &= \mathbf{b}_m / \|\mathbf{b}_m\| = \left(\mathbf{a}_m - \sum_{i=1}^{m-1} (\mathbf{e}_i^T \mathbf{a}_m) \mathbf{e}_i \right) / \left\| \mathbf{a}_m - \sum_{i=1}^{m-1} (\mathbf{e}_i^T \mathbf{a}_m) \mathbf{e}_i \right\| \end{aligned}$$

と求めることができる。

3.6 m 個の固有ベクトルアルゴリズム

前節で説明したグラム＝シュミットの直交化を用いて m 個の固有ベクトルを求めるアルゴリズムを記す。

1. 求めたい固有ベクトルの数 m を決める。カウンタ p を $p \leftarrow 1$ とする。
2. \mathbf{w}_p を $\|\mathbf{w}_p\| = 1$ となるように乱数で決定してやる。
3. $\mathbf{w}_p \leftarrow \mathbf{w}_p - \alpha(E[\mathbf{x}\mathbf{x}^T])\mathbf{w}_p$ (α :学習係数)
4. $\mathbf{w}_p \leftarrow \mathbf{w}_p - \sum_{i=1}^{p-1} (\mathbf{w}_i^T \mathbf{w}_p) \mathbf{w}_i$
5. $\mathbf{w}_p \leftarrow \mathbf{w}_p / \|\mathbf{w}_p\|$.
6. もし \mathbf{w}_p が収束していなければ 3 に戻る。
7. $p \leftarrow p + 1$ とする。もし、 $p \leq m$ ならば 2 に戻る。

3.7 白色化

次に固有ベクトルの導出および PCA アルゴリズムが分かったところで、この節では ICA の前処理として有効な白色化について説明する。白色化とは、確率ベクトル $\mathbf{x} = (x_1, x_2, \dots, x_n)$ に対して行列 V により、

$$\mathbf{z} = V\mathbf{x}$$

確率ベクトル $\mathbf{z} = (z_1, z_2, \dots, z_n)$ が $E[\mathbf{z}] = 0$, $E[C_z] = I$ になる時である。またこの時の行列 V を白色化行列と呼ぶ。白色化行列 V は以下で与えられる。

$$V = D^{-1/2} E^T$$

$D^{-1/2} = \text{diag}(d_1^{-1/2}, \dots, d_n^{-1/2})$ を C_x の固有値をそれぞれ $-1/2$ 乗したもからなる対角行列、 E は C_x の固有ベクトルを並べたものであり、 $E^T E = I$, $E E^T = I$ を満足するものとする。これ

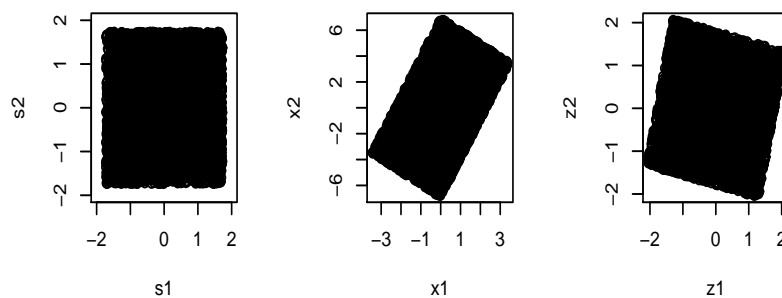
により、 $E^T C_x E = D$ を使って、 z の平均、共分散行列を求めれば

$$\mu_z = E[z] = VE[x] = 0$$

$$C_z = E[zz^T] = D^{-1/2} E^T E[xx^T] E D^{-1/2} = D^{-1/2} E^T C_x E D^{-1/2} = I$$

であることが確かめられる。

実際に、統計的に独立な要素をもつ二次元の一様分布（左の図）にたいして混合行列 A をかけたもの（中心の図）に対して白色化したもの（右の図）を以下に示す



観測値 x に対する白色化のプログラムは簡略化するため、固有値および固有ベクトルを R で用意されている関数 (`eigen $ value`, `eigen $ vector`) を使用してプログラムを書いた。また観測値 x を中心化のプログラムについても記述する。

白色化

```
whitening <- function(x){  #x: 観測値

  C <- var(x);  #C:x の共分散行列 Cx
  E <- eigen(C)$vectors;  #E: Cx の固有ベクトル行列
  D <- matrix(0,length(x[1,]),length(x[1,]));  #D: 固有値の-1/2 乗を格納する行列

  for(i in 1:length(x[1,])){

    D[i,i] <- (sqrt(eigen(C)$values[i]))^-1;

  }

  V <- D %*% t(E);  #V: 白色化行列
  z <- matrix(0,length(x[,1]),length(x[1,]));  #z: 無相関なベクトル z を格納
```

する行列

```
for(i in 1:length(x[,1])){  
  
  z[i,] <- V %% matrix(x[i,],length(x[1,]),1);  
  
}  
  
return(z);  
  
}
```

中心化

```
centration <- function(x){  
  
  x0 <- matrix(0,length(x[,1]),length(x[1,]));  #x0: 中心化された観測値 x を格納  
                                                    する行列  
  
  for(i in 1:length(x[1,])){  
  
    x0[,i] <- x[,i] - mean(x[,i]);  
  
  }  
  
  return(x0);  
  
}
```

4 情報理論

本章では、情報理論の基本概念および ICA の方法論で必要となるネゲエントロピーの近似について述べる。

4.1 エントロピー

エントロピーは情報理論の基本的概念である。離散値確率変数 x に対してエントロピーは、

$$H(x) = - \sum_{i=1}^n p(x = a_i) \log p(x = a_i)$$

で定義される。また、確率変数のエントロピーは、その確率変数が観測によって得られる情報の大きさと解釈され、確率変数が不規則であるほどエントロピーは大きく、不規則性がないほどエントロピーは小さい。離散エントロピーで最大を与えるのは、確率 $p(x = a_i) = 1/N$ である。関係式 $\log \alpha \leq \alpha - 1$ を利用して、任意の確率 $p(x = b_i)$ 、 $\alpha = p(x = a_i)/p(x = b_i)$ とし、先ほどの関係式に代入して

$$\log p(x = a_i) - \log p(x = b_i) \leq p(x = a_i)/p(x = b_i) - 1$$

両辺に $p(x = b_i)$ をかけて和をとれば

$$- \sum_{i=1}^n p(x = b_i) \log p(x = b_i) \leq - \sum_{i=1}^n p(x = b_i) \log p(x = a_i) = \log N$$

であり、左辺は $p(x = a_i) = 1/N$ の離散エントロピー $-\sum_{i=1}^n p(x = a_i) \log p(x = a_i) = -\sum_{i=1}^n 1/N \log 1/N = \log N$ と一致する。

4.2 微分エントロピー

前の節では離散値の確率変数に対するエントロピーについて定義したが、これは連続値確率変数やベクトルに拡張することができる。これを微分エントロピーと呼び、

$$H(x) = - \int p_x(\xi) \log p_x(\xi) d\xi$$

で定義される。微分エントロピーは、エントロピーと同様に確率変数が不規則であるほど微分エントロピーは大きく、不規則性がないほど小さいので、不規則性の尺度として解釈できる。また、 $E[x] = 0, \sigma_x^2 = 1$ の分布のなかで最大のエントロピーを確率分布密度関数は、標準ガウス関数 $\phi(x) = \exp(-x^2/2)/\sqrt{2\pi}$ である。離散エントロピーと同様に、関係式 $\log \alpha \leq \alpha - 1$ に任意の確率密度関数 $p(x)$ を導入して、 $\alpha = \phi(x)/p(x)$ を先ほどの関係に代入して

$$\log \phi(x) - \log p(x) \leq \phi(x)/p(x) - 1$$

この両辺に $p(x)$ をかけて x で積分すれば

$$\begin{aligned} - \int_{-\infty}^{+\infty} p(x) \log p(x) dx &\leq - \int_{-\infty}^{+\infty} p(x) \log \phi(x) \\ &= - \int_{-\infty}^{+\infty} p(x) \log(1/\sqrt{2\pi} + (-x^2/2)) dx \\ &= (\log 2\pi + 1)/2 \end{aligned}$$

であり、最後の式は標準ガウス密度関数 $\phi(x)$ のエントロピー

$$\begin{aligned} - \int_{-\infty}^{+\infty} \phi(x) \log \phi(x) &= - \int_{-\infty}^{+\infty} \phi(x) \{\log(1/\sqrt{2\pi} + (-x^2/2))\} dx \\ &= (\log 2\pi + 1)/2 \end{aligned}$$

と一致する。

4.3 ネゲエントロピー

ネゲエントロピーとは、分布の非ガウス性の基準として以下のように定義される。

$$J(x) = H(v) - H(x)$$

v は x と同じ共分散行列をもつガウス確率ベクトルである。 J が大きい時は x の分布のエントロピーは小さく、 J が小さい時は x の分布のエントロピーは大きい。

4.4 最大エントロピー法

有限個の測定値から確率密度関数 $p(x)$ を正確に求めることはできないが、束縛条件が期待値をとるような形である

$$\int p(x) F^i(x) dx = c_i \quad i = 1, \dots, n$$

が与えられている時、上の制約条件を満たす確率密度のなかでエントロピーを最大にする関数 $p(x)$ は見つけることを最大エントロピー法呼ぶ。また、エントロピー最大の密度とは、測定値と適合し、かつデータに関する仮定が最も少ないような密度であると解釈される。

汎関数

$$I = - \int p(x) \log p(x) dx + \sum_{i=1}^{n-1} \lambda_i \left(\int F^i(x) p(x) dx - c_i \right) + \lambda_n \left(\int p(x) dx - 1 \right)$$

は上のようになり、これを $p(x)$ で微分すれば

$$\frac{\sigma I}{\sigma p(x)} = \int (-\log p(x) + 1 + \sum_{i=1}^{n-1} \lambda_i F^i(x) + \lambda_n) dx = 0$$

より最大のエントロピーをもつ確率密度関数

$$p(x) = A \exp\left(\sum_{i=1}^n \lambda_i F^i(x)\right)$$

を求めることが出来る ($A = \exp(1 + \lambda_{n+1})$) 先ほどの制約条件に

$$F^{n+1}(x) = x \quad c_{n+1} = 0$$

$$F^{n+2}(x) = x^2 \quad c_{n+2} = 1$$

を追加すれば確率密度関数は

$$\begin{aligned} p(x) &= A \exp(\lambda_{n+1}x + \lambda_{n+2}x^2 + \sum_{i=1}^n \lambda_i F^i(x)) \\ &= A \exp(-x^2/2 + \lambda_{n+1}x + (\lambda_{n+2} + 1/2)x^2 + \sum_{i=1}^n \lambda_i F^i(x)) \\ &= \tilde{A} \phi(x) \exp(\lambda_{n+1}x + (\lambda_{n+2} + 1/2)x^2 + \sum_{i=1}^n \lambda_i F^i(x)) \end{aligned}$$

と書くことができ ($\tilde{A} = \sqrt{2\pi}A$, $\phi(x)$ は標準ガウス密度である)、 $\exp(x) \simeq 1 + x$ の近似を使えば

$$p(x) = \tilde{A} \phi(x) (1 + \lambda_{n+1}x + (\lambda_{n+2} + 1/2)x^2 + \sum_{i=1}^n \lambda_i F^i(x))$$

さらに計算を単純化させるために、 $F_i(i, \dots, n)$ は標準ガウス分布 $\phi(x)$ に対して

$$\begin{aligned} \int_{-\infty}^{+\infty} \phi(x) F^i(x) F^j(x) dx &= \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \\ \int_{-\infty}^{+\infty} \phi F^i(x) x^k dx &= 0 \quad k = 0, 1, 2 \end{aligned} \tag{4}$$

のような直行性の制約条件を加える。これにより

$$\begin{aligned} \int_{-\infty}^{+\infty} p(x) dx &= \tilde{A} (1 + (\lambda_{n+2} + 1/2)) = 1 \\ \int_{-\infty}^{+\infty} x p(x) dx &= \tilde{A} \lambda_{n+1} = 0 \\ \int_{-\infty}^{+\infty} x^2 p(x) dx &= \tilde{A} (1 + 3(\lambda_{n+2} + 1/2)) = 1 \\ \int_{-\infty}^{+\infty} F^i(x) p(x) dx &= \tilde{A} \lambda_i = c_i \quad i = 1, \dots, n \end{aligned}$$

これを解けば、 $\tilde{A} = 1, \lambda_{n+1} = 0, \lambda_{n+2} = 1/2, \lambda_i = c_i$ を得る。ゆえに近似的最大エントロピー密度

$$p(x) \simeq \phi(x)(1 + \sum_{i=1}^n c_i F^i(x))$$

を得ることができる。

4.5 ネゲエントロピーの近似

前節で導出して確率密度関数 $p(x)$ を用いて、エントロピーを計算する。

$$-\int_{-\infty}^{+\infty} p(x) \log p(x) dx = \int \phi(x) (1 + \sum_{i=1}^n F^i(x)) (\log(1 + \sum_{i=1}^n c_i F^i(x)) + \log \phi(x)) dx$$

ここで、近似式 $(1+x)\log(1+x) \simeq x + x^2/2$ (高次の項を無視) を用いれば

$$\begin{aligned} -\int_{-\infty}^{+\infty} p(x) \log p(x) dx &= -\int_{-\infty}^{+\infty} \phi(x) \log \phi(x) dx - \int_{-\infty}^{+\infty} \phi(x) \sum c_i F^i(x) \log \phi(x) \\ &\quad - \int_{-\infty}^{+\infty} \phi(x) (\sum c_i F^i(x) + 1/2 (\sum c_i F^i(x))^2) dx \\ &= H(v) - \sum c_i^2 / 2 \end{aligned}$$

が得られる。よって、ネゲエントロピーの近似は

$$\begin{aligned} J(x) &\simeq H(v) - (H(v) - 1/2 \sum c_i^2) \\ &= 1/2 \sum c_i^2 \end{aligned}$$

である。次に直交性の制約を満足する $F^i(x)$ は、実質的に任意の線形独立な関数の集合 $G^i (i = 1, \dots, m)$ に単項式 $x^k (k = 0, 1, 2)$ を付け加えてグラム＝シュミットの直行法を施すと得られる。例えば G^1 を奇関数、 G^2 を偶関数とすれば F^1, F^2 は

$$F^1(x) = (G^1(x) + \beta_1 x) / \sigma_1$$

$$F^2(x) = (G^2(x) + \alpha_2 x^2 + \gamma_2) / \sigma_2$$

のようになり、(4) の意味で直行化している。また

$$p(x) - \phi(x) = \phi(x) (\sum_{i=1}^n c_i F^i(x))$$

の両辺に $F^i(x)$ をかけて積分すると

$$E[G^i(x)] - E[G^i(v)] = c_i$$

が得られる。これを近似的ネゲエントロピーに代入すれば、

$$J(x) \simeq k_1(E[G^1(x)] - E[G^1(v)])^2 + k_2(E[G^2(x)] - E[G^2(v)])^2$$

になり ($k_1 = 1/2\sigma_1^2, k_2 = 1/2\sigma_2^2$)、関数 G^1 は奇関数であることを用いれば

$$J(x) \simeq k_1(E[G^1(x)])^2 + k_2(E[G^2(x)] - E[G^2(v)])^2$$

が得られる。

5 独立成分分析

本章では、独立成分分析の前処理として有効である白色化と非ガウス性の尺度として導入したネゲエントロピーを用いて独立成分分析の解 W を求める。

5.1 独立成分分析の原理

基本的な ICA モデル

$$\mathbf{x} = A\mathbf{s}$$

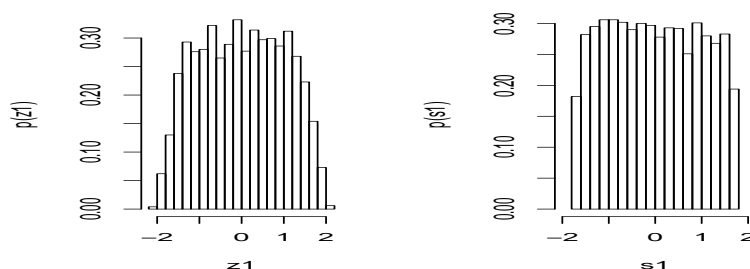
から信号源 \mathbf{s} を推定するには、 \mathbf{s} に対して独立性を仮定するだけでよい。独立性とは、結合密度関数

$$p(\mathbf{s}) = p(s_1)p(s_2)\cdots p(s_n)$$

を上のように、それぞれの周辺密度関数の積であらわせることである。独立成分分析では、この独立性を無相関性と非ガウス性を用いることにより再現できる。無相関性とは共分散が 0（平均が 0 の時）であることで、非ガウス性とは分布がガウス性とは異なることである。無相関性を使う理由としては、統計的に独立な変数は相関がないことは明らかで、独立成分分析を行う前処理として、成分を無相関にする白色化は非常に有効であるからである。実際に、統計的に独立な要素をもつ二次元の一樣分布 \mathbf{s} に未知の行列をかけた観測値 \mathbf{x} に対して白色化行列 V をかけると

$$\mathbf{z} = V\mathbf{x}$$

により、無相関なベクトル \mathbf{z} を得ることができる。 \mathbf{z} の一成分 z_1 と \mathbf{s} の一成分 s_1 を比較した場合



z_1 の方が中心極限定理よりガウス性を帯びている。このことから、我々は \mathbf{z} に対して推定値 $\mathbf{y} = W\mathbf{z}$ が非ガウス性を最大にもつような W を \mathbf{z} にかけてやれば良いことが分かる。分布の非ガウス性の尺度としてネゲエントロピー J を導入した。このネゲエントロピー J を最大にする W こそが求める分布の非ガウス性を最大にする。

5.2 ネゲエントロピーを尺度とした ICA

前節の情報理論でネゲエントロピーの近似を導出した。

$$J(y) \simeq k_1 E[G^1(y)]^2 + k_2 E[(G^2(y) - G^2(v))^2]$$

G^1 は奇関数で G^2 は偶関数である。ここで、 y の従う確率密度が偶関数に近いと仮定すれば先ほどのネゲエントロピーはさらに近似できて

$$J(y) \propto E[(G^2(y) - G^2(v))^2]$$

となる。一般的に以下 G が大変有効であると分かっている。

$$G_1(y) = \frac{1}{\alpha} \log \cosh \alpha y$$

$$G_2(y) = -\exp(-y^2/2)$$

ここで $1 \leq \alpha \leq 2$ は何か適切な定数で、しばしば 1 が選ばれる。

5.3 確率的最急降下法を用いた ICA アルゴリズム

目的関数 $J(\mathbf{w})$ は

$$J(\mathbf{w}) = (E[G(\mathbf{w}^T \mathbf{z})] - E[G(v)])^2$$

として与えられる。 $J(\mathbf{w})$ に対して \mathbf{w} の勾配をとれば

$$\begin{aligned} \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} &= 2(E[G(\mathbf{w}^T \mathbf{z})] - E[G(v)])(\mathbf{z} E[g(\mathbf{w}^T \mathbf{z})]) \\ &= r E[g(\mathbf{w}^T \mathbf{z})] \end{aligned}$$

が得られる ($r = 2E[G(\mathbf{w}^T \mathbf{z}) - E[G(v)]]$, $g: G$ の導関数)。独立成分分析では分散を、 $E[yy^T] = w^T E[zz^T] w = 1$ とするので \mathbf{w} のノルムが 1 という考慮を入れる。非ガウス性が最大の一つの方向を求めるための確率的勾配アルゴリズムの手順は

1. データの平均を 0 にする中心化を行う。
2. データの白色化して、 $y = \mathbf{w}^T \mathbf{z}$ とする。
3. \mathbf{w} の初期値は $\|\mathbf{w}\| = 1$ になるように乱数で決定する。
4. $\Delta w \propto r E[g(\mathbf{w}^T \mathbf{z})]$ とする。ここで、 $G(\mathbf{w}^T \mathbf{z}) = \log \cosh(\mathbf{w}^T \mathbf{z})$
5. $\mathbf{w} \leftarrow \mathbf{w} / \|\mathbf{w}\|$
6. 収束していなければ、4 に戻る。

5.4 fastICA アルゴリズム

最急勾配法は収束が遅く、よい学習係数を選ばなければならない。学習係数の選択が悪いと、実際に収束しなくなる場合がある。そのため収束を速く、学習係数を選ばない方法として不動点アルゴリズムがある。最急勾配法の収束点において、その勾配は \mathbf{w} の方向を向いていることと \mathbf{w} の正規化の際に係数 r が消えることを利用すると更新則は

$$\mathbf{w} \leftarrow E[\mathbf{z}g(\mathbf{w}^T \mathbf{z})]$$

$$\mathbf{w} \leftarrow \mathbf{w}/\|\mathbf{w}\|$$

となる。

5.4.1 不動点反復法

実は、不動点アルゴリズムをネゲエントロピーに用いるとあまり収束性はあまりよくないことが知られているが、近似的なニュートン法を用いることによりそれを回避することができる。ここで $E[G(\mathbf{w}^T \mathbf{z})]$ の最適点は $\|\mathbf{w}\| = 1$ という制約条件の下で、ラグランジュ関数の勾配が 0 になるところである。

$$E[\mathbf{z}g(\mathbf{w}^T \mathbf{z})] + \beta \mathbf{w} = 0$$

左の関数を F とおき、ラグランジュ関数の二次勾配について求めれば

$$\frac{\partial F}{\partial \mathbf{w}} = E[\mathbf{z}\mathbf{z}^T g'(\mathbf{w}^T \mathbf{z})] + \beta I$$

データ \mathbf{z} は白色化されているので

$$E[\mathbf{z}\mathbf{z}^T g'(\mathbf{w}^T \mathbf{z})] \simeq E[\mathbf{z}\mathbf{z}^T]E[g'(\mathbf{w}^T \mathbf{z})] = E[g'(\mathbf{w}^T \mathbf{z})]I$$

という近似が成り立っても悪くはないと考えれば、以下の近似的なニュートン法が得られる。

$$\mathbf{w} \leftarrow \mathbf{w} - (E[\mathbf{z}g(\mathbf{w}^T \mathbf{z})] + \beta \mathbf{w}) / (E[g'(\mathbf{w}^T \mathbf{z})] + \beta)$$

この両辺に $\beta + E[g'(\mathbf{w}^T \mathbf{z})]$ をかけることにより

$$\mathbf{w} \leftarrow E[\mathbf{z}g(\mathbf{w}^T \mathbf{z})] - E[g'(\mathbf{w}^T \mathbf{z})]\mathbf{w}$$

が得られ、これが fastICA における不動点反復法である。

5.5 一成分における fastICA アルゴリズム

前節で説明した不動点反復法を用いた fastICA を以下に記述する。

1. データの平均を 0 にする中心化を行う。

2. データを白色化したものを \mathbf{z} とする。
3. 初期値 \mathbf{w} が $\|\mathbf{w}\| = 1$ になるように乱数を用いて決める。
4. $w \leftarrow E[\mathbf{z}g(\mathbf{w}^T\mathbf{z})] - E[g'(\mathbf{w}^T\mathbf{z})]\mathbf{w}$
5. $\mathbf{w} \leftarrow \mathbf{w}/\|\mathbf{w}\|$
6. 収束 ($\|w_{t-1}^T w_t\| \geq r, 0 \leq r \leq 1$) していなければ 4 に戻る。以下に一成分に対する fastICA アルゴリズム (fastICA1D) のプログラムを記述する。

fastICA1D <- function(z,n,r){ #z:白色後のデータ n:反復回数 r:収束半径

 w <- matrix(0,length(z[1,]),1); #アルゴリズム 3

 W <- c(0);

 for(i in 1:length(z[1,])){

 w[i,] <- rnorm(1);

 W <- W + w[i,]^2;

 }

 w <- w/sqrt(W);

 for(i in 1:n){ #アルゴリズム 4

 A <- matrix(0,length(z[1,]),1);

 w0 <- w;

 for(t in 1:length(z[,1])){

 A <- A + (matrix(z[t,],length(z[1,]),1)*tanh(c(t(w)%*%matrix(z[t,],length(z[1,]),1)))
 -(1-tanh(c(t(w)%*%matrix(z[t,],length(z[1,]),1)))^2)*w);

 }

 w <- A/length(z[,1])

 W <- c(0)

 for(t in 1:length(z[1,])){

```

        W <- W + w[t,]^2;

    }

    w <- w/sqrt(W);    #アルゴリズム 5

    if(abs(c(t(w0)%*%w))>r) break;    #アルゴリズム 6

}

y <- matrix(0,length(z[,1]),1)    #独立成分の推定値の一成分

for(i in 1:length(z[,1])){

    y[i,] <- c(t(w)%*%matrix(z[i,],length(z[1,]),1));

}

return(y);

}

```

5.6 対称的直交化

第2章の主成分分析で複数の固有ベクトルを求める際に、逐次的直交化について説明したが、逐次的直交化には最初のほうのベクトルの推定誤差が、直交化により後続のベクトルの推定に累積するという欠点がある。そこで推定ベクトル \mathbf{w}_i 達をひとつずつ計算するのではなく、並列に推定する対称的直交化について述べる。まず、全ての \mathbf{w}_i について1単位アルゴリズムの反復アルゴリズムを1回適用して、 $W = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n)^t$ について以下の操作を行う。

$$W \leftarrow (WW^T)^{-1/2}W$$

$(WW^T)^{-1/2}$ は固有値分解

$$WW^T = E \text{diag}(d_1, d_2, \dots, d_n) E^T$$

を用いて

$$(WW^T)^{-1/2} = E \text{diag}(d_1^{-1/2}, d_2^{-1/2}, \dots, d_n)^{-1/2} E^T$$

によって与えられ、以下により W の要素 $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ は直行していることが分かる。

$$\begin{aligned} W_t W_t^T &= (W_{t-1} W_{t-1}^T)^{-1/2} W_{t-1} W_{t-1}^T ((W_{t-1} W_{t-1}^T)^{-1/2})^T \\ &= E \text{diag}(d_1, d_2, \dots, d_n)^{-1/2} E^T W_{t-1} W_{t-1}^T E \text{diag}(d_1, d_2, \dots, d_n)^{-1/2} E^T \\ &= I \end{aligned}$$

この操作を \mathbf{w}_i が収束するまで繰り返す。

5.7 複数成分における fastICA アルゴリズム

以下に、複数の独立成分を推定する fastICA アルゴリズムのソースコードを記す。

1. 観測信号 \mathbf{x} を中心化 $E[\mathbf{x}] = 0$ する。
2. 観測信号 \mathbf{x} を白色化したものを \mathbf{z} とする。
3. $\mathbf{w}_i (i = 1, \dots, n)$ の初期値それぞれが $\|\mathbf{w}_i\| = 1$ となるように乱数で決める。行列 W を下の 5 により直交化する。
4. 全ての $i = 1, \dots, n$ について
 $\mathbf{w}_i \leftarrow E[\mathbf{w}_i g(\mathbf{w}_i^T \mathbf{z})] - E[g'(\mathbf{w}_i^T \mathbf{z})] \mathbf{w}_i$ とする。ここで、 $g(\mathbf{w}_i^T \mathbf{z}) = \tanh(\mathbf{w}_i^T \mathbf{z})$ を用いる。
5. $W = (\mathbf{w}_1, \dots, \mathbf{w}_n)^T$ の対称的直交化
 $W \leftarrow (WW^T)^{-1/2} W$ を行う。
6. もし収束 ($W_{t-1}^T W_t \leq R, 0 \leq R \leq I$) していなければ 4 に戻る

以下にプログラム (fastICA) を記述する。

```
fastICA <- function(z,n,R){    #z:白色化後のデータ n:反復回数 R:収束半径
```

```
    W0 <- matrix(0,length(z[1,]),length(z[1,]));    #アルゴリズム 3
```

```
    W <- matrix(0,length(z[1,]),length(z[1,]));
```

```
    for(i in 1:length(z[1,])){
```

```
        A <- c(0);
```

```
        for(t in 1:length(z[1,])){
```

```
            W0[i,t] <- rnorm(1);
```

```
            A <- A + W0[i,t]^2;
```

```

}

W0[i,] <- W0[i,] / sqrt(A);

}

E <- eigen(W0 %*% t(W0))$vectors;   #アルゴリズム 5
D <- matrix(0,length(z[1,]),length(z[1,]));

for(i in 1:length(z[1,])){

  D[i,i] <- (sqrt(eigen(W0 %*% t(W0))$values[i]))^-1;

}

W0 <- E %*% D %*% t(E) %*% W0;

for(s in 1:n){   #アルゴリズム 4

  for(i in 1:length(z[1,])){

    B <- matrix(0,1,length(z[1,]));

    for(t in 1:length(z[,1])){

      B <- B+matrix(z[t,],1,length(z[1,]))*tanh(c(matrix(W0[i,],1,length(z[1,]))
        %*%matrix(z[t,],length(z[1,]),1))-(1-(tanh(c(matrix(W0[i,],1,length(z[1,]))
        %*%matrix(z[t,],length(z[1,]),1)))^2))*matrix(W0[i,],1,length(z[1,]));

    }

    W[i,] <- B/length(z[,1]);
  }

  E <- eigen(W %*% t(W))$vectors;   #アルゴリズム 5
  D <- matrix(0,length(z[1,]),length(z[1,]));

```

```

for(i in 1:length(z[1,])){

  D[i,i] <- (sqrt(eigen(W %*% t(W))$values[i]))^-1;

}

W <- E %*% D %*% t(E) %*% W;

COS <- t(W)%*%W0; #アルゴリズム 6
W0 <- W;

k <- c(0);

for(o in 1:length(W[,1])){

  if(R<abs(COS[o,o])){
    k <- k+1 ;
  }
}

if(k==length(W[,1])){
  break;
}
}

y <- matrix(0,length(z[,1]),length(z[1,])); #複数の独立成分の推定値

for(i in 1:length(z[1,])){

  for(t in 1:length(z[,1])){

    y[t,i] <- matrix(W0[i,],1,length(z[1,]))%*%matrix(z[t,],length(z[1,]),1);
  }
}

return(y);
}

```


6 ICA の応用

ICA は、画像処理、通信技術、脳科学、金融など様々な分野で用いられている。ここでは ICA の応用として、実際に脳波計（EEG）から得られるデータを用いて ICA の処理を行った。

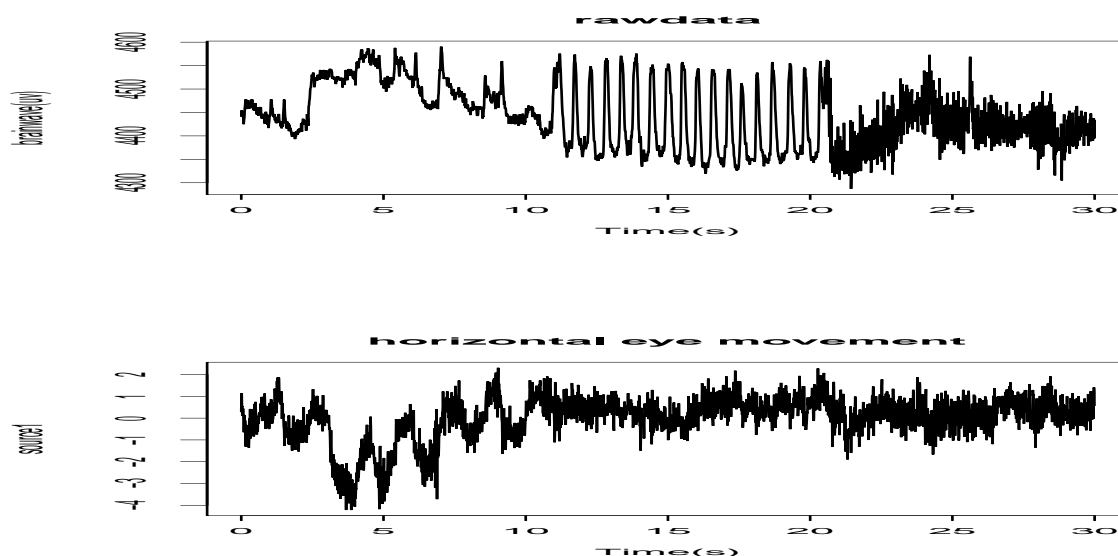
6.1 ICA による脳波データのノイズ除去

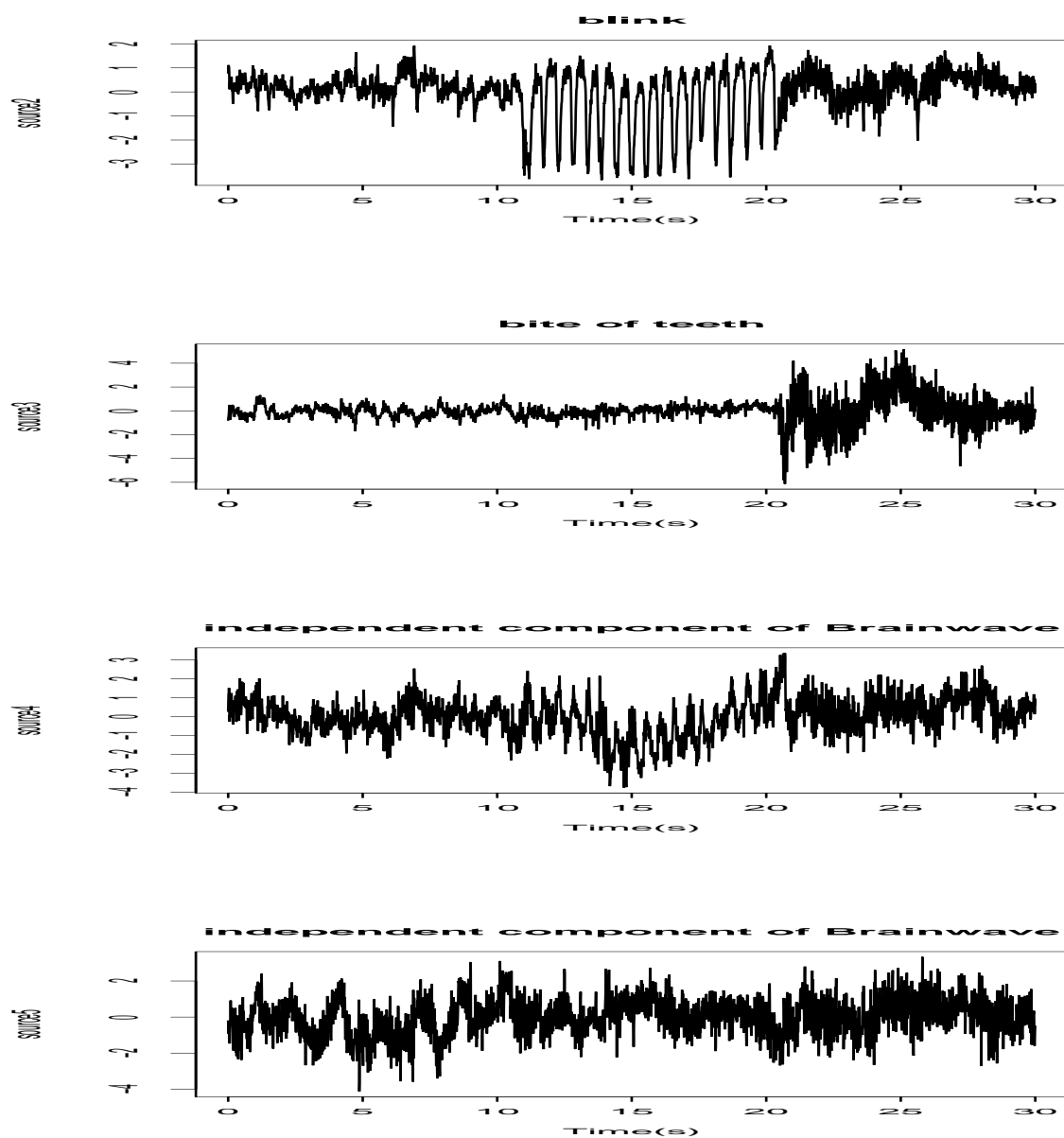
EEG から得られるデータは、脳波や筋電位などの混合の電位である。混合のデータから脳波のみを取り出したい場合、脳波や筋電位をそれぞれ独立であると仮定すれば、ICA の仮定にあてはまりそうである。基本的な ICA モデルは

$$\mathbf{x} = \mathbf{A}\mathbf{s}$$

とし、 $\mathbf{x} = (x_1, \dots, x_n)$ を脳波データ（※今回、 n は EEG の電極数であり、実験では 14 個の電極を使用した）、 $\mathbf{s} = (s_1, \dots, s_n)$ を脳波の独立成分や筋電位などの独立成分と仮定する。

この仮定において、電極数とそれぞれの独立成分の数が同じであるはずがないという批判はあるだろうが、そこはひとまず伏せて基本的な ICA モデルによって、それぞれの独立成分が推定できるのかを検証した。本実験では、脳波計として 14 電極の EEG（EPOC 社 emotiv）を使用した。被験者は眼球の水平運動、まばたき、歯の噛み締めめの 3 つの運動をそれぞれ順番に 10 秒間、合計 30 秒を 1 セットとし、合わせて 10 回の実験を行った。下図は、処理前と処理後の EEG のデータをプロットしたものである。一番上の処理前のデータを表し、残りの 5 つは ICA 処理後の独立成分（最後の 2 つが脳波の独立成分であると考えられる）である。





EEG から得られた混合の電位より ICA を用いることによって、筋電位が背景脳波よりも大きい時ですら、それぞれ分離できそうである。ゆえに脳波だけを観測したい場合に ICA は非常に有効であると考えられる。

参考文献

- [1] アーポ・バビリネン, エルキ・オヤ, ユハ・カールネン, ” 詳解 独立成分分析 ” , 東京電気大学出版局, 2013
- [2] アタナシアス・パポリス, ” 確率論の応用 ” , 東海大学出版会, 1990
- [3] 金谷健一, ” これなら分かる最適化数学 ” , 共立出版, 2005

謝辞

まず本論文を書くにあたり、指導教官である守真太郎先生に深く感謝を致します。ただ脳波を解析したいという漠然とした目標を持っていた当時の私に対して、独立成分分析というテーマを与えて下さったり、進学先の大学院で研究する分野でも同テーマが使われているなど、先生には足を向けて寝られない程、感謝しています。そして、同研究室の先生や先輩、同期にも感謝いたします。とくに同期の永井君には、自分の勉強があるにも関わらず、私の趣味の散歩に嫌な顔を見せずに付き合ってくれてたことを本当に嬉しく思います。