

月4 2013 冬学期 [4830-1032]

第8回

音声音響信号処理

(独立成分分析によるブラインド信号分離)

亀岡弘和

東京大学大学院情報理工学系研究科

日本電信電話株式会社

NTTコミュニケーション科学基礎研究所

講義内容(キーワード)

- 信号処理、符号化、標準化の実用システム例の紹介
- 情報通信の基本(誤り検出、訂正符号、変調、IP)
- 符号化技術の基本(量子化、予測、変換、圧縮)
- 音声分析・合成・認識・強調、音楽信号処理
- 統計的信号処理の基礎(スペクトル、ガウス過程、最尤推定)
- ガウス性確率変数の基本性質
- 時間周波数分析(短時間フーリエ変換、ウェーブレット変換)
- ウィナーフィルタとカルマンフィルタ
- 音声生成過程のモデル(ソースフィルタ理論と藤崎モデル)
- 自己回帰モデルと線形予測分析
- 独立成分分析によるブラインド音源分離
- 非負値行列因子分解によるスペクトログラムの分解表現
- スペクトル間擬距離
- 最適化アルゴリズム(EMアルゴリズム、補助関数法)

講義スケジュール

10/ 7 守谷先生担当

10/15 (火) 守谷先生担当

10/21 守谷先生担当

10/28 (休講)

11/ 5 (火) 線形予測分析と自己回帰モデル

11/11 時間周波数解析

11/18 (休講)

11/25 非負値行列因子分解

12/ 2 統計的手法による音声強調

⇒ **12/ 9 独立成分分析によるブラインド信号分離**

12/16 吉井和佳 氏 (産業技術総合研究所)

1/15 (水)

1/20 戸田智基 准教授 (奈良先端科学技術大学院大学)

成績評価

■レポート課題

- 本講義に関連する論文を1つ選び、発表資料形式(パワーポイント等)にまとめて学期末に提出してください。提出先は最終講義にてお知らせします。
- 「どの程度本質を理解しているか」「要点が分かりやすく記述されているか」「なぜその論文を重要と考えたか」を評価の規準にして採点します。
- 毎回の講義後にその回の講義に関連する論文を1つ挙げる予定です。それらの中から選んでも良いですし、自分で自由に探してきててもOKです。

■講義の感想

- レポートとともに講義に対する感想文も一緒に提出して下さい。
- ※講義資料は講義用ホームページにアップしていく予定。

本日の話題

■ ブラインド音源分離

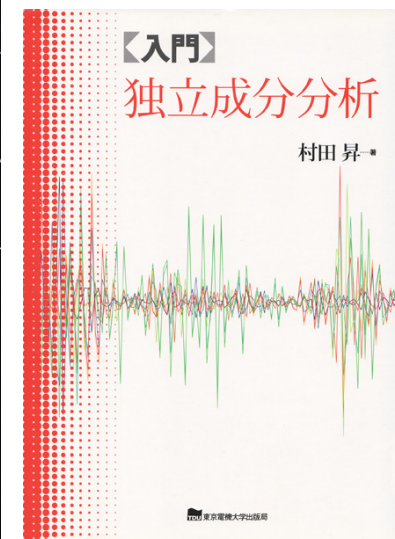
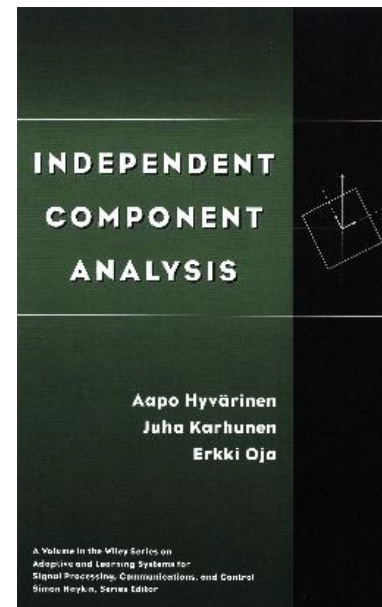
- 複数のマイクロホンで取得した観測信号から同時に鳴っている複数の音源信号を分離獲得する技術

■ 独立成分分析

- ブラインド音源分離を行うための統計的手法

■ 参考資料

- A. Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.
- 村田昇, 入門 独立成分分析, 東京電機大学出版局, 2004.
- 澤田宏, “【チュートリアル】独立成分分析入門 ～音の分離を題材として～,” 部分空間法研究会2010, 2010年7月.



目次

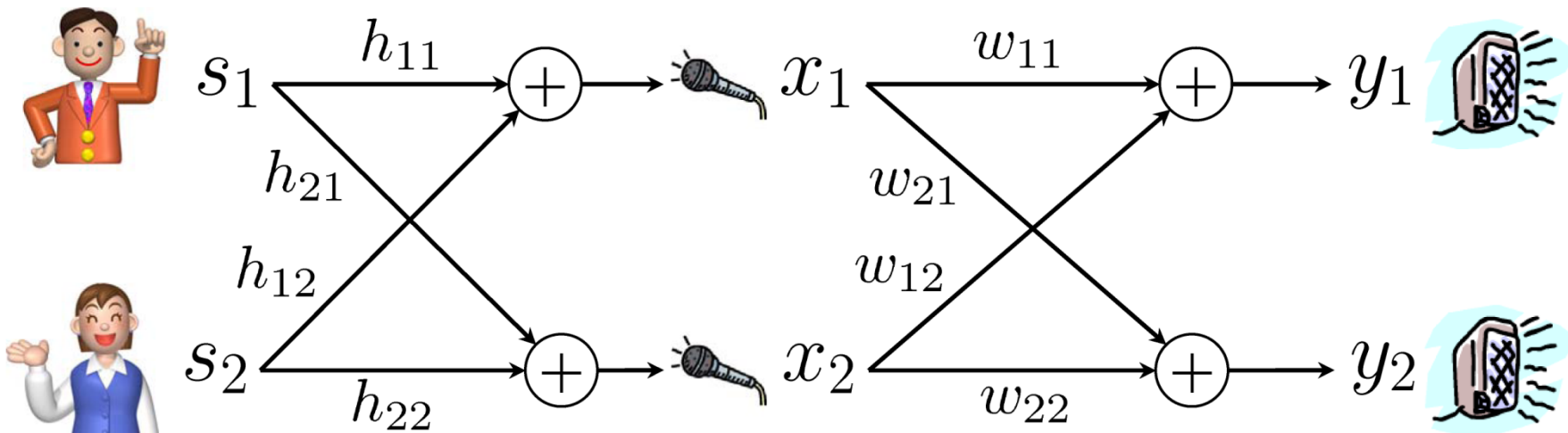
- 独立成分分析について
 - 定式化, 歴史, 応用
- 信号の統計的性質
 - 信号を混ぜるー 中心極限定理
- 独立成分分析のアルゴリズム
 - 白色化 + FastICA
 - 最尤推定法 by Natural Gradient

目次

- 独立成分分析について
 - 定式化, 歴史, 応用
- 信号の統計的性質
 - 信号を混ぜるー 中心極限定理
- 独立成分分析のアルゴリズム
 - 白色化 + FastICA
 - 最尤推定法 by Natural Gradient

ブラインド音源分離 (Blind Source Separation)

- 混ぜり合った信号 x_1, x_2 から元の信号を取り出す
- どのように混ぜたかに関する情報 H は利用できない



音の分離

■カクテルパーティー効果

- 人の聞き分け能力



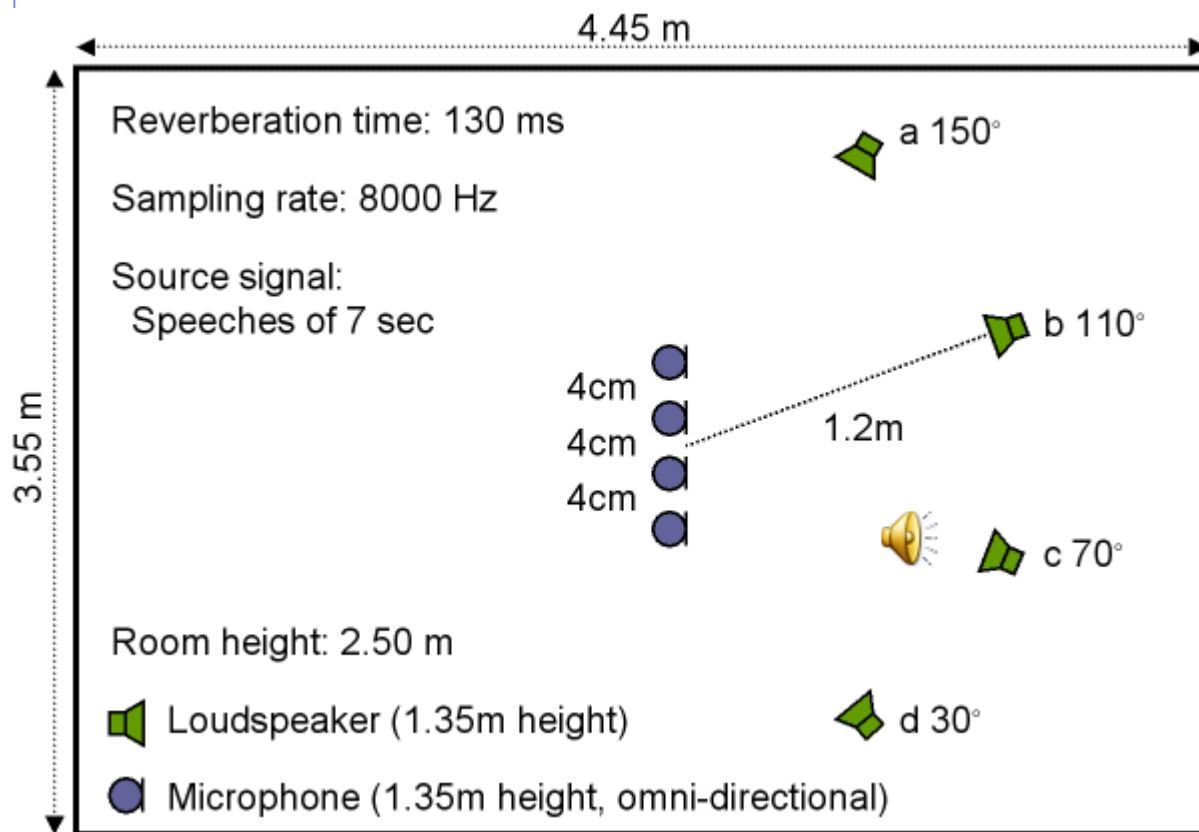
■音声認識

- マイクロホンと口の間の距離が大きくなるにつれて増大してくる妨害音を抑圧・除去

■音楽／楽器音分析

- 例)オーケストラの一つ一つの演奏をチェック

ブラインド音源分離の動作結果例



a	b	a	b	c	a	b	c	d
x1	x2	x1	x2	x3	x1	x2	x3	x4
y1	y2	y1	y2	y3	y1	y2	y3	y4

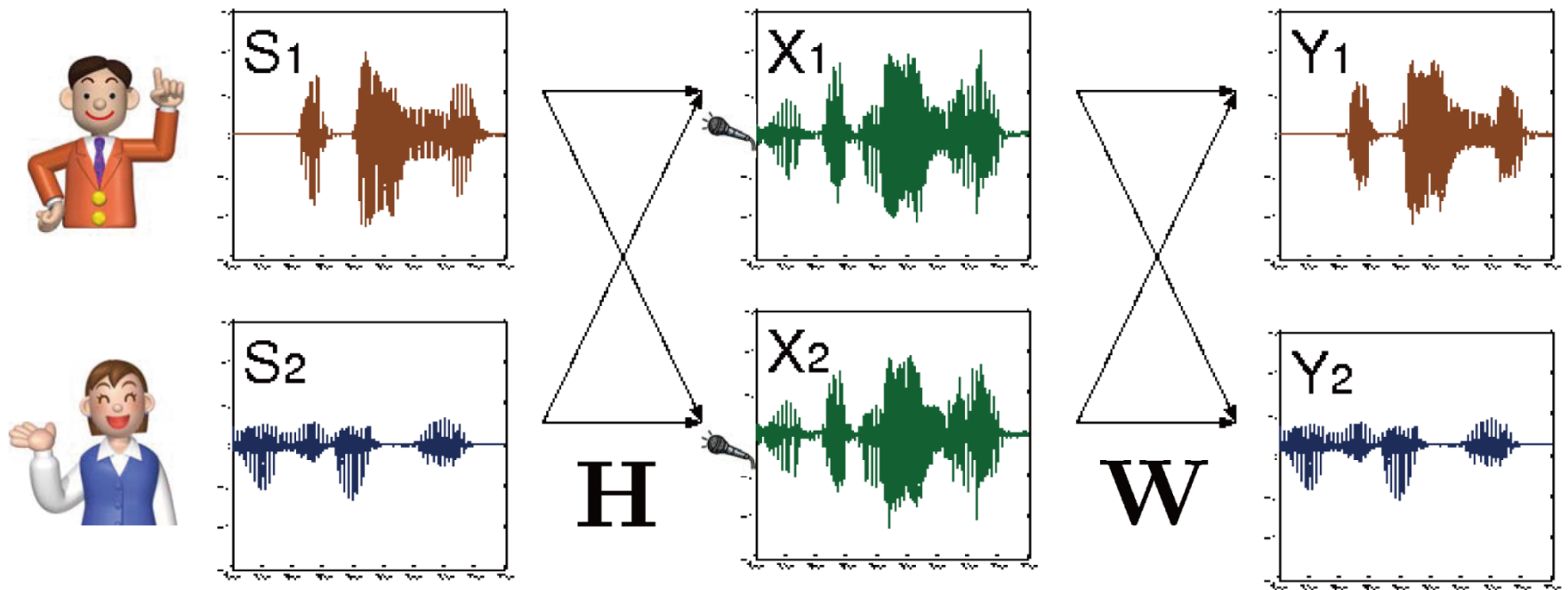
音の分離

- 音のみを手がかりとして混ざった音を分離
 - ⇔ 視覚(カメラ)に関わる情報は利用できない

- 何を手がかりにするか?
 - 音源の性質
 - ◆ 音声の場合: 声質, 話し方の特徴, 次に来る言葉を予測
 - ◆ 楽器の場合: 音の高さ/特徴, リズム, 繰り返しパターン
 - 空間情報
 - ◆ 音源の方向, 距離 ← 複数マイク(耳)

- 分離方法
 - 線形フィルタ: ビームフォーマ, 独立成分分析
 - 非線形処理: 時間周波数マスキング

独立成分分析 (Independent Component Analysis)



元の信号は独立

- 一方から他方が推測できない

依存関係がある

- 同じ成分が双方に入っている
- 一方から他方が推測できる

なるべく独立に

- 具体的にはどうやる？

独立成分分析によるブラインド音源分離

- I 個の音源信号 \mathbf{s} が混合行列 \mathbf{H} により混ざり合い、 J 個の観測信号 \mathbf{x} が T 個得られたとする。

$$\mathbf{x}(t) = \mathbf{H}\mathbf{s}(t) \quad t = 1, \dots, T$$
$$\mathbf{s} = \begin{bmatrix} s_1 \\ \vdots \\ s_I \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_I \end{bmatrix}$$

- 分離行列 \mathbf{W} により分離信号 \mathbf{y} を生成する。

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_I \end{bmatrix}$$

- \mathbf{W} の計算は観測信号 \mathbf{x} のみから行う。
 I 個の分離信号 y_1, \dots, y_I が互いに独立になるようにする。

独立成分分析 ー背景

■歴史

- 1980年代：フランスの研究者ら中心
 - ◆ 非線形無相関化, 高次統計量
- 1990年代中盤から世界的な広がり
 - ◆ 理論的枠組の充実, 効率的なアルゴリズム
- 国際会議ICA: 1999年から1年毎に1回開催

■応用

- 信号分離: 音, 脳波, 無線信号, など
 - ◆ ブラインド音源分離 (BlindSourceSeparation)
- 特徴抽出: 自然画像, 音, など

目次

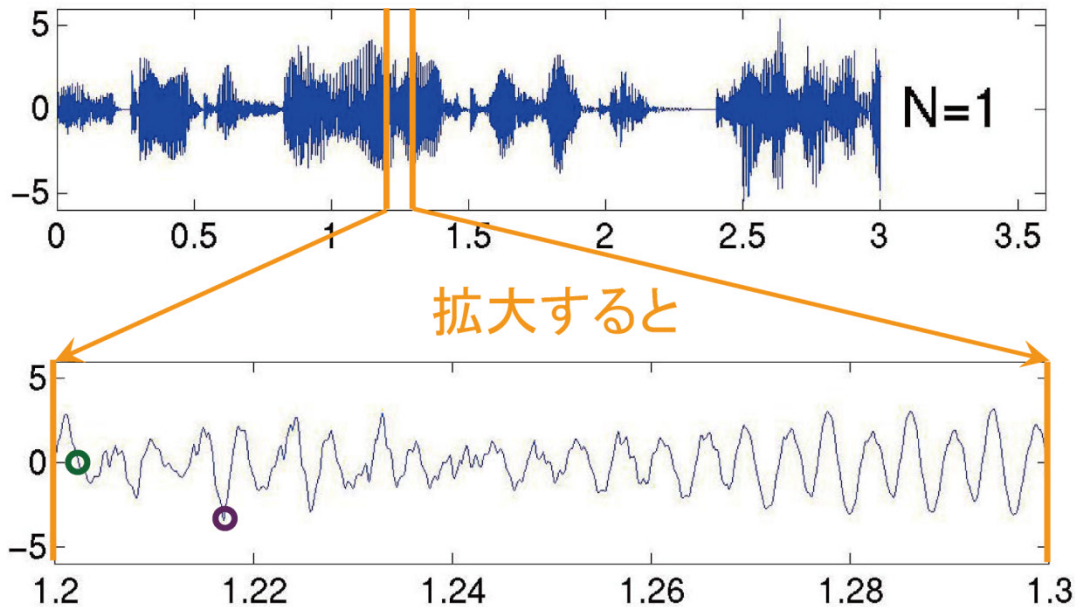
- 独立成分分析について
 - 定式化, 歴史, 応用
- 信号の統計的性質
 - 信号を混ぜるー 中心極限定理
- 独立成分分析のアルゴリズム
 - 白色化 + FastICA
 - 最尤推定法 by Natural Gradient

目次

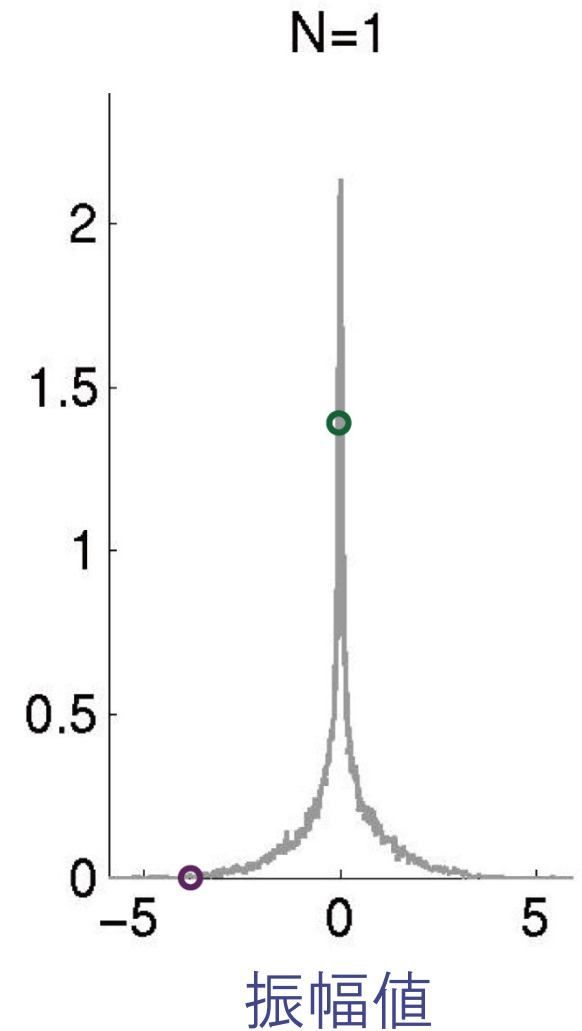
- 独立成分分析について
 - 定式化, 歴史, 応用
- 信号の統計的性質
 - 信号を混ぜるー 中心極限定理
- 独立成分分析のアルゴリズム
 - 白色化 + FastICA
 - 最尤推定法 by Natural Gradient

音の統計的性質

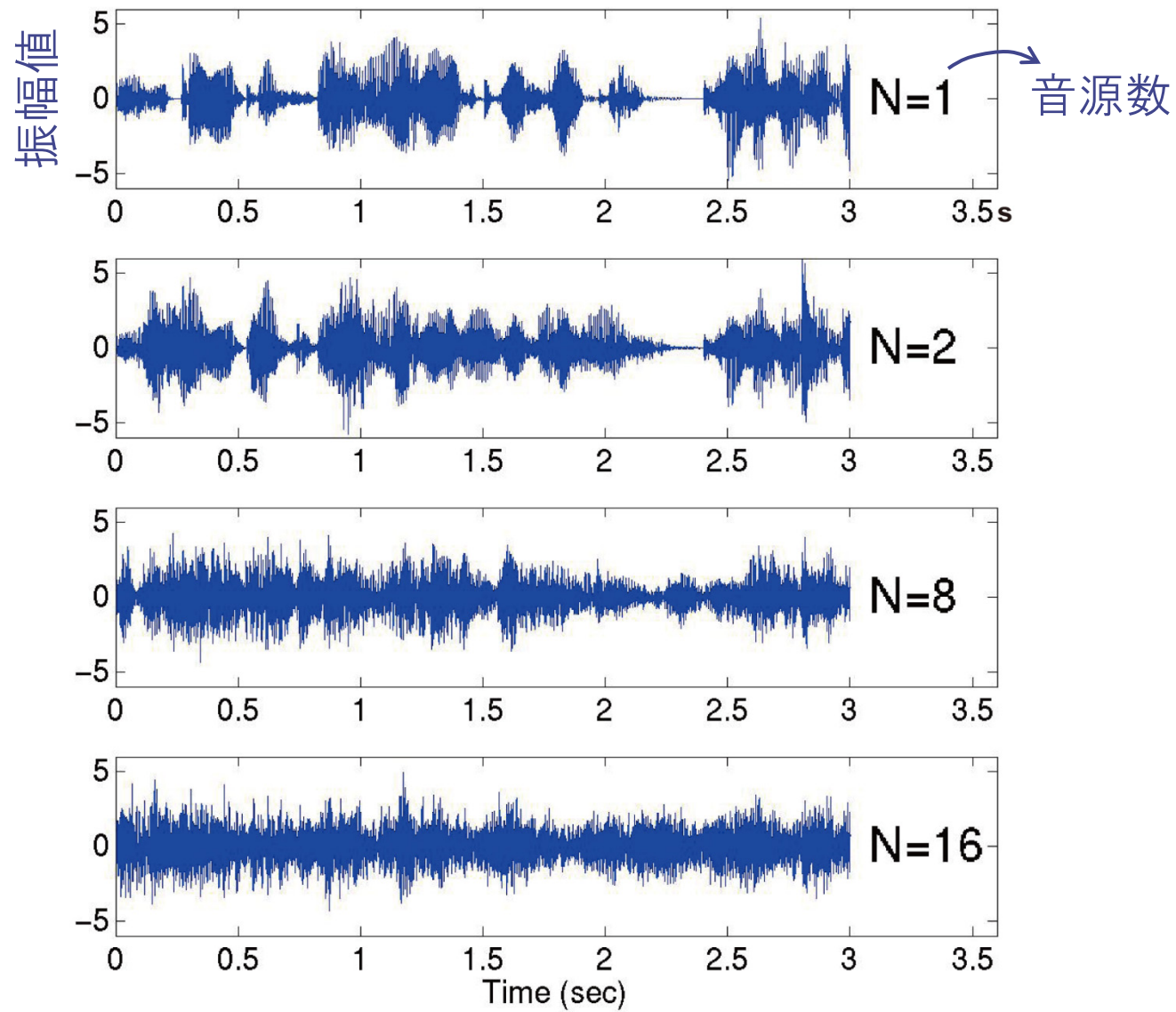
音声の波形



振幅値のヒストグラム

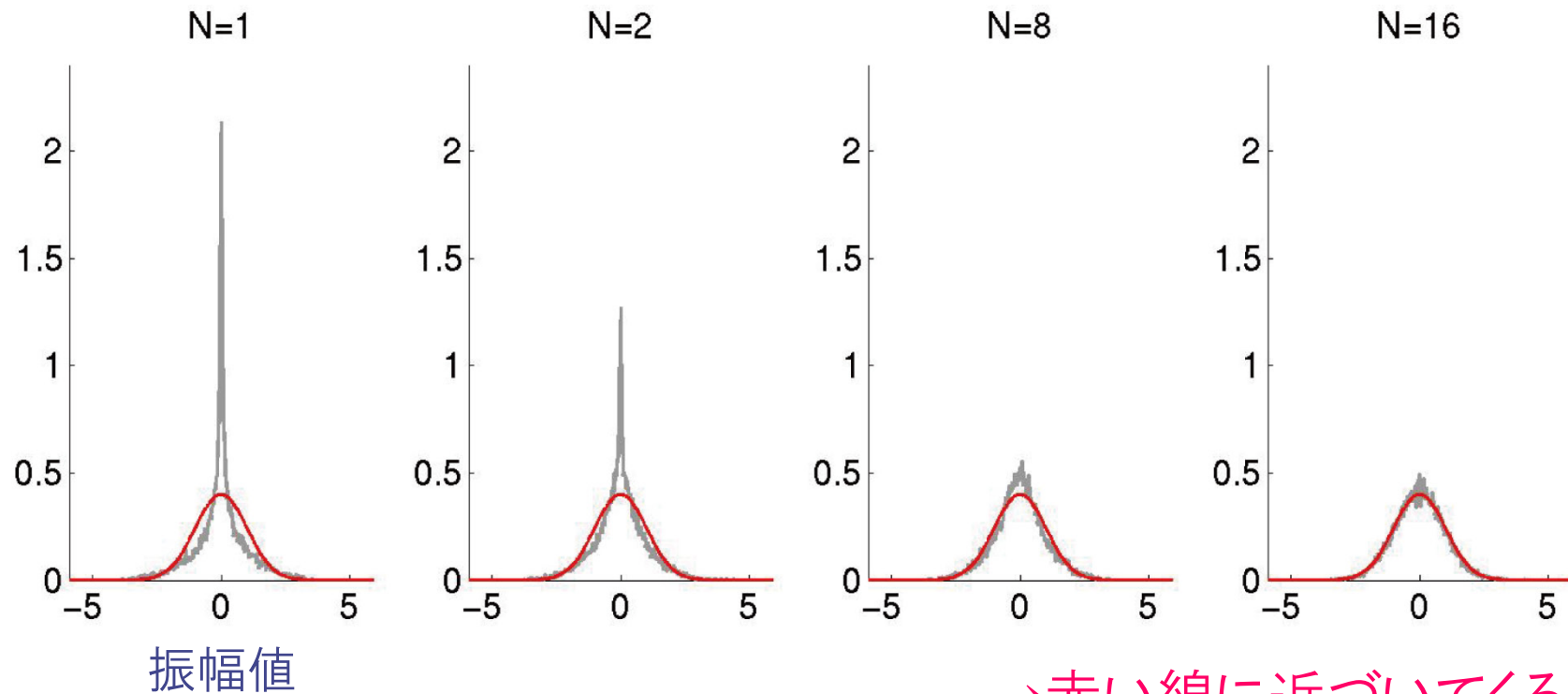


音を混ぜてみる



混ぜた音の統計的性質

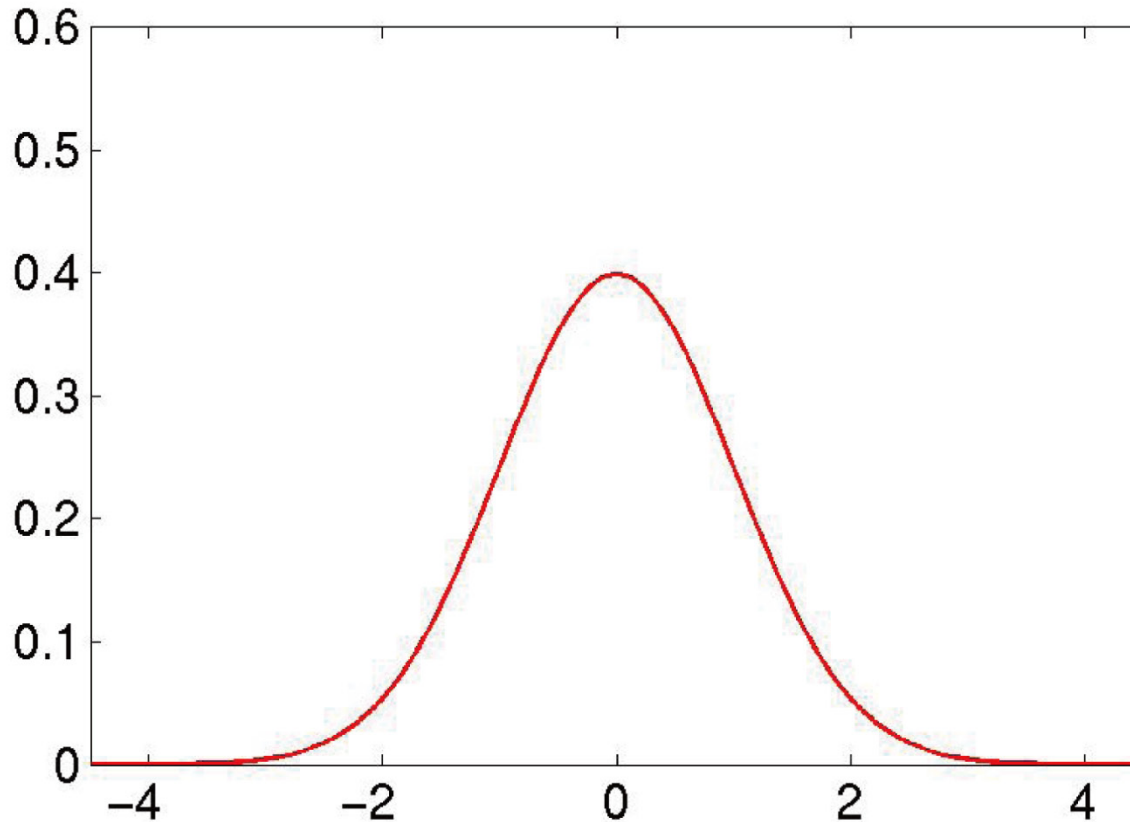
■ 振幅値のヒストグラム



→ 赤い線に近づいてくる

中心極限定理 (Central Limit Theorem)

- 元々の信号がどのような統計的性質をもったものであれ、多くの信号を足し合わせると、その振幅値の分布は正規分布に近づく。



平均0
分散1の
正規分布

中心極限定理 (Central Limit Theorem)

- 確率変数 x_1, x_2, \dots, x_N が独立に同一な確率密度関数 $p_x(x)$ に従うならば, $z = (x_1 + x_2 + \dots + x_N) / \sqrt{N}$ の確率密度関数は $N \rightarrow \infty$ でガウス分布に近づく
- ただし, x の平均は0とする

中心極限定理の略証 (1/2)

■ x の特性関数を $\Theta_x(ju)$ とする

■ $y = x_1 + x_2 + \cdots + x_N$ の特性関数 $\Theta_y(ju)$

$$\Theta_y(ju) = \Theta_x(ju)^N \quad (\because \text{重畳積分定理})$$

■ $z = y/\sqrt{N}$ の確率密度関数 $p_z(z)$

$$p_z(z) = \sqrt{N}p_y(\sqrt{N}z)$$

■ $z = y/\sqrt{N}$ の特性関数 $\Theta_z(ju)$

$$\Theta_z(ju)$$

$$= \int_{-\infty}^{\infty} p_z(z)e^{jzu} dz = \int_{-\infty}^{\infty} \sqrt{N}p_y(\sqrt{N}z)e^{jzu} dz$$

$$= \int_{-\infty}^{\infty} p_y(y)e^{jy(u/\sqrt{N})} dy = \Theta_y(ju/\sqrt{N}) = \Theta_x(ju/\sqrt{N})^N$$

中心極限定理の略証 (2/2)

- $\log \Theta_x(ju)$ のマクローリン展開

$$\log \Theta_x(ju) = \sum_{n=0}^{\infty} \underbrace{k_n}_{n\text{次キユムラント}} \frac{(ju)^n}{n!}$$

- $\log \Theta_z(ju)$ のマクローリン展開

$$\begin{aligned} \log \Theta_z(ju) &= N \log \Theta_x(ju/\sqrt{N}) \\ &= \sum_{n=0}^{\infty} \underbrace{\left(\frac{Nk_n}{N^{n/2}} \right)}_{z\text{のキユムラントを} \\ &\quad x\text{のキユムラントで表したもの}} \frac{(ju)^n}{n!} \end{aligned}$$

- 3次以上のキユムラントは $N \rightarrow \infty$ で0に収束
⇒ 分布がガウス分布に収束

(∵ ガウス分布に従う確率変数の3次以上のキユムラントはすべて0)

正規分布(Gauss分布)

■ 確率密度関数

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

■ 統計学における最も重要な分布の一つ

■ さまざまな性質がある

- 平均 μ と分散 σ^2 によって確率密度関数が一意に決まる
- 最も「ランダム」な分布
 - ◆ 同じ分散をもつ分布の中でエントロピーが最大
- 中心極限定理

エントロピー(平均情報量)

■ある事象の情報量

$$I(x) = \log \frac{1}{p(x)} = -\log p(x)$$

めったに起こらないこと
ほど情報量は大きい

■エントロピー

$$H(x) = \mathbb{E}\{-\log p(x)\}$$

- 各事象の情報量の平均値
- 事象のランダムさを表す指標
- 正規分布の場合

$$H(x) = \log \sqrt{2\pi\sigma^2} + \frac{\mathbb{E}\{(x - \mu)^2\}}{2\sigma^2} = \frac{1}{2} \log 2\pi\sigma^2 + \frac{1}{2}$$

エントロピーが最大になる確率密度関数

■問題:

■ Maximize $H(x) = - \int p(x) \log p(x) dx$ with respect to $p(x)$

Subject to $\int p(x) dx = 1, \int xp(x) dx = 0, \int x^2 p(x) dx = 1$

■ラグランジュ未定乗数法

■ ラグランジアンをpに関して偏微分して0と置く

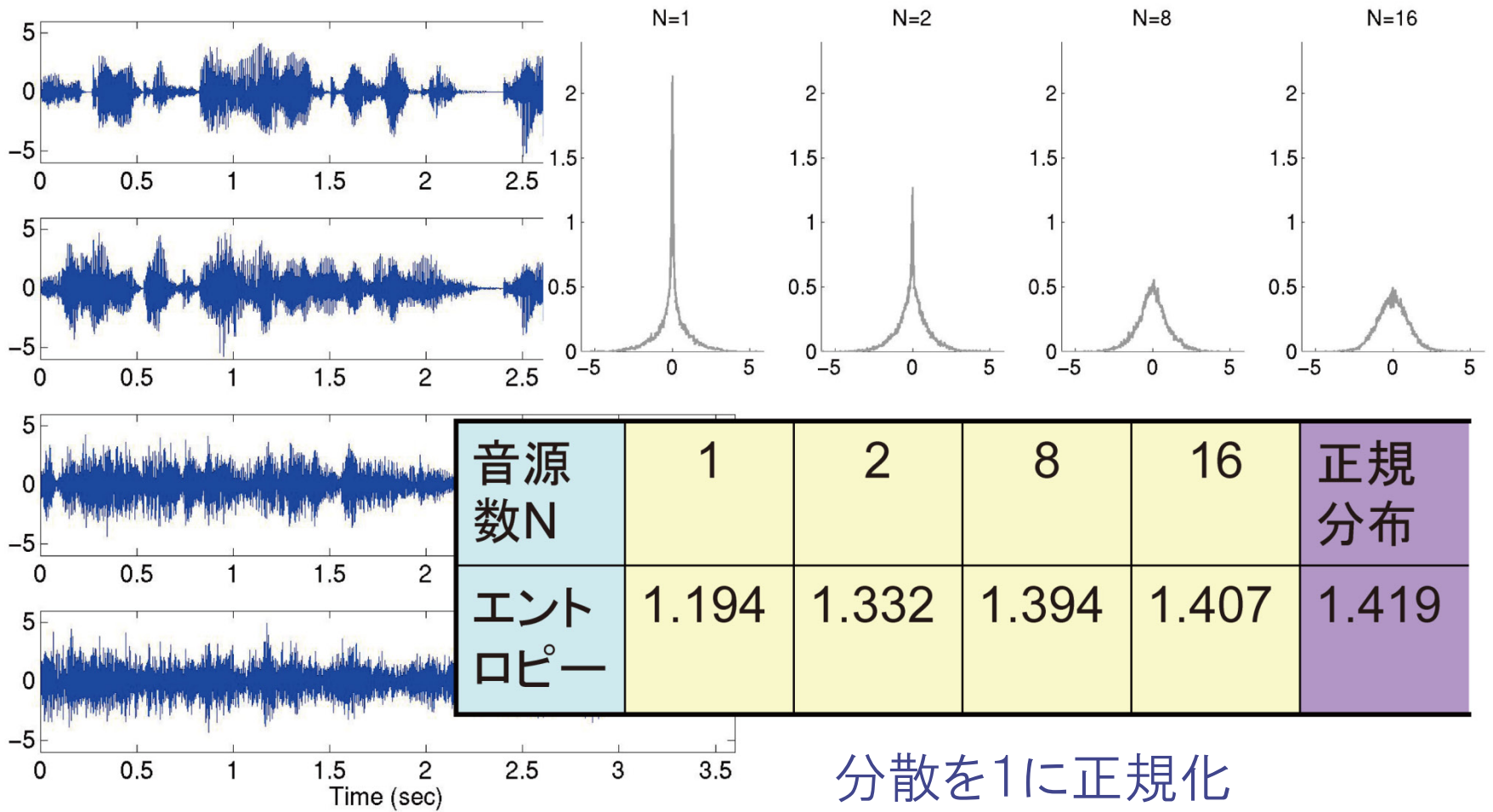
$$\begin{aligned} \mathcal{L}[p] = & - \int p(x) \log p(x) dx \\ & + \lambda_0 \left(\int p(x) dx - 1 \right) + \lambda_1 \left(\int xp(x) dx \right) + \lambda_2 \left(\int x^2 p(x) dx - 1 \right) \end{aligned}$$

$$\frac{\partial \mathcal{L}}{\partial p} = -1 - \log p(x) + \lambda_0 + \lambda_1 x + \lambda_2 x^2 = 0$$

$$\Rightarrow p(x) = \exp(1 - \lambda_0 - \lambda_1 x - \lambda_2 x^2)$$

たしかに正規分布型になっている！

混ぜた音のエントロピー



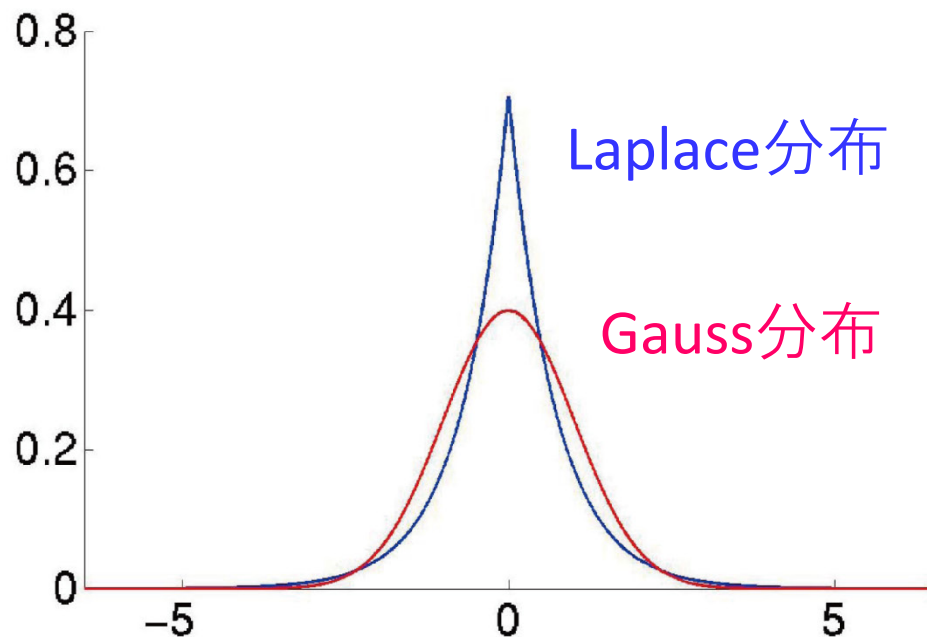
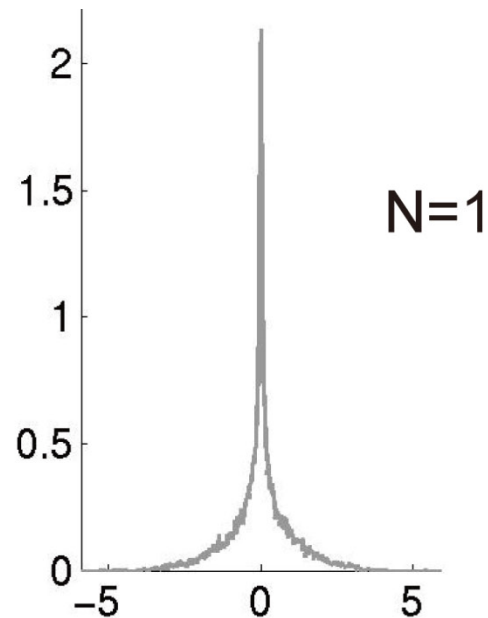
音源の統計的性質

- 我々の身の回りにある音（音声や楽音等）の波形は振幅値0の頻度が高い

- Laplace分布によるモデル化

$$p(x) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right)$$

平均: $\mu = 0$
分散: $2b^2 = 1$



エントロピーの近似計算

- 真の分布を良く近似する分布を用いてエントロピーを計算

$$H(x) = \mathbb{E}\{-\log p(x)\} \simeq -\frac{1}{T} \sum_{t=1}^T \log p(x(t))$$

音源数N	1	2	8	16	正規分布
エントロピー	1.194	1.332	1.394	1.407	1.419
ラプラス分布による近似	1.286	1.356	1.420	1.457	1.474
ガウス分布による近似	1.419	1.419	1.419	1.419	1.419

良好→

分散正規化のため変化無し

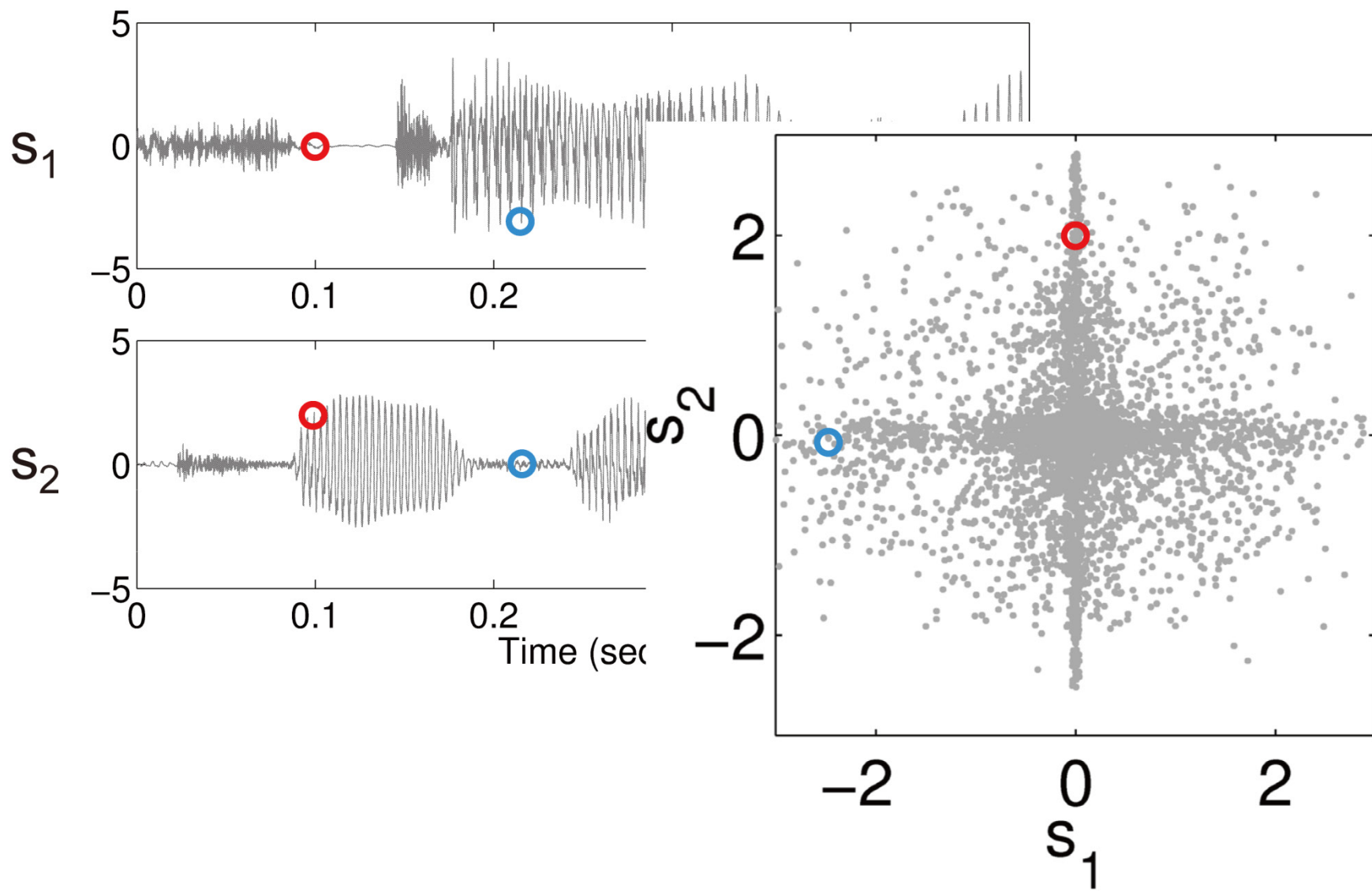
目次

- 独立成分分析について
 - 定式化, 歴史, 応用
- 信号の統計的性質
 - 信号を混ぜるー 中心極限定理
- 独立成分分析のアルゴリズム
 - 白色化 + FastICA
 - 最尤推定法 by Natural Gradient

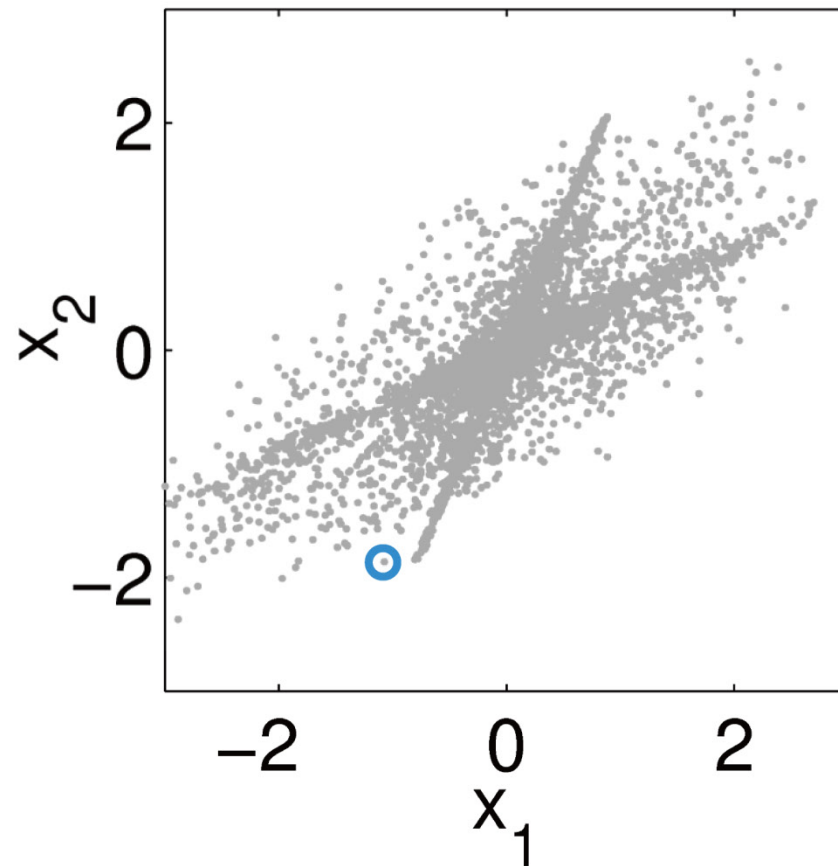
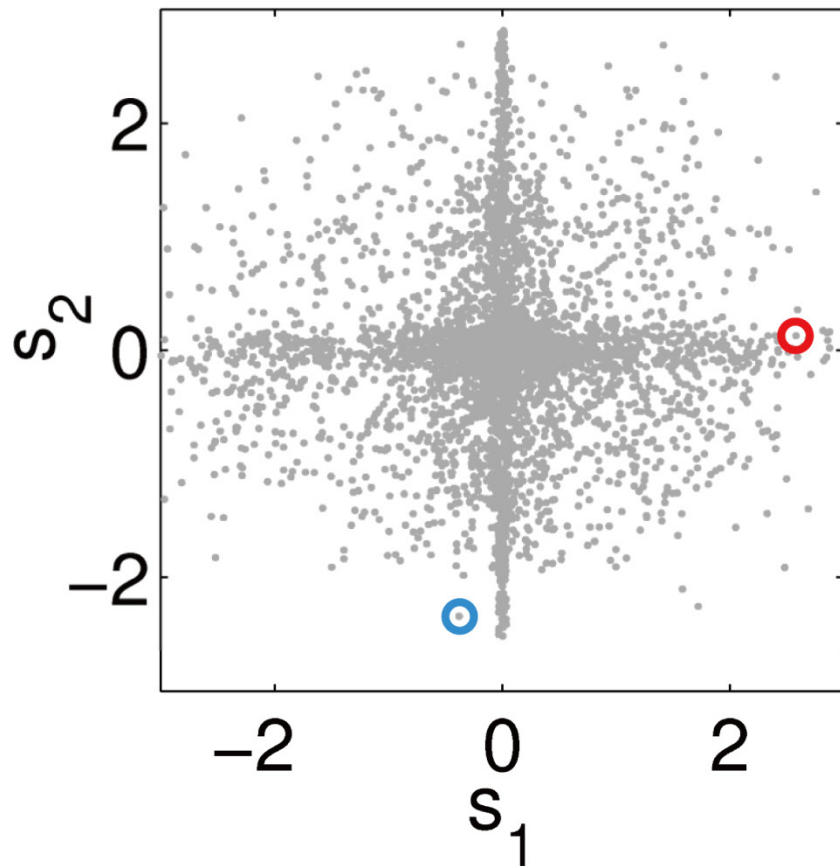
目次

- 独立成分分析について
 - 定式化, 歴史, 応用
- 信号の統計的性質
 - 信号を混ぜるー 中心極限定理
- 独立成分分析のアルゴリズム
 - 白色化 + FastICA
 - 最尤推定法 by Natural Gradient

音源信号の散布図



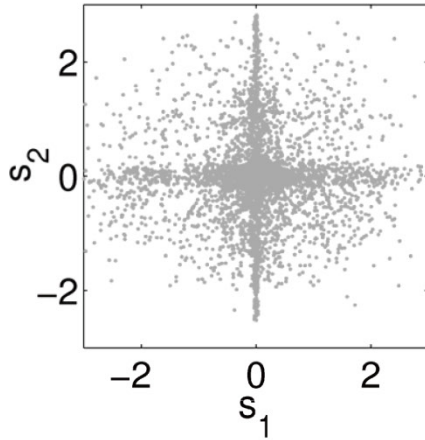
音源信号と混合信号の散布図



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.92 & 0.31 \\ 0.41 & 0.73 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

白色化 + FastICA

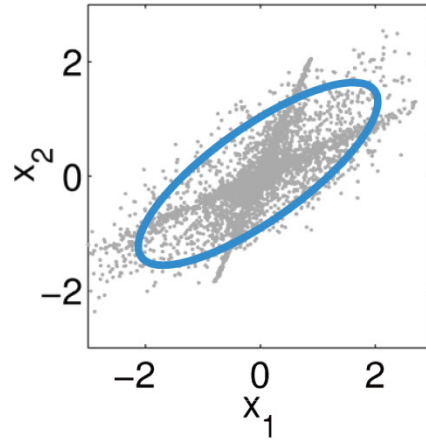
音源信号



H

混合行列
(未知)

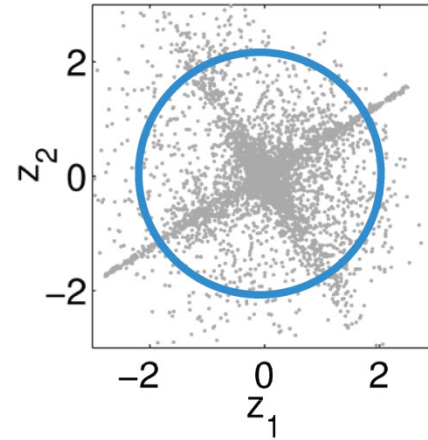
混合信号



V

白色化行列
(楕円を円に)

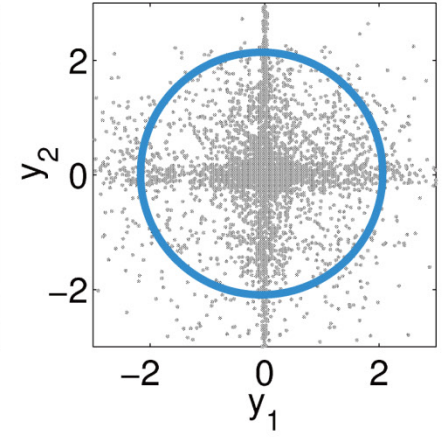
白色化信号



U

ユニタリ行列
(回転)

分離信号



FastICA

相関と相関行列

■ 相関

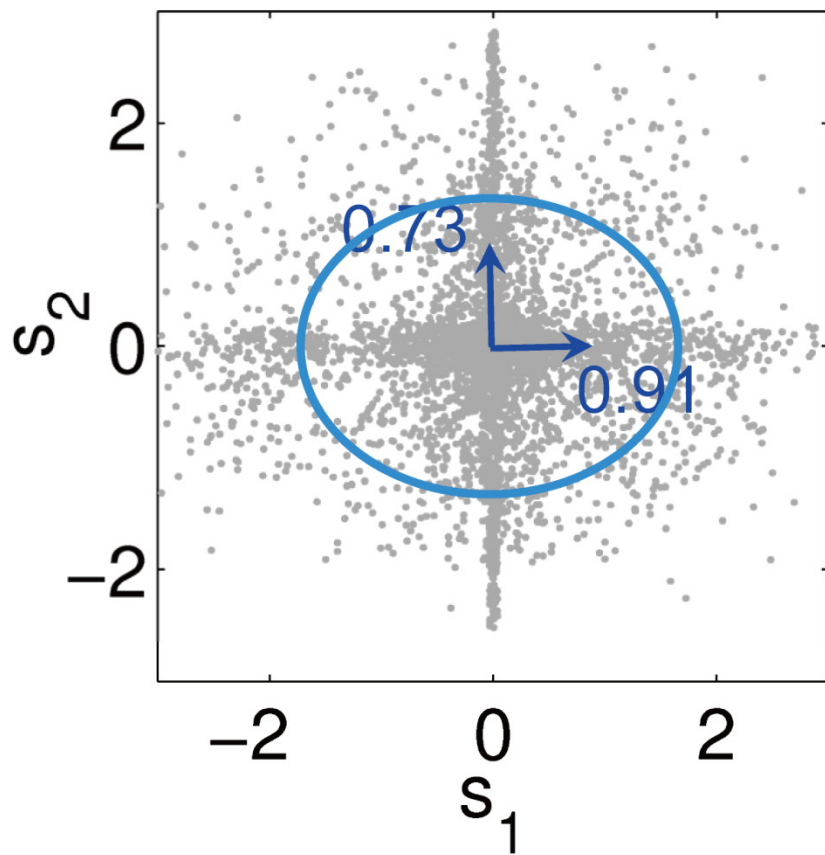
$$r_{i,j} = \mathbb{E}\{x_i x_j\} \simeq \frac{1}{T} \sum_{t=1}^T x_i(t) x_j(t)$$

■ 相関行列

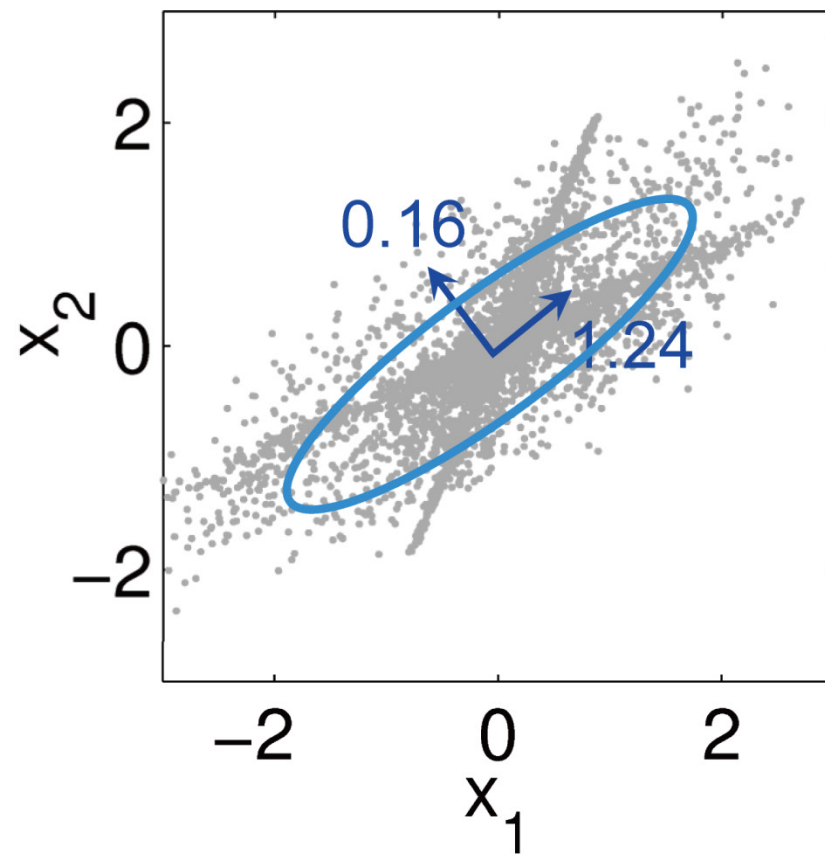
$$\mathbf{R}_x = \mathbb{E}\{\mathbf{x} \mathbf{x}^T\} = \begin{bmatrix} \mathbb{E}\{x_1 x_1\} & \cdots & \mathbb{E}\{x_1 x_J\} \\ \vdots & \ddots & \vdots \\ \mathbb{E}\{x_J x_1\} & \cdots & \mathbb{E}\{x_J x_J\} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_J \end{bmatrix}$$

相関行列の固有ベクトルと固有値

$$R_s = \begin{bmatrix} 0.91 & 0.01 \\ 0.01 & 0.73 \end{bmatrix}$$



$$R_x = \begin{bmatrix} 0.85 & 0.52 \\ 0.52 & 0.55 \end{bmatrix}$$



白色化

- 相関行列が単位行列になるように $x(t)$ を変換

$$z(t) = \mathbf{V} x(t)$$

$$\mathbf{R}_z = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

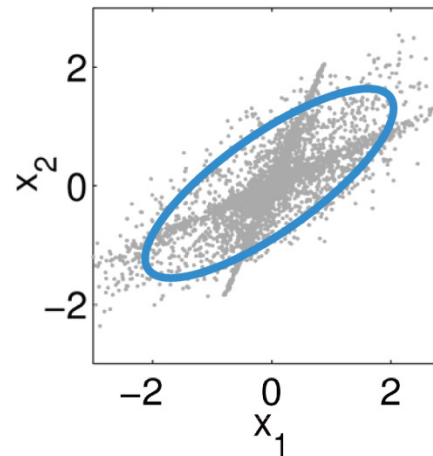
- 無相関化

$$\mathbb{E}\{z_i z_j\} = 0$$

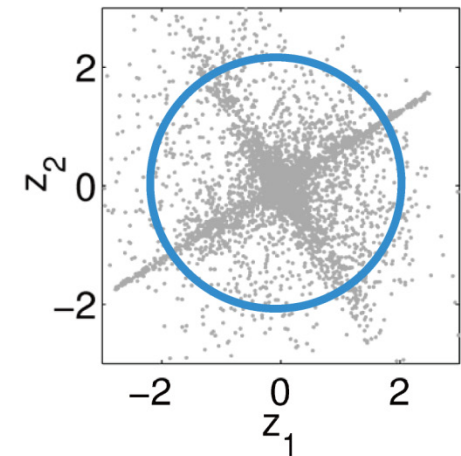
- 分散の正規化

$$\mathbb{E}\{z_i z_i\} = 1$$

混合信号



白色化信号



\mathbf{V}

白色化行列
(楕円を円に)

白色化行列の求め方

- $z = Vx$ により $R_z = I$ となる白色化行列 V を求めたい
 - z の相関行列を計算してみると...

$$R_z = \mathbb{E}\{zz^T\} = \mathbb{E}\{Vxx^TV^T\} = VR_xV^T$$

- x を固有値分解してみると...

$$R_x = EDE^T$$

相関行列の固有値分解

$$R = EDE^T$$

$$Re_i = \lambda_i e_i$$

$$D = \text{diag}(\lambda_1, \dots, \lambda_J)$$

$$E = [e_1 \cdots e_J] : E^T E = I \text{ を満たす}$$

正規直交基底行列

- $R_z = VR_xV^T$ を単位行列にしたいわけなので $V = D^{-1/2}E^T$

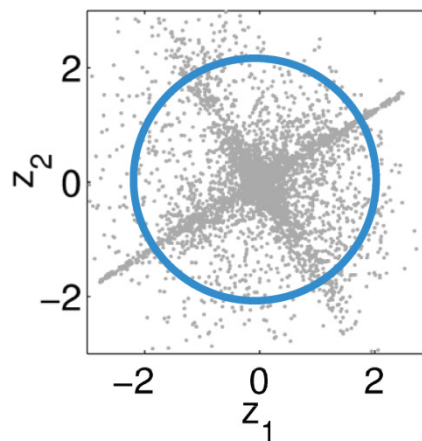
ユニタリ変換(回転)

- 相関行列が単位行列になるように $z(t)$ を変換

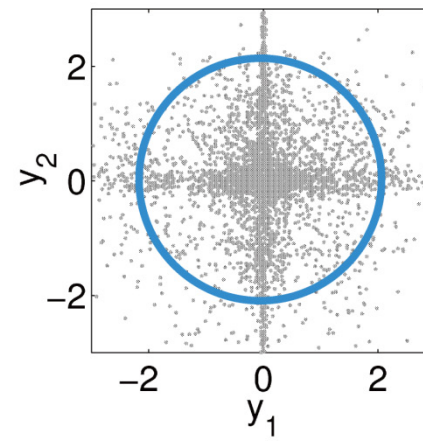
$$\mathbf{y}(t) = \mathbf{U} \mathbf{z}(t)$$

$$\mathbf{U} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

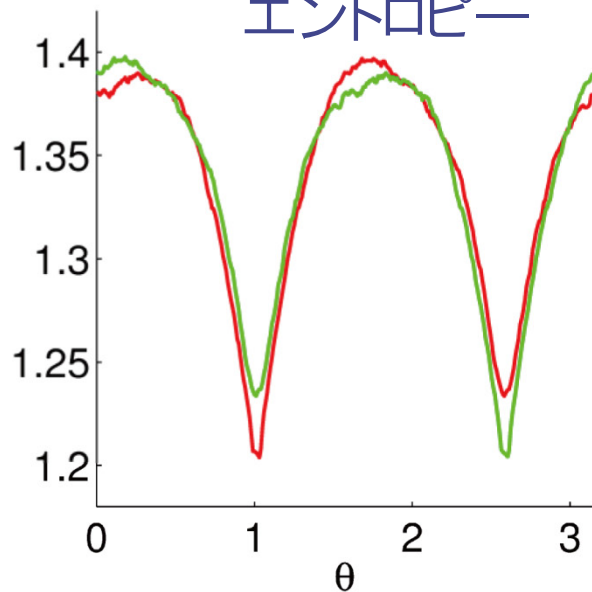
白色化信号



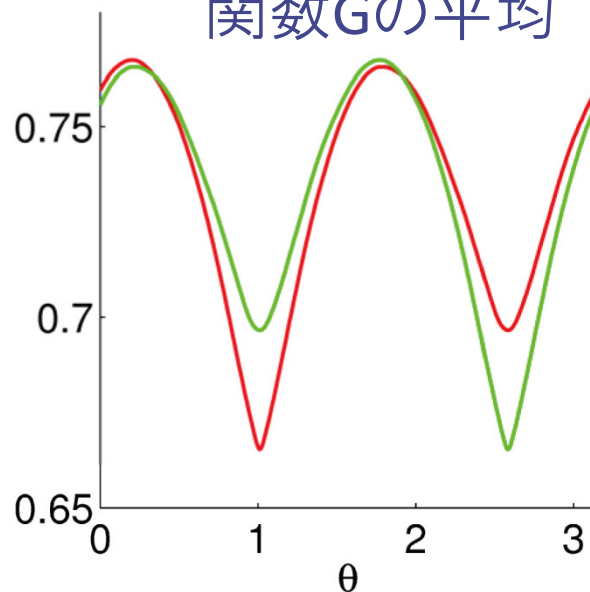
分離信号



エントロピー



関数Gの平均



$$G(y_i) = -\log p(y_i)$$

$$p(y_i) \propto \exp(-|y_i|)$$

Laplace分布による
情報量の近似

FastICAアルゴリズム

- 非線形関数 $G(y_i) = -\log p(y_i)$ の平均を最小化
 - 求めたいのは $U = [\mathbf{u}_1 \cdots \mathbf{u}_i]^T$
 - 解空間をユニタリ行列に限定

各 i について収束するまで以下を繰り返す

$$y_i \leftarrow \mathbf{u}_i^T \mathbf{z} \quad \text{分離信号の計算}$$

$$\mathbf{u}_i \leftarrow \mathbb{E}\{G''(y_i)\} \mathbf{u}_i - \mathbb{E}\{G'(y_i) \mathbf{z}\} \quad \text{Newton法による降下}$$

$$\mathbf{u}_i \leftarrow \mathbf{u}_i - \sum_{k=1}^{i-1} (\mathbf{u}_k^T \mathbf{u}_i) \mathbf{u}_k \quad \text{Gram-Schmidtの直交化}$$

$$\mathbf{u}_i \leftarrow \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|_2} \quad \text{ノルムの正規化}$$

非線形関数Gについて

- FastICAでは一階微分と二階微分が必要
(Newton法を使うため)
- Laplace分布を用いて定義したGは不連続のため不都合

$$G(y) = |y| \rightarrow \text{微分できない}$$

- 代わりに微分可能な近似関数を使用することが多い

$$G(y) = \sqrt{y^2 + \alpha} \quad \rightarrow \alpha = 0 \text{ で } |y| \text{ となる}$$

$$G'(y) = \frac{y}{\sqrt{y^2 + \alpha}}$$

$$G''(y) = \frac{1}{\sqrt{y^2 + \alpha}} \left(1 - \frac{y^2}{y^2 + \alpha} \right)$$

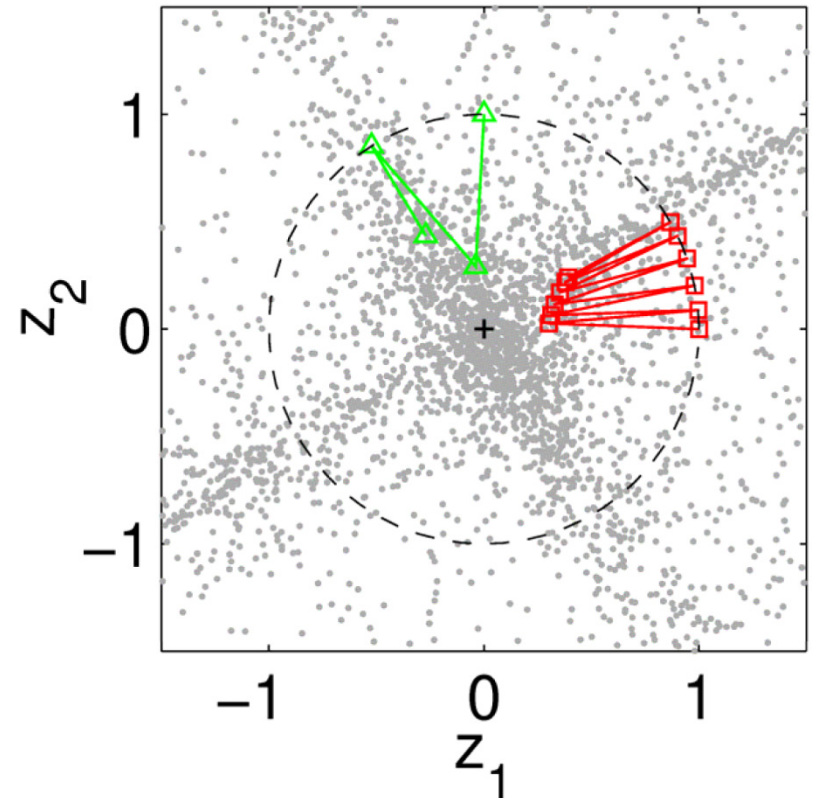
FastICAアルゴリズムによる推定の様子

■ 赤(□)

- $\mathbf{u}_1 = [1 \ 0]^T$ を初期値に設定
- Newton法による更新: 原点に向かっている
- ノルム1に正規化: 単位円上に引き戻されている
- 5回の繰り返しで良好な解へ

■ 緑(△)

- $\mathbf{u}_2 = [0 \ 1]^T$ を初期値に設定
- 直交化により, 1回だけで解に到達



白色化+FastICAのまとめ

■手順

- 観測信号を白色化(観測信号の相関行列の固有値分解)

$$\mathbf{z}(t) = \mathbf{V} \mathbf{x}(t)$$

- ユニタリ行列を射影勾配法(Newton法+解空間への射影)

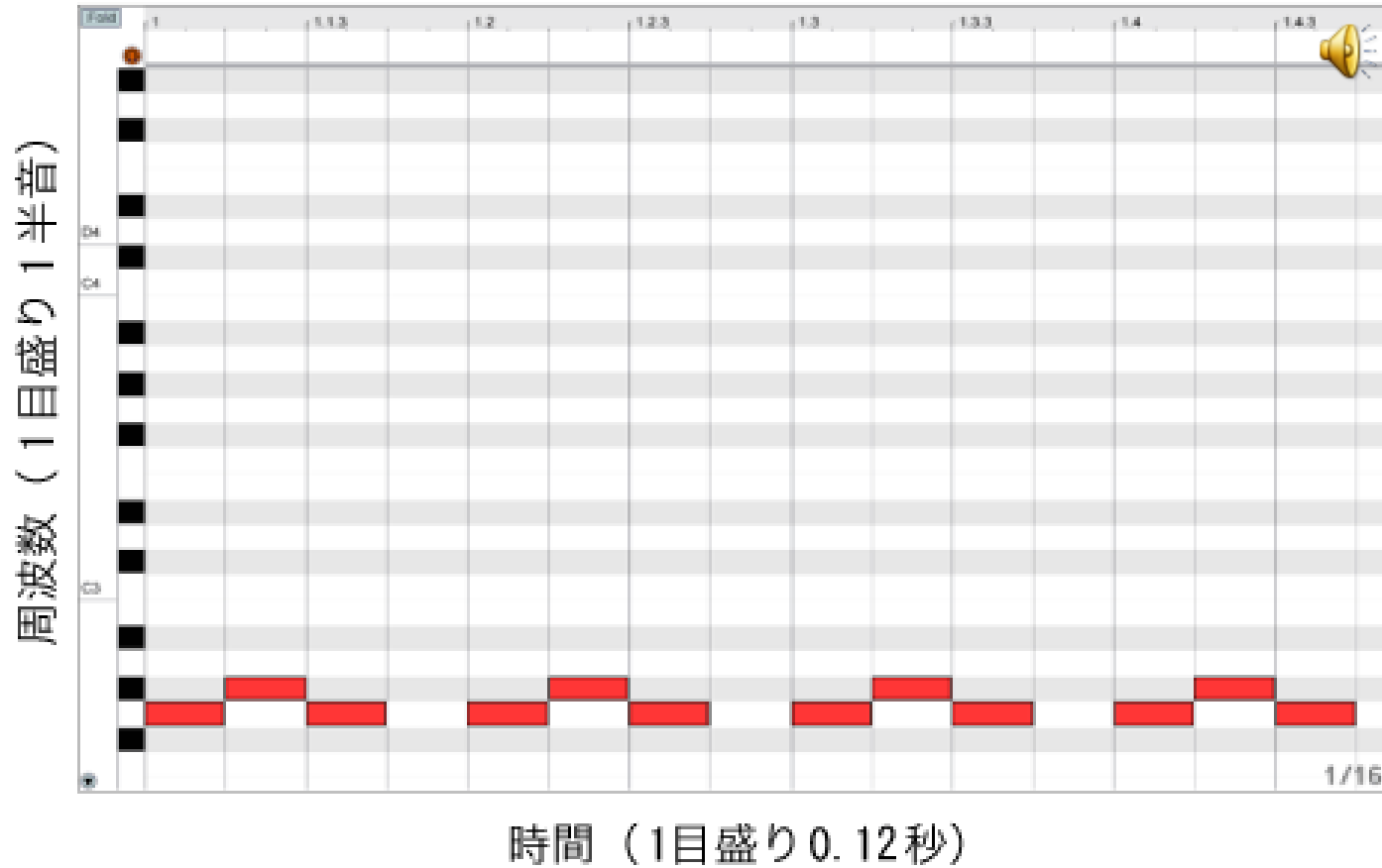
$$\mathbf{y}(t) = \mathbf{U} \mathbf{z}(t)$$

- 求めたかったのは分離行列: $\mathbf{y}(t) = \mathbf{W} \mathbf{x}(t)$

$$\mathbf{W} = \mathbf{U} \mathbf{V} \quad (\text{限定された形})$$

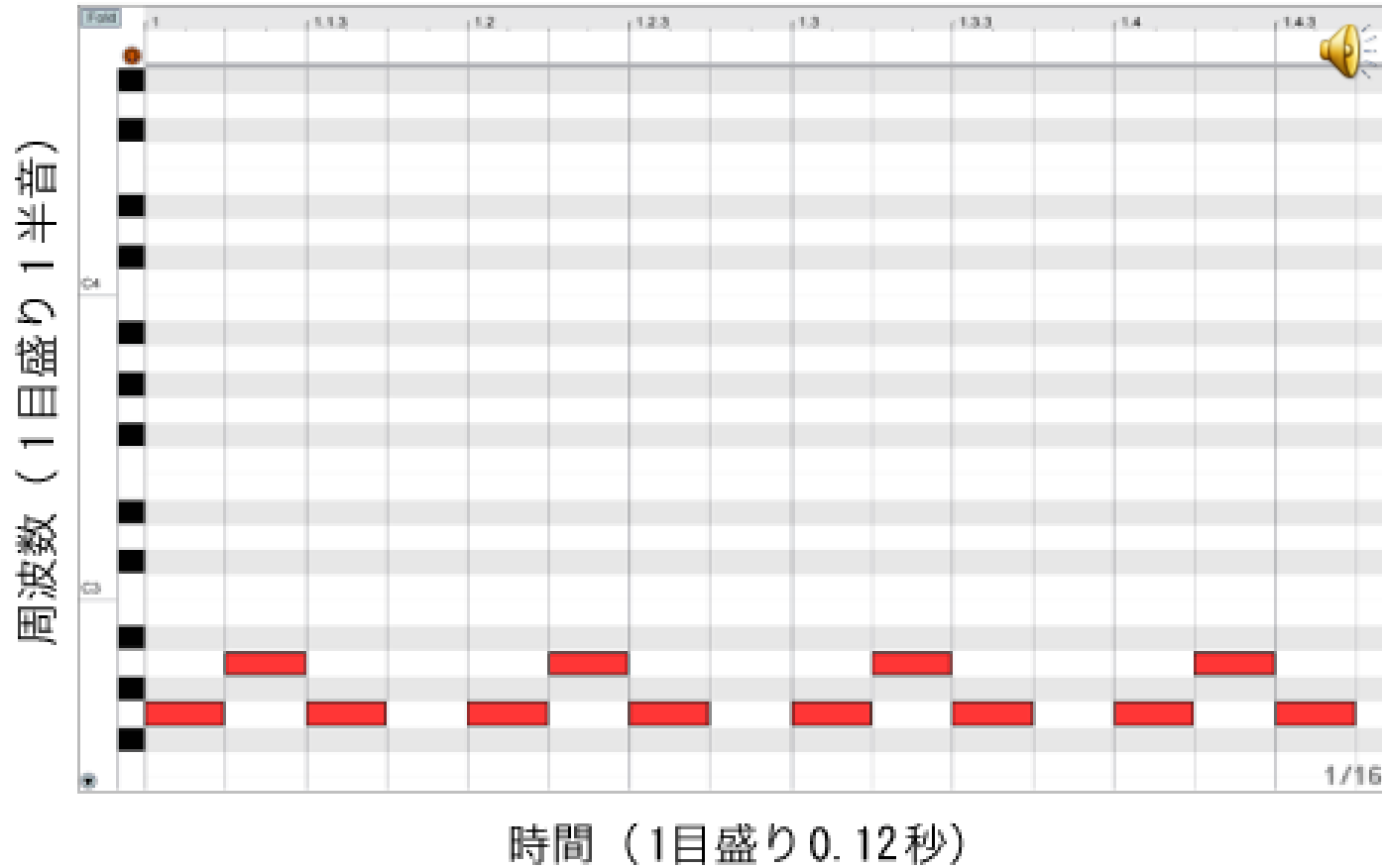
Coffee Break: 音脈分凝について

周波数差：1 半音



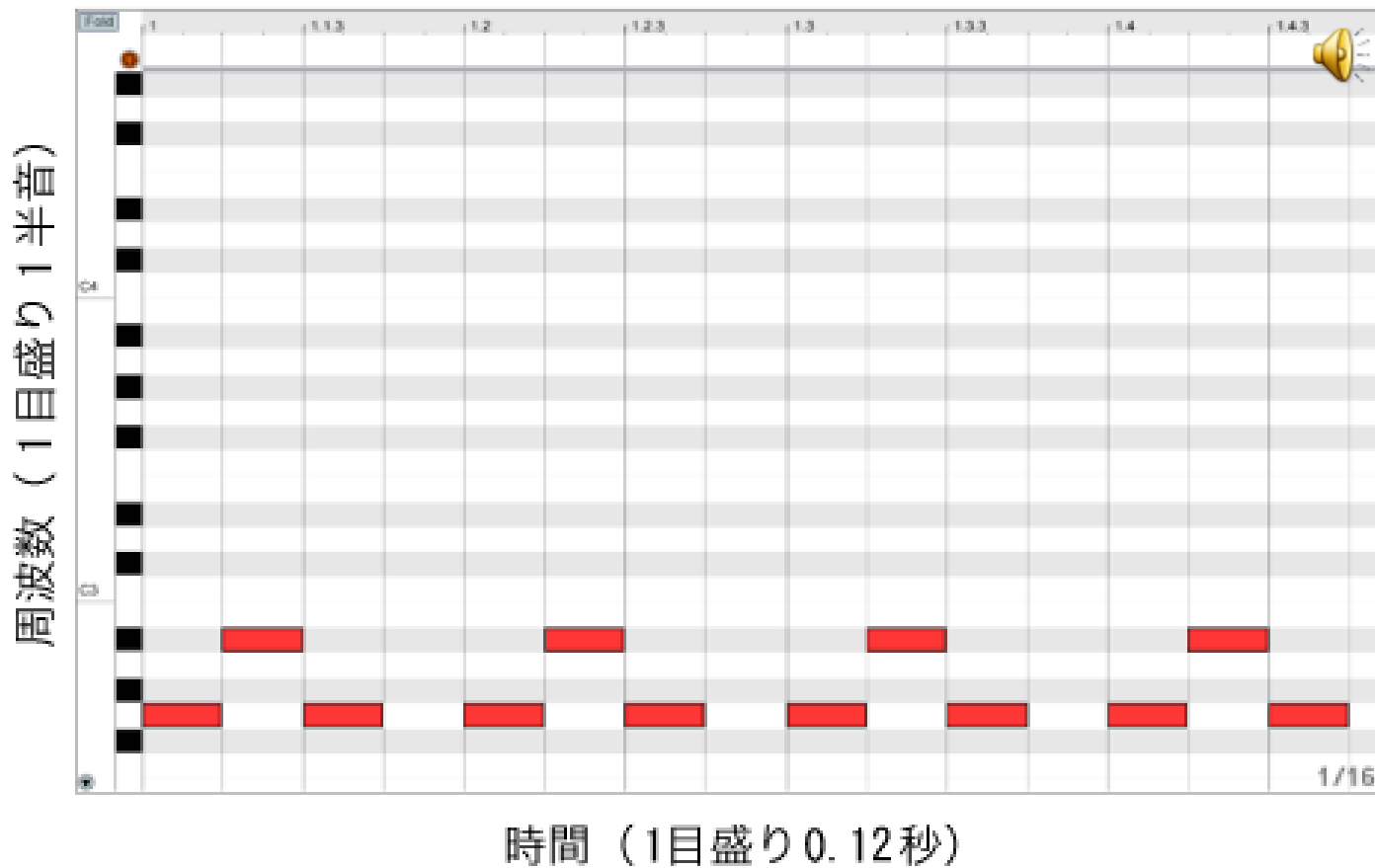
Coffee Break: 音脈分凝について

周波数差：2半音



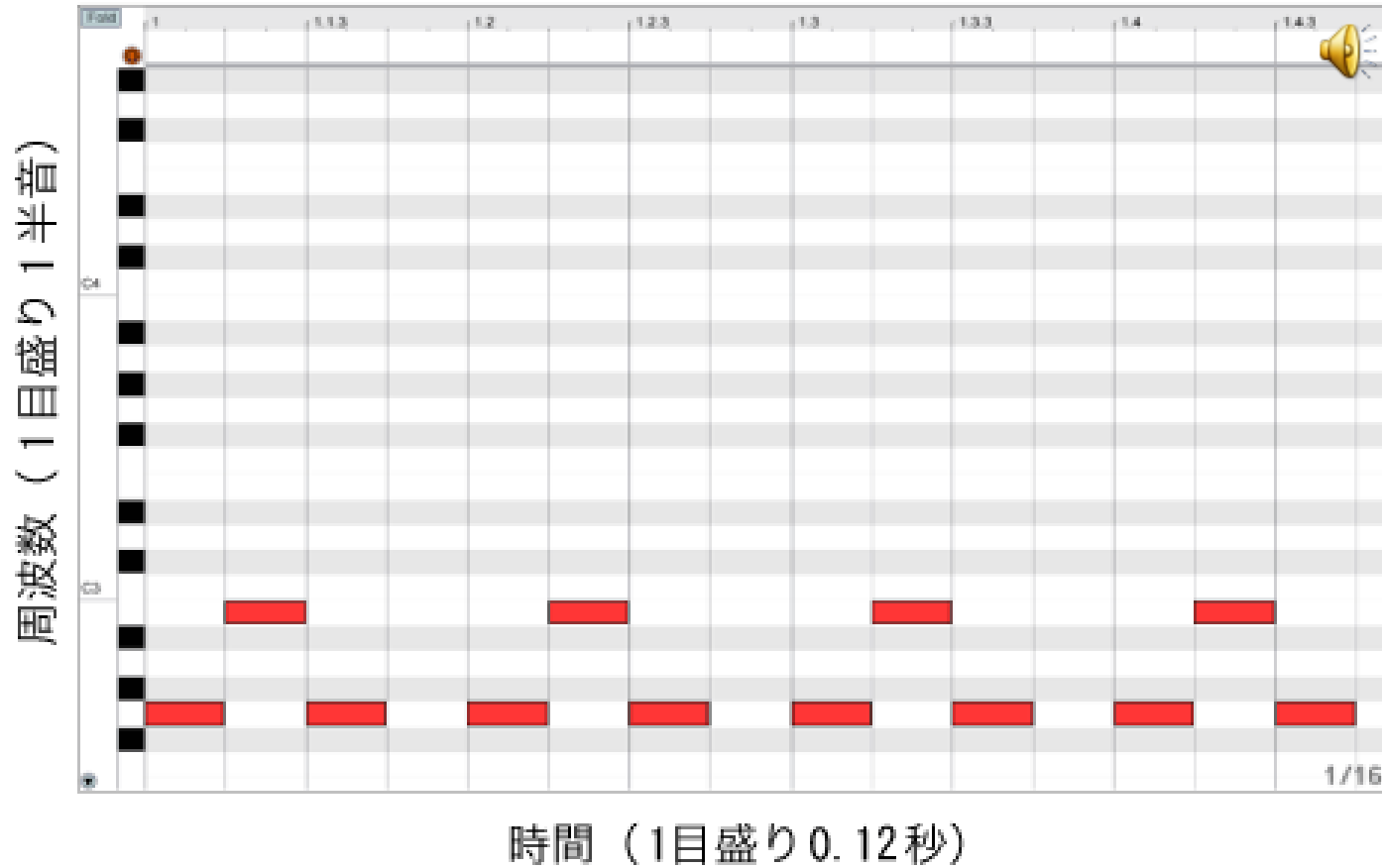
Coffee Break: 音脈分凝について

周波数差：3半音



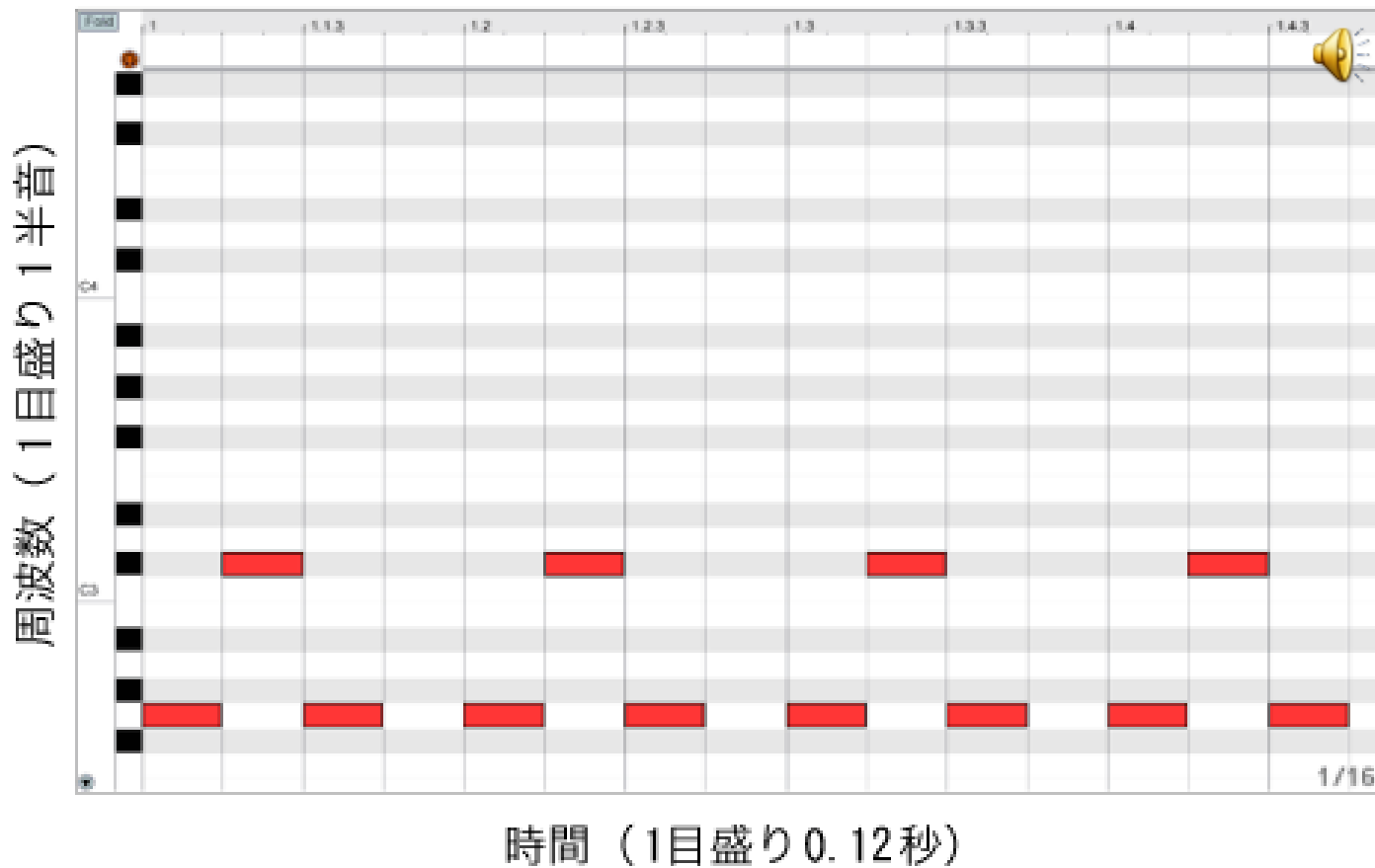
Coffee Break: 音脈分凝について

周波数差：4半音



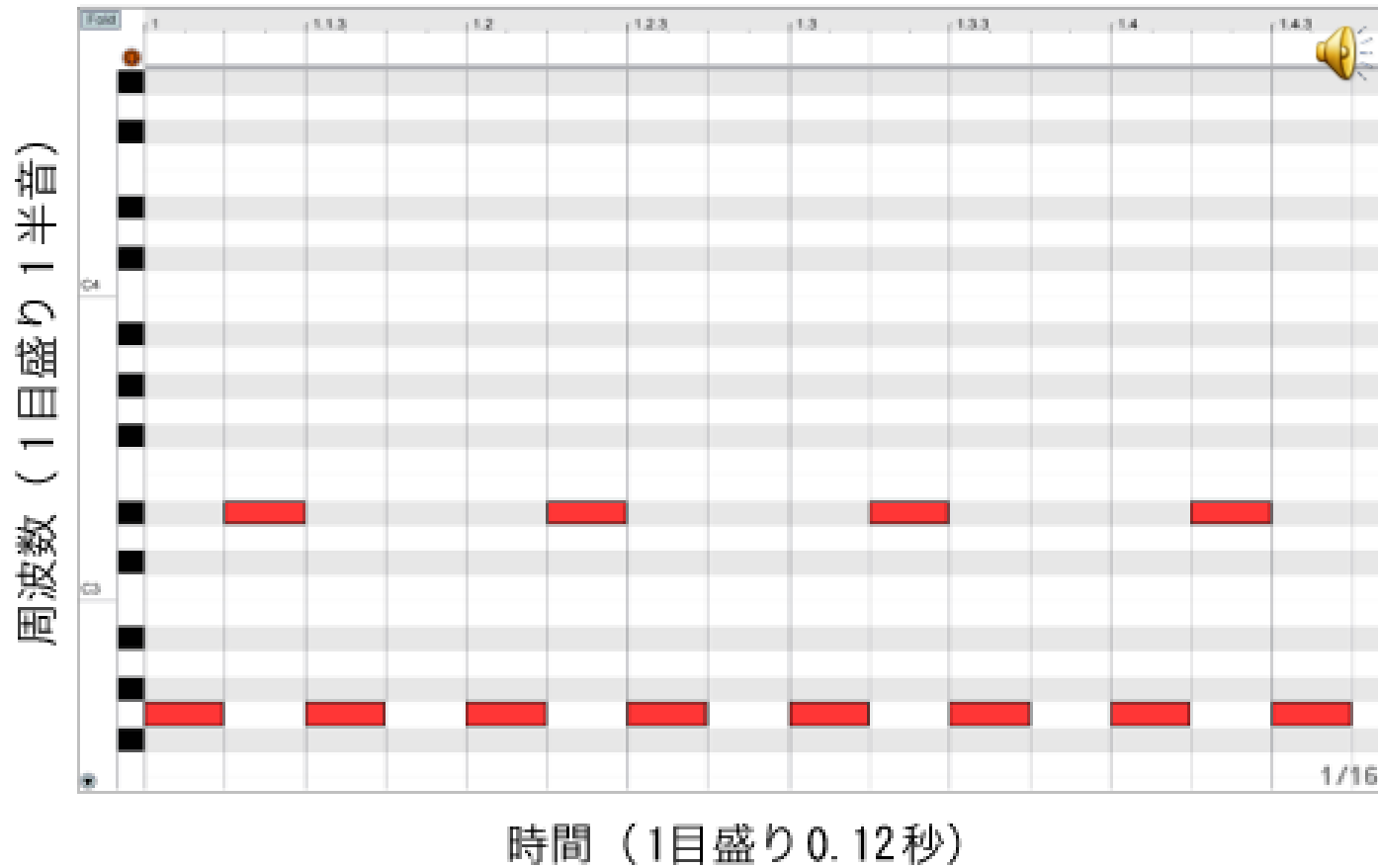
Coffee Break: 音脈分凝について

周波数差：6半音



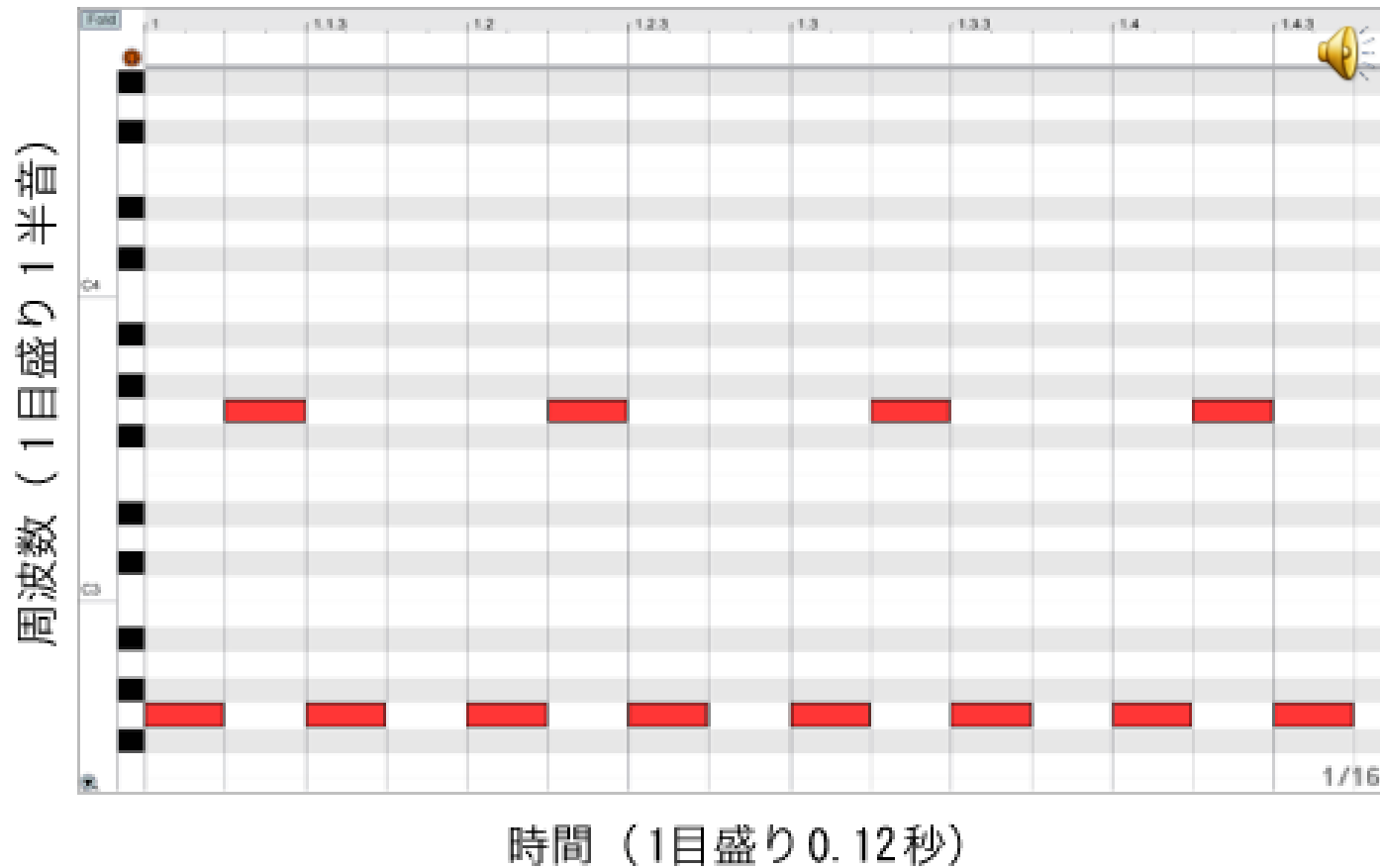
Coffee Break: 音脈分凝について

周波数差：8半音



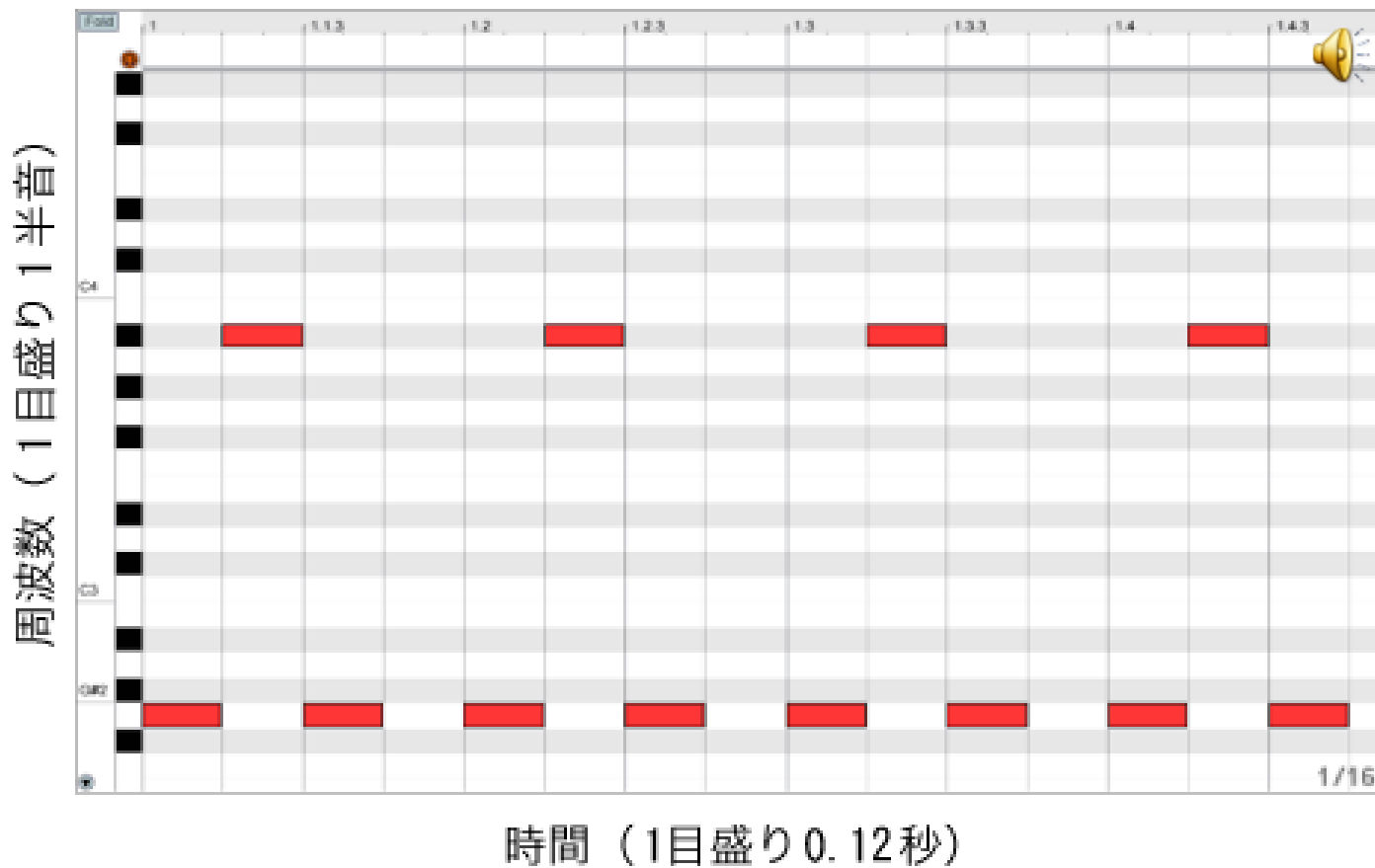
Coffee Break: 音脈分凝について

周波数差：1 2半音



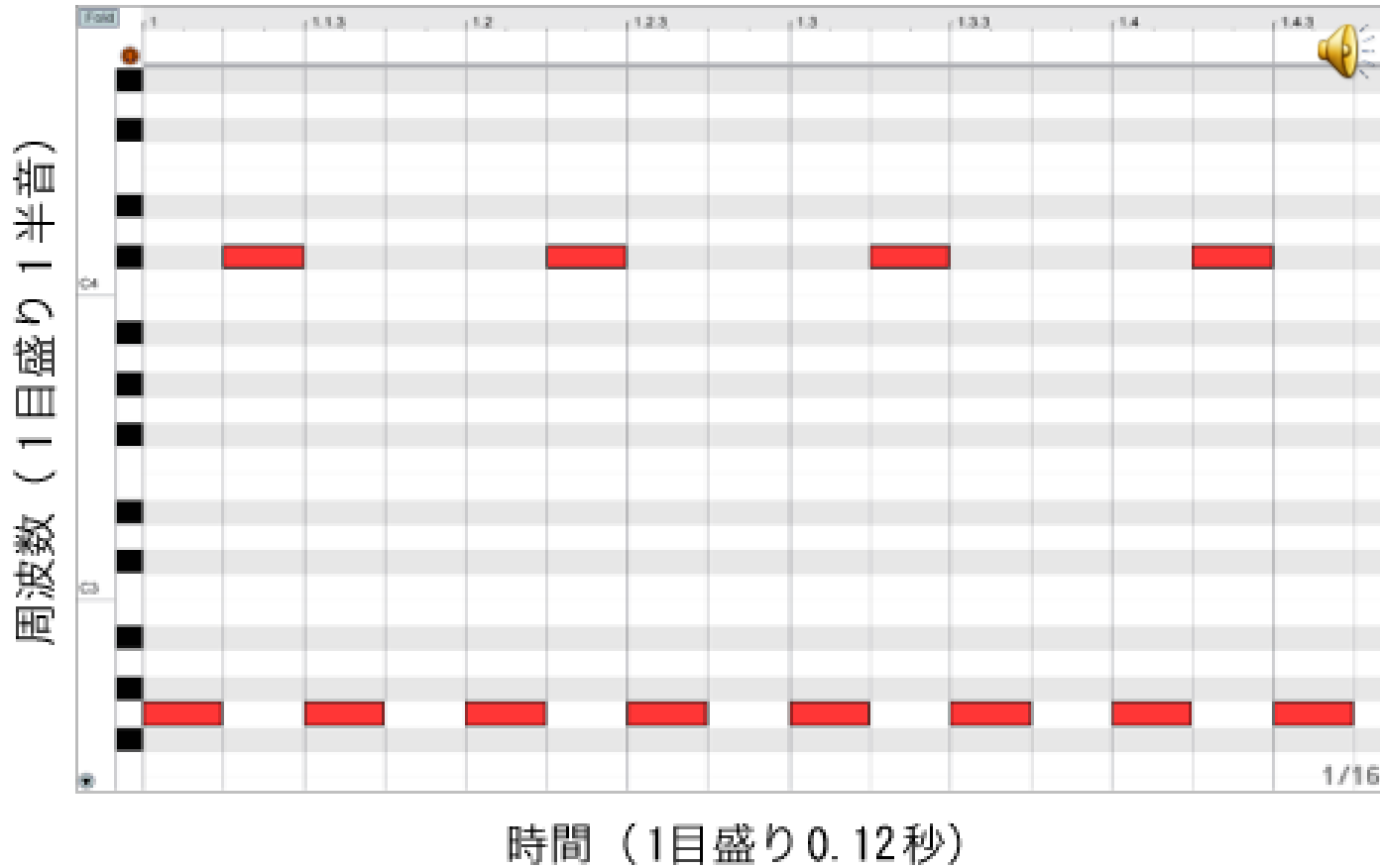
Coffee Break: 音脈分凝について

周波数差：1.5半音



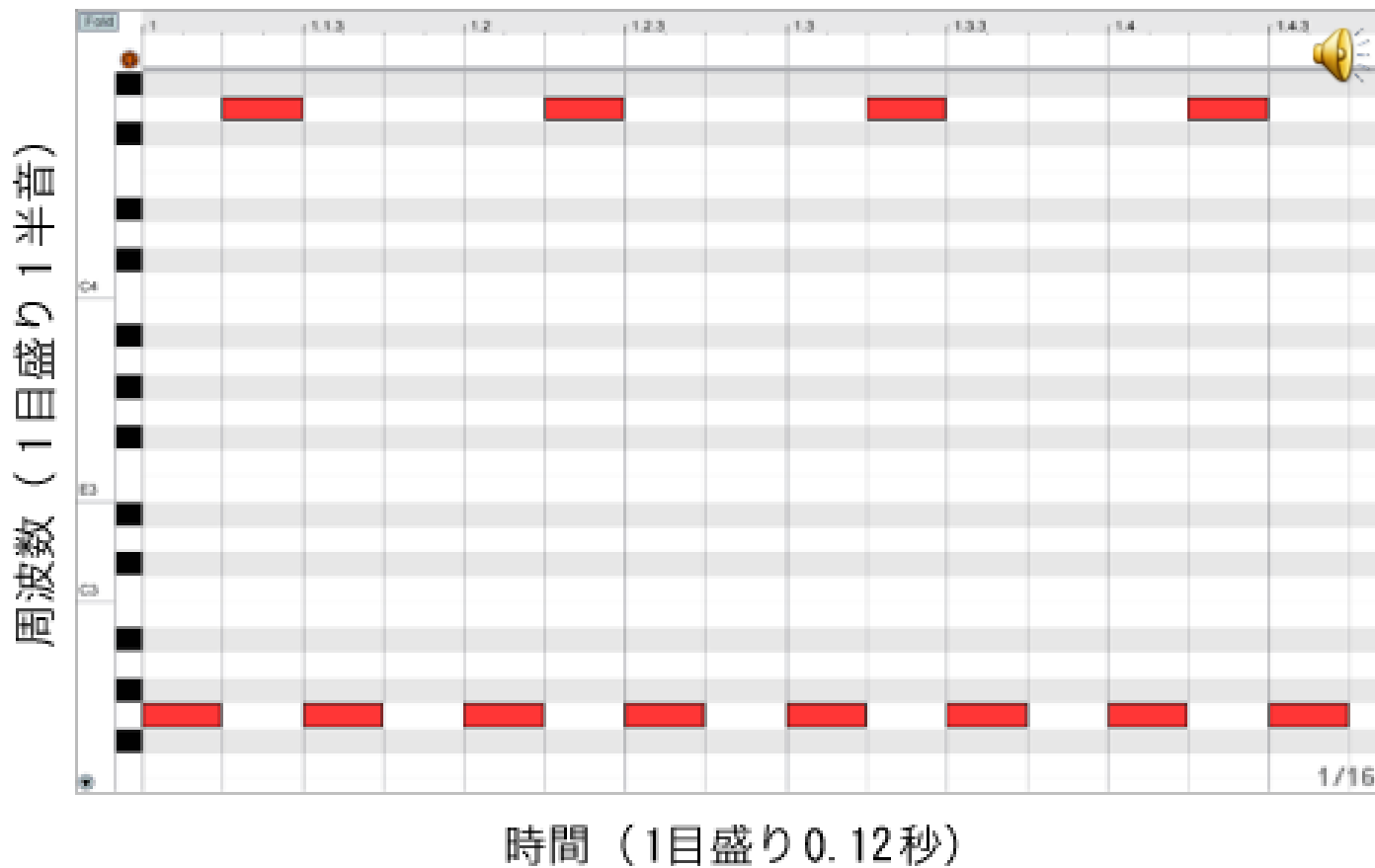
Coffee Break: 音脈分凝について

周波数差：1.8半音



Coffee Break: 音脈分凝について

周波数差：24半音



Coffee Break: 音脈分凝について

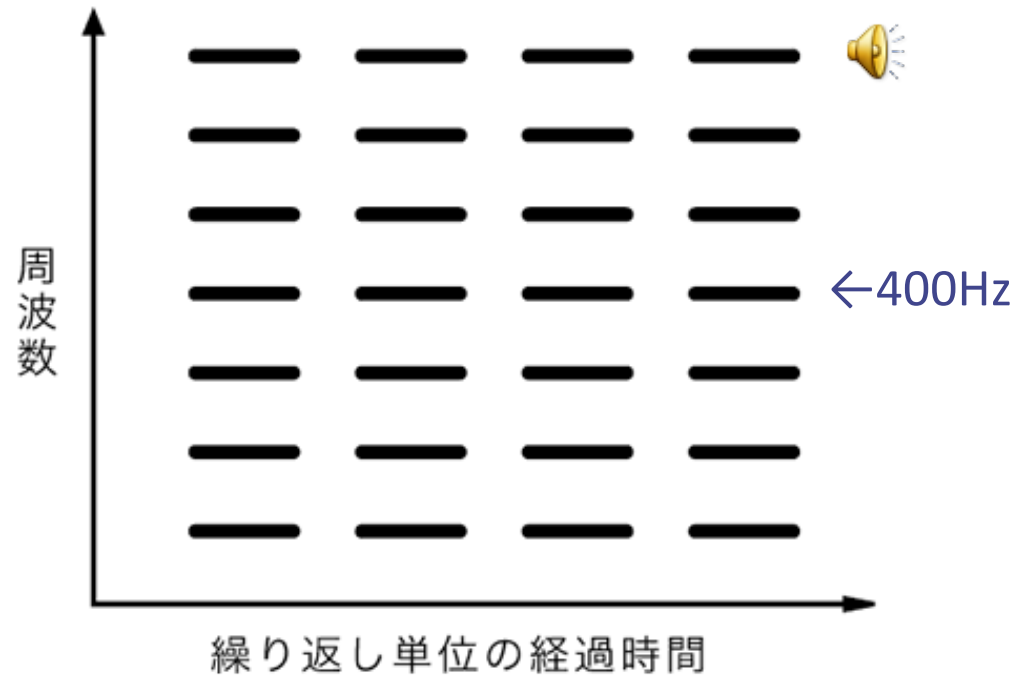
テンポ 遅い

速い

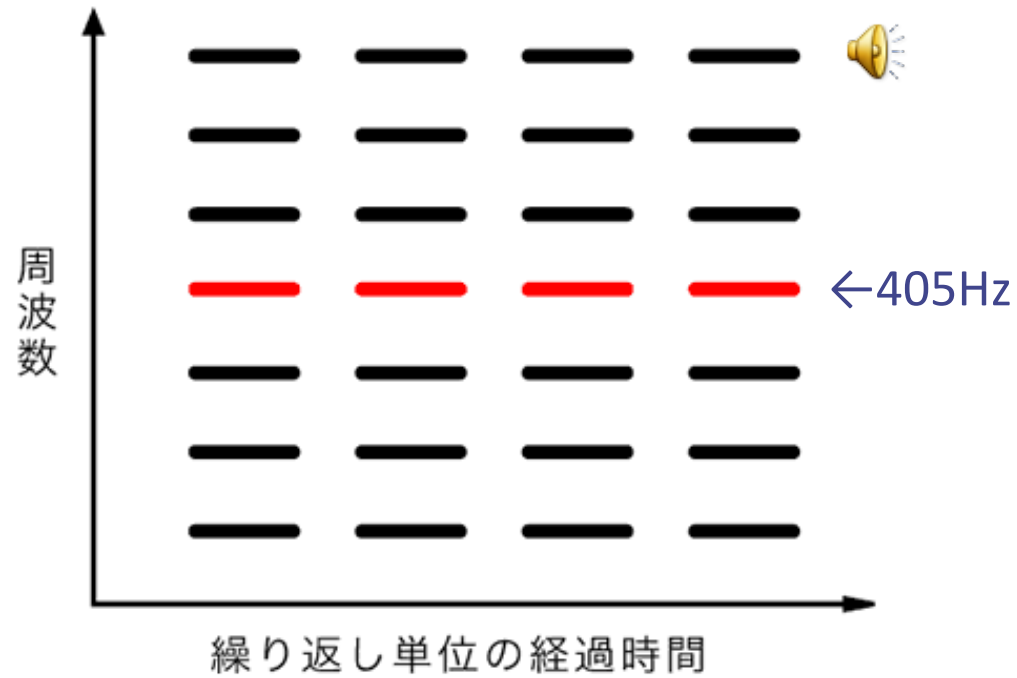


バスクリン作曲『ドニゼッティ『ポリウート』の主題による幻想曲』より

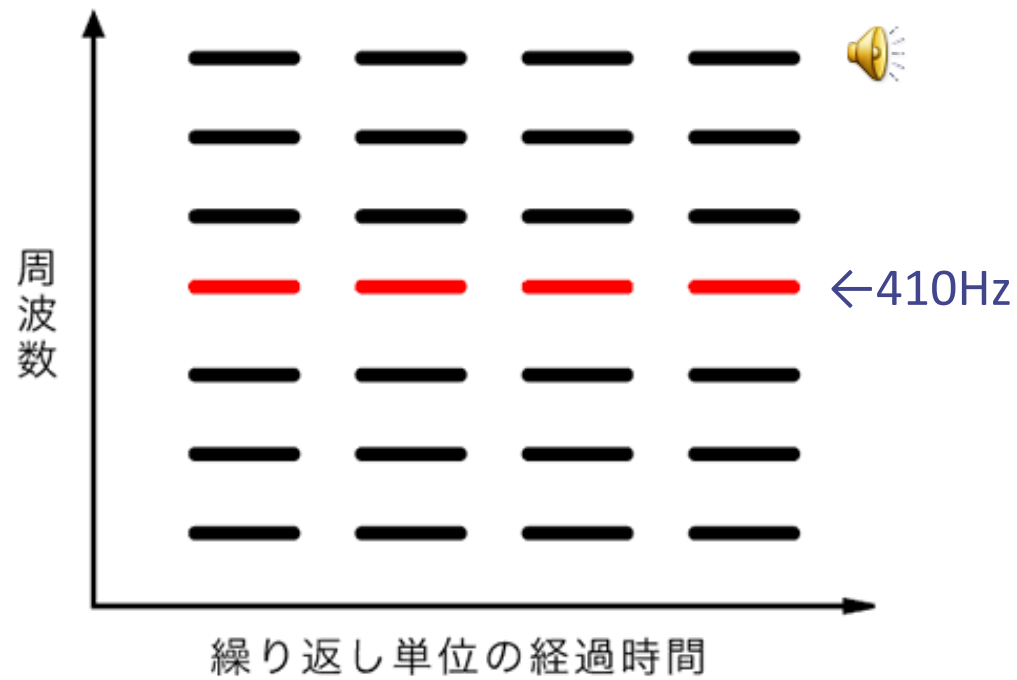
Coffee Break: 音脈分凝について



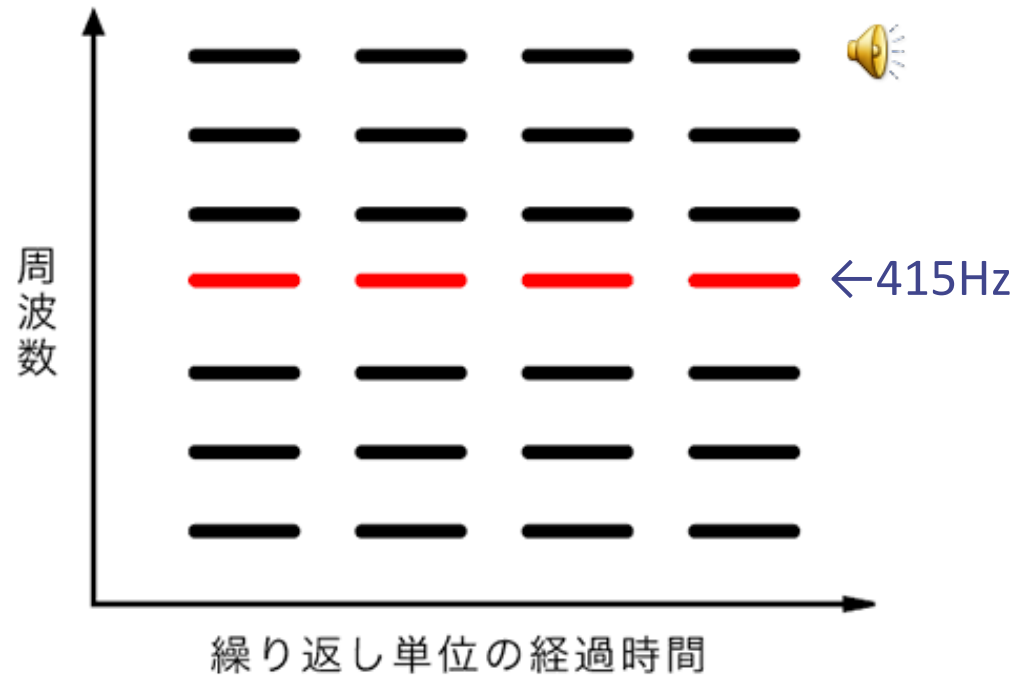
Coffee Break: 音脈分凝について



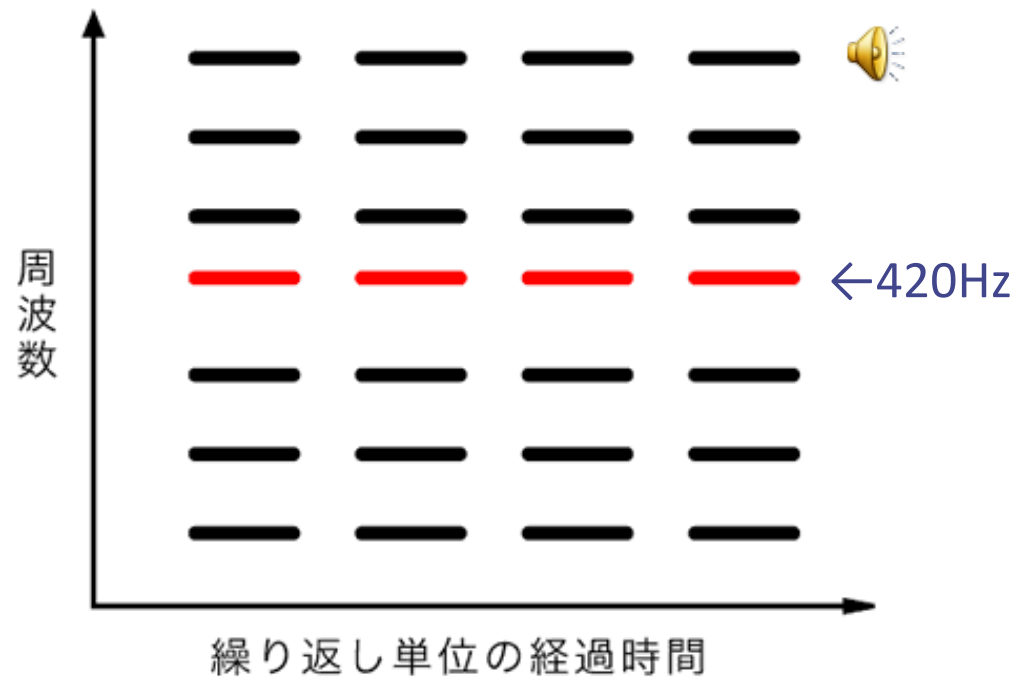
Coffee Break: 音脈分凝について



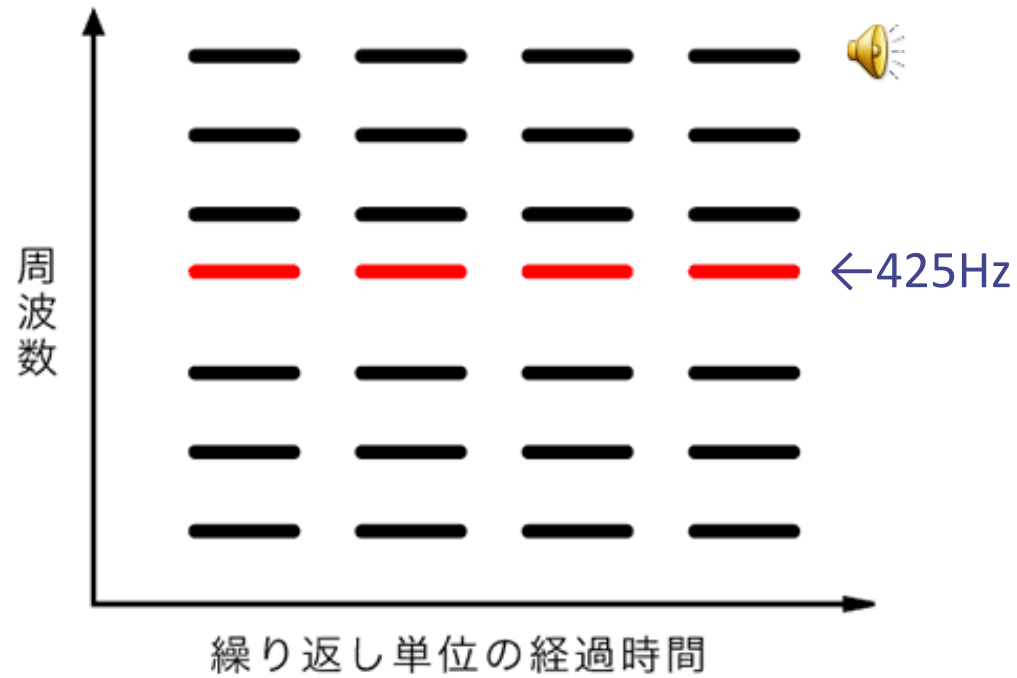
Coffee Break: 音脈分凝について



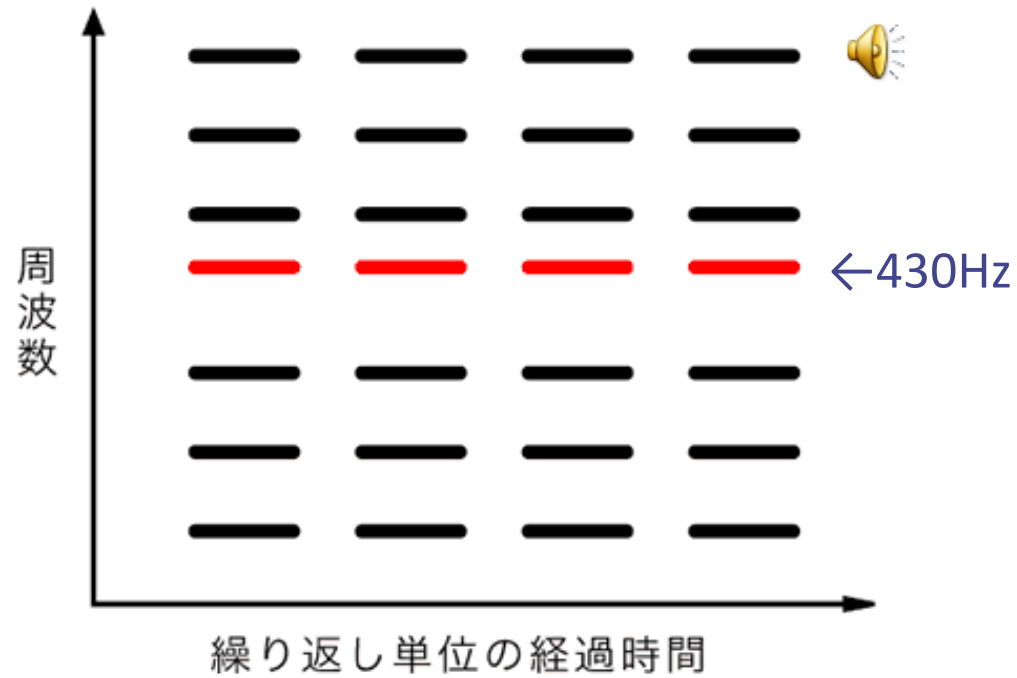
Coffee Break: 音脈分凝について



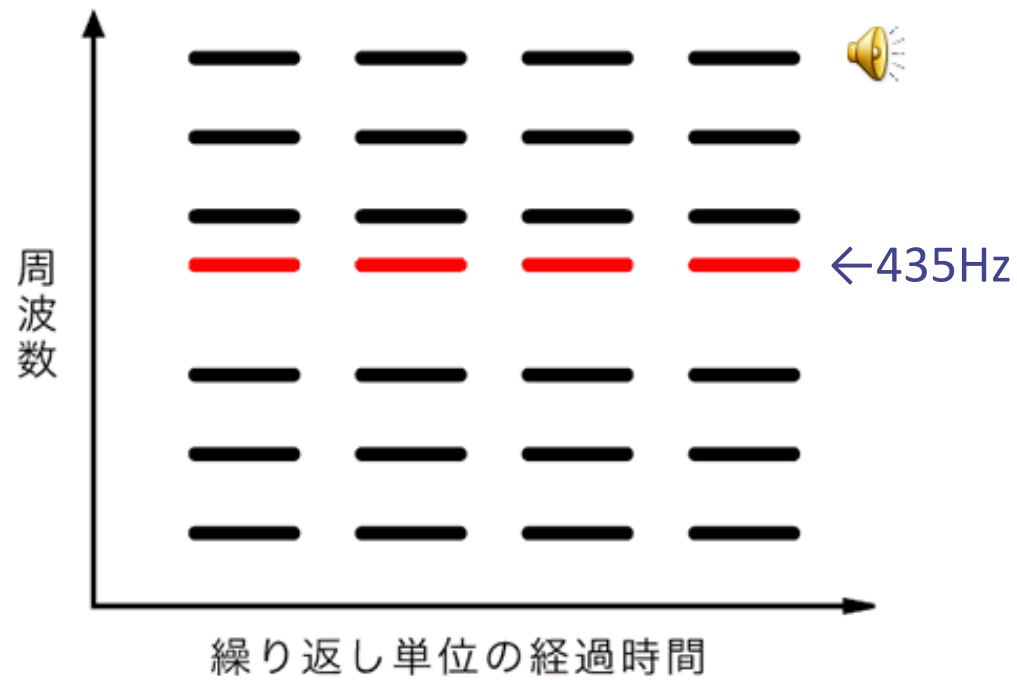
Coffee Break: 音脈分凝について



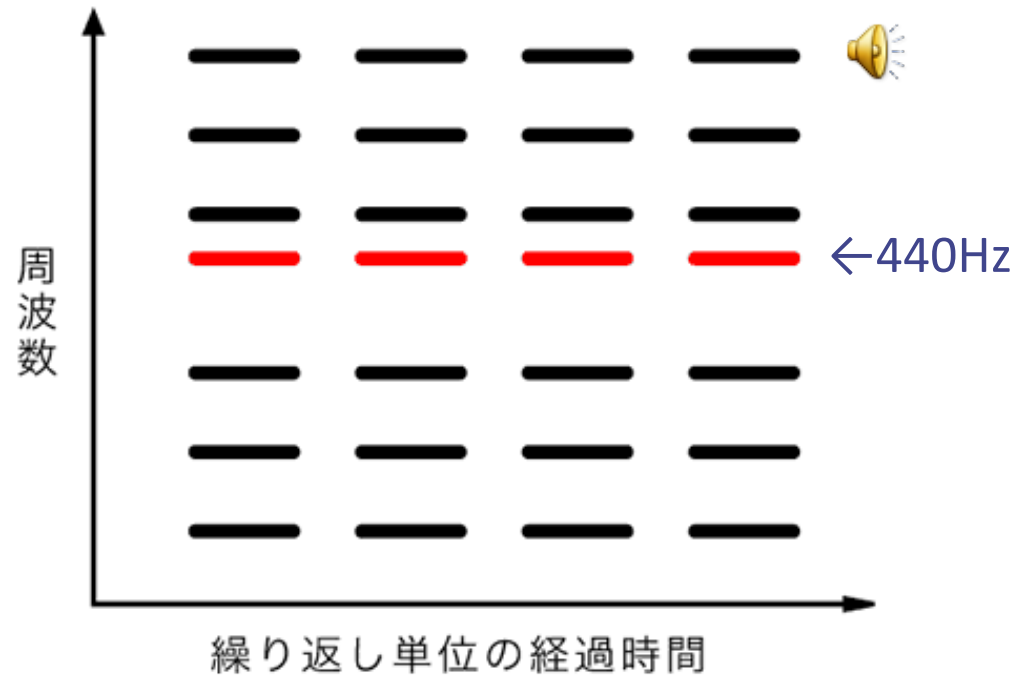
Coffee Break: 音脈分凝について



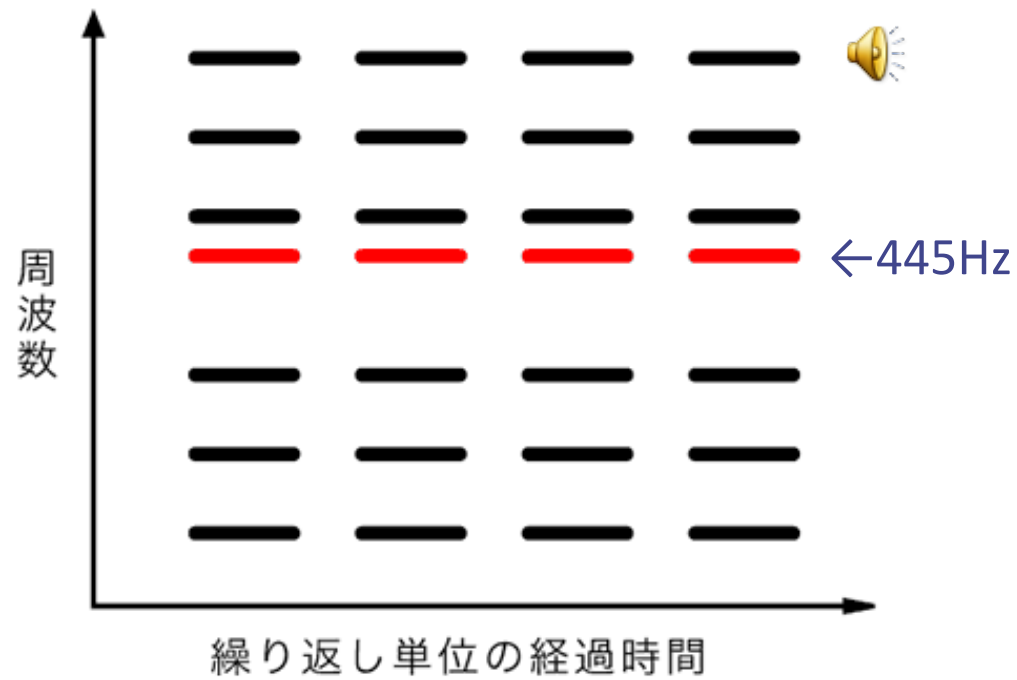
Coffee Break: 音脈分凝について



Coffee Break: 音脈分凝について



Coffee Break: 音脈分凝について



最尤推定法

- 分離行列 \mathbf{W} を直接推定
- 観測信号 $X = \{\mathbf{x}(t)\}_{1 \leq t \leq T}$ に対する \mathbf{W} の尤度関数

$$\mathcal{L}(\mathbf{W}) = \prod_{t=1}^T p(\mathbf{x}(t) | \mathbf{W}) \rightarrow \text{maximize}$$

- 線形変換と確率密度関数

$$\mathbf{y}(t) = \mathbf{W} \mathbf{x}(t) \longrightarrow \underline{p(\mathbf{y}(t))} = \frac{1}{|\det \mathbf{W}|} p(\mathbf{x}(t))$$

- 音源信号の独立性と非Gauss性を仮定

$$p(\mathbf{y}(t)) = \prod_{i=1}^I p(y_i(t)) \quad p(y_i(t)) : \text{Laplace分布など}$$

 $\log \mathcal{L}(\mathbf{W}) = T \log |\det \mathbf{W}| + \sum_{t=1}^T \sum_{i=1}^I \log p(y_i(t))$

最尤推定法

- 目的関数:

$$\mathcal{J}(\mathbf{W}) = \log |\det \mathbf{W}| + \frac{1}{T} \sum_t \sum_i \log p(y_i(t))$$

- 勾配法により \mathbf{W} を反復的に更新

$$\mathbf{W} \leftarrow \mathbf{W} + \eta \frac{\partial \mathcal{J}(\mathbf{W})}{\partial \mathbf{W}} \quad (\eta \text{ はステップサイズ})$$

$$\frac{\partial \mathcal{J}(\mathbf{W})}{\partial \mathbf{W}} = (\mathbf{W}^T)^{-1} - \frac{1}{T} \sum_t \Phi(\mathbf{y}(t)) \mathbf{x}(t)^T$$

$$\Phi(\mathbf{y}(t)) = \begin{bmatrix} \phi(y_1(t)) \\ \vdots \\ \phi(y_I(t)) \end{bmatrix}$$

$$\phi(y) = -\frac{\partial \log p(y)}{\partial y}$$

$\phi(y)$ の具体形

Laplace分布の場合: $\phi(y) = \text{sign}(y)$

先の近似分布の場合: $\phi(y) = \frac{y}{\sqrt{y^2 + \alpha}}$

自然勾配(Natural Gradient)法

- 勾配法では \mathbf{W} の逆行列計算が厄介

$$\frac{\partial \mathcal{J}(\mathbf{W})}{\partial \mathbf{W}} = (\mathbf{W}^T)^{-1} - \frac{1}{T} \sum_t \Phi(\mathbf{y}(t)) \mathbf{x}(t)^T$$

- 自然勾配(Natural Gradient)

$$\frac{\partial \mathcal{J}(\mathbf{W})}{\partial \mathbf{W}} \mathbf{W}^T \mathbf{W} = \left(\mathbf{I} - \frac{1}{T} \sum_t \Phi(\mathbf{y}(t)) \mathbf{x}(t)^T \right) \mathbf{W}$$

- 逆行列計算が不要
- Equivariance Property
混合行列 \mathbf{H} の影響 (特異点に近くて不安定, etc.) を受けない
- アルゴリズムは以下のとおり

$$\mathbf{y}(t) \leftarrow \mathbf{W} \mathbf{x}(t)$$

$$\mathbf{W} \leftarrow \mathbf{W} + \eta \left(\mathbf{I} - \frac{1}{T} \sum_t \Phi(\mathbf{y}(t)) \mathbf{y}^T \right) \mathbf{W}$$

自然勾配の導出 (1/2)

■ 最急降下方向

- W を $W + \epsilon dW$ に変化させたときに最も $\mathcal{J}(W + \epsilon dW) - \mathcal{J}(W)$ を小さくする dW が最急降下方向
- モデルの接空間がユークリッド的でない場合は「偏微分方向 \neq 最急降下方向」

■ dW のノルムの定義

- W における変化を単位行列 I に移し変えて

$$W + \epsilon dW = (I + \epsilon dW W^{-1})W$$

と変形し, 実質的な変化分を $dW W^{-1}$ と捉える

- $dW W^{-1}$ のFrobeniusノルム

$$\|dW W^{-1}\|_F^2 = \text{tr}(W^{-T} dW^T dW W^{-1}) = \text{tr}(W^{-1} W^{-T} dW^T dW)$$

を dW のノルムと定義

自然勾配の導出 (2/2)

■ 最急降下方向の導出

- $\text{tr}(\mathbf{W}^{-1}\mathbf{W}^{-\top}\text{d}\mathbf{W}^{\top}\text{d}\mathbf{W}) = 1$ という制約の下で $\mathcal{J}(\mathbf{W} + \epsilon\text{d}\mathbf{W}) - \mathcal{J}(\mathbf{W}) \simeq \epsilon\text{tr}(\nabla\mathcal{J}(\mathbf{W})^{\top}\text{d}\mathbf{W})$ を最小化する $\text{d}\mathbf{W}$ を求める最適化問題として定式化

- Lagrangeの未定乗数法

$$\frac{\partial}{\partial(\text{d}\mathbf{W})} \{ \epsilon(\nabla\mathcal{J}(\mathbf{W})^{\top}\text{d}\mathbf{W}) - \lambda(1 - \text{tr}(\mathbf{W}^{-1}\mathbf{W}^{-\top}\text{d}\mathbf{W}^{\top}\text{d}\mathbf{W})) \} = 0$$

→ Lagrangeの未定乗数

$$\Rightarrow \epsilon\nabla\mathcal{J}(\mathbf{W})^{\top} + 2\lambda\mathbf{W}^{-1}\mathbf{W}^{-\top}\text{d}\mathbf{W}^{\top} = 0$$

$$\Rightarrow \text{d}\mathbf{W} \propto -\nabla\mathcal{J}(\mathbf{W})\mathbf{W}^{\top}\mathbf{W}$$

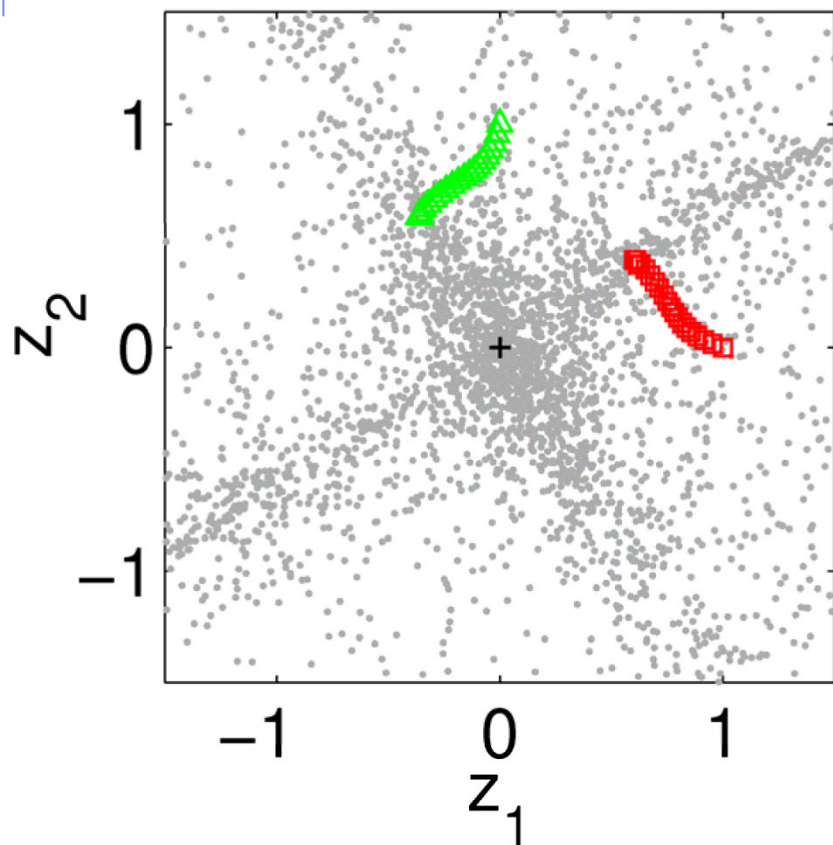
→ 逆行列をキャンセルしてくれる

\mathbf{W} の変化を単位行列に移し変えて考えているため、
 \mathbf{W} に依らず一定の収束特性を示す
 \Rightarrow 等価性 (Equivalence)

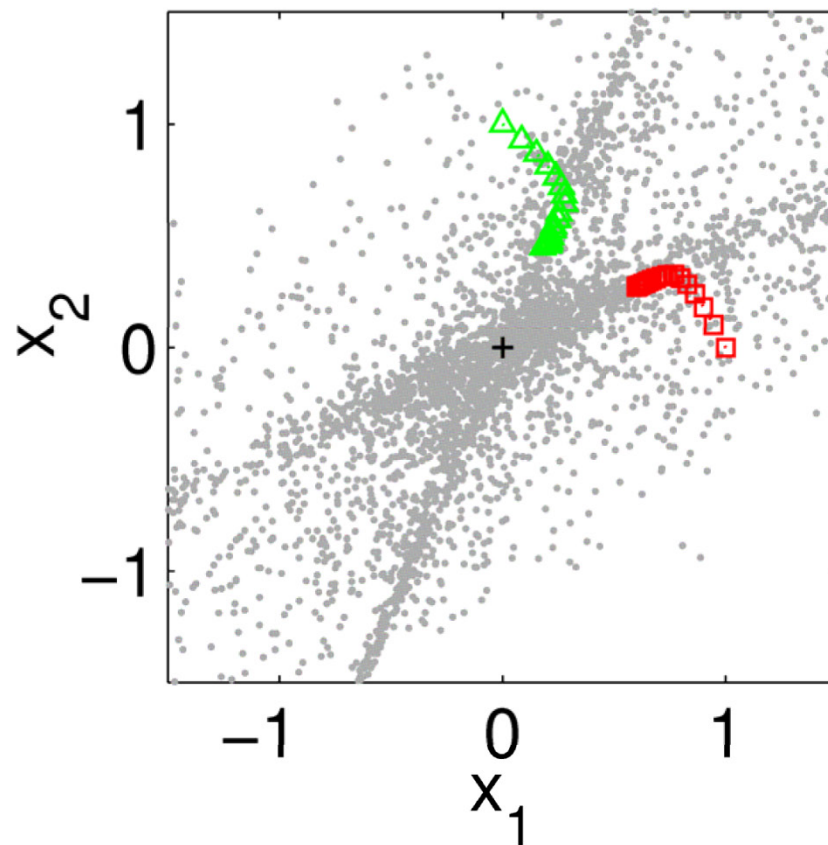
自然勾配法による推定の様子

- 赤(□): $w_1 = [1 \ 0]^T$ を初期値に設定
- 緑(△): $w_2 = [0 \ 1]^T$ を初期値に設定
- 観測信号の白色化をせずとも良好な解に収束

Whitened mixtures



Original mixtures



目次

- 独立成分分析について
 - 定式化, 歴史, 応用
- 信号の統計的性質
 - 信号を混ぜるー 中心極限定理
- 独立成分分析のアルゴリズム
 - 白色化 + FastICA
 - 最尤推定法 by Natural Gradient

まとめ

■ 中心極限定理

- 音をたくさん混ぜていくと振幅値の分布は正規分布に近づく

■ 独立成分分析

- 独立 = エントロピー減 = 正規分布から遠ざかる
- 非Gauss性の分布の例: Laplace分布
- 効率的なアルゴリズムの紹介
 - ◆ FastICA, 自然勾配法

レポート課題の対象論文

- S. Amari, A. Cichocki and H.H. Yang, “A new learning algorithm for blind signal separation,” In Advances in Neural Information Processing Systems (NIPS), vol. 8, pp. 757-763 (1996).