

非負値行列因子分解入門 ～音響信号処理を題材として～

亀岡弘和

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
kameoka@cs.brl.ntt.co.jp

部分空間法研究会 2010年7月26日

目次

1. 非負値行列因子分解(NMF)とは

- 何に使えるのか(音響信号処理を題材として)
- どのような性質があるのか
- どのように求めるのか
- 統計モデルとしての解釈とアルゴリズムの一般化
- 他の技術とどう関係しているのか

2. 複素NMF [H. Kameoka et al., 2008]

- NMFにおける問題点
- 解決のためのアイデア
- NMFとどう関係しているのか

非負値行列因子分解(NMF)とは

- 非負値行列を2つの非負値行列の積で表現

$$Y \simeq H U$$

- 行列因子分解の応用場面

- ブラインド信号分離(Blind Signal Separation)

H : 混合行列 U の各行: 源信号

- クラスタリング

H の各列: 基底ベクトル U の各列: 結合係数

なぜ非負値なのか？その意図は？

- データ行列の非負性

- 実世界には非負値データが多い
(例) パワースペクトル, 画素値, 度数,

- 基底行列の非負性

- 「非負値データの構成要素もまた非負値データであるべき(でないと物理的に意味をなさない！)」

という考え方

(例) 負の値をもったパワースペクトルなんて解釈のしようがない

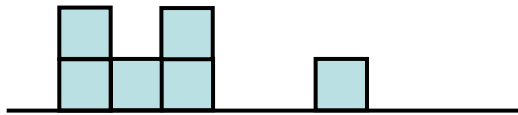
- 係数行列の非負性

- 構成要素の混ざり方は「足し算」のみ
- 係数行列をスパースに誘導 → 基底の情報量をアップ

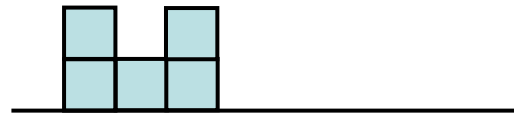
係数行列の非負性について

- 係数が負の値を取っても良い場合

観測データ



基底ベクトル



係数

$\times 1$



$\times 1$



$\times (-2)$

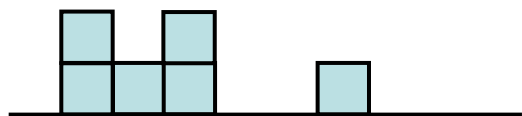
+)



係数行列の非負性について

- 係数が負の値を取ってはならない場合

観測データ



基底ベクトル



+)



係数

$\times 1$

$\times 0$

$\times 0$

スパース

PCAとNMFの比較

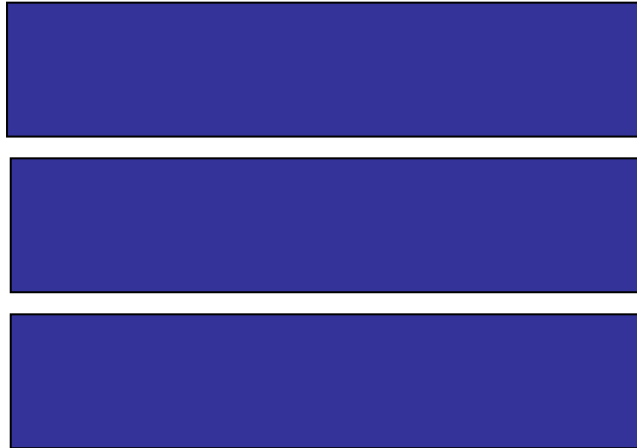
HH
SN
BD



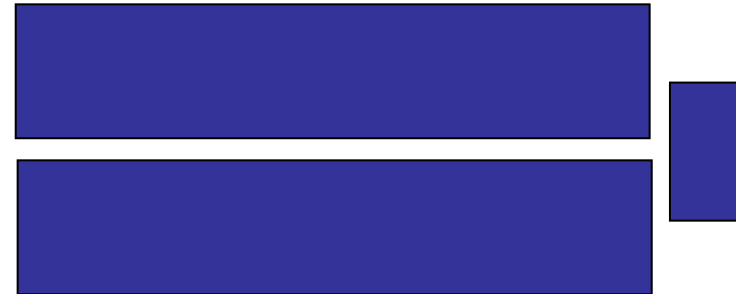
観測データの平均値ベクトル
(ハイハット/スネアドラム/バスドラムの音量パターン)

→時刻

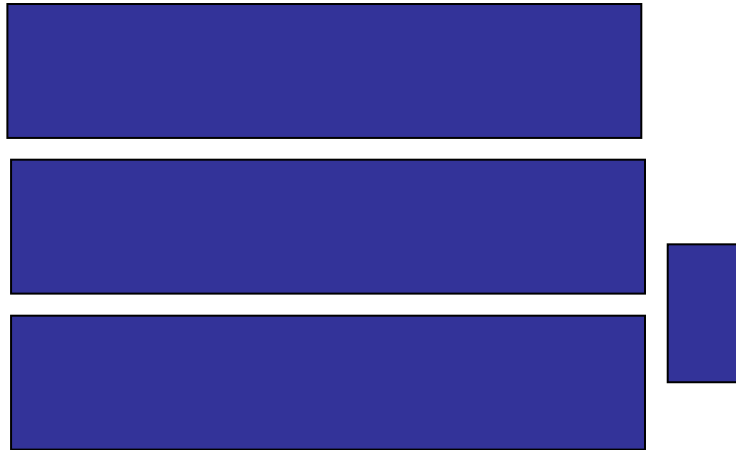
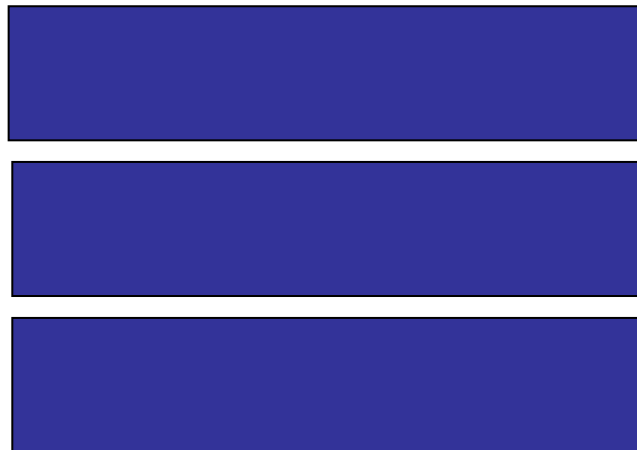
PCA
基底



各図が1つのベクトルに対応



NMF
基底



各図は[D.P.W Ellis & J. Arroyo, 2004]より抜粋

NMFが生まれた背景

- 画像処理分野で生まれた技術 [D.D. Lee & H.S. Seung, 1999]
 - 顔画像から顔パーツを抽出するのが目的
 - 概念自体は90年代前半に登場 [P. Paatero & U. Tapper, 1994]
- 音のスペクトルを画像と見なして適用(後述)
→ 音声分離・自動採譜等
[P. Smaragdis et al., 2003]
以降極めて多数...
- 効率的な反復アルゴリズム
[D.D. Lee & H.S. Seung, 2000]

$$H_{k,i} \leftarrow H_{k,i} \frac{[YU^T]_{k,i}}{[HUU^T]_{k,i}}$$

$$U_{i,j} \leftarrow U_{i,j} \frac{[H^T Y]_{i,j}}{[H^T H U]_{i,j}}$$

[D.D. Lee & H.S. Seung, 1999]より抜粋

NMFで音声スペクトログラムを分解してみる

- 音響信号 $y(u) \in \mathbb{R}$

↓ 短時間フーリエ変換
(時間周波数分解)

$$Y_{\omega,t} = \int W(u)y(u+t)e^{j\omega u}du \in \mathbb{C}$$

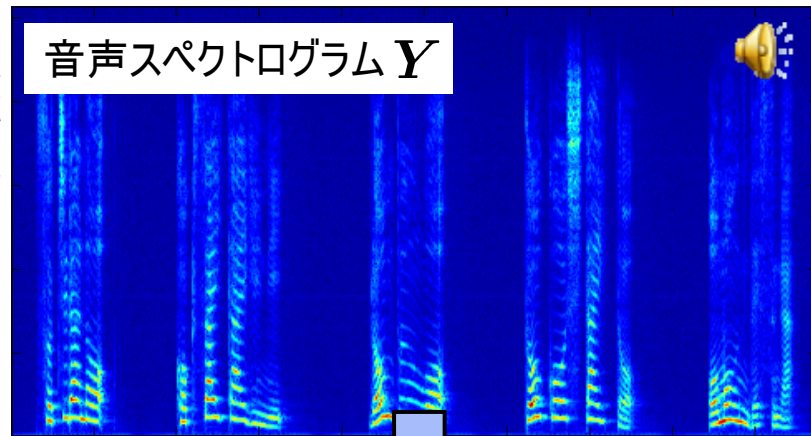
↓ 絶対値をとる

$|Y_{\omega,t}|$: 時刻 t に周波数 ω の成分が
どれほど含まれているか

$$Y \simeq HU \quad (H \succeq 0, U \succeq 0)$$

↑ 周波数

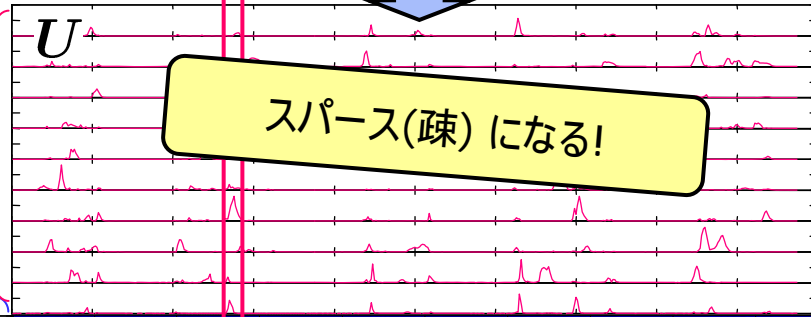
音声スペクトログラム Y



各基底の
アクティビティ

U

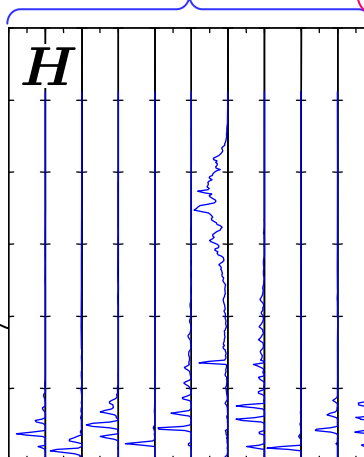
スパース(疎)になる!



基底スペクトル

↑ 周波数

H



低ランクスペクトログラム HU

基底数10

基底数30



繰り返し現れるスペクトル
パターンが獲得される

時刻 →

何に使えるのか？ (1/3)

- 自動採譜 [P. Smaragdis et al., 2003]

J.S. Bach: Fuge #16 in G minor

U

[P. Smaragdis et al., 2003]より抜粋

何に使えるのか？(2/3)

- 教師ありモノラル音源分離 [P. Smaragdis et al., 2007]

H_{speech}

音声(学習データ)

H_{chime}

チャイム(学習データ)

各図は[P. Smaragdis et al., 2007]より抜粋

Y

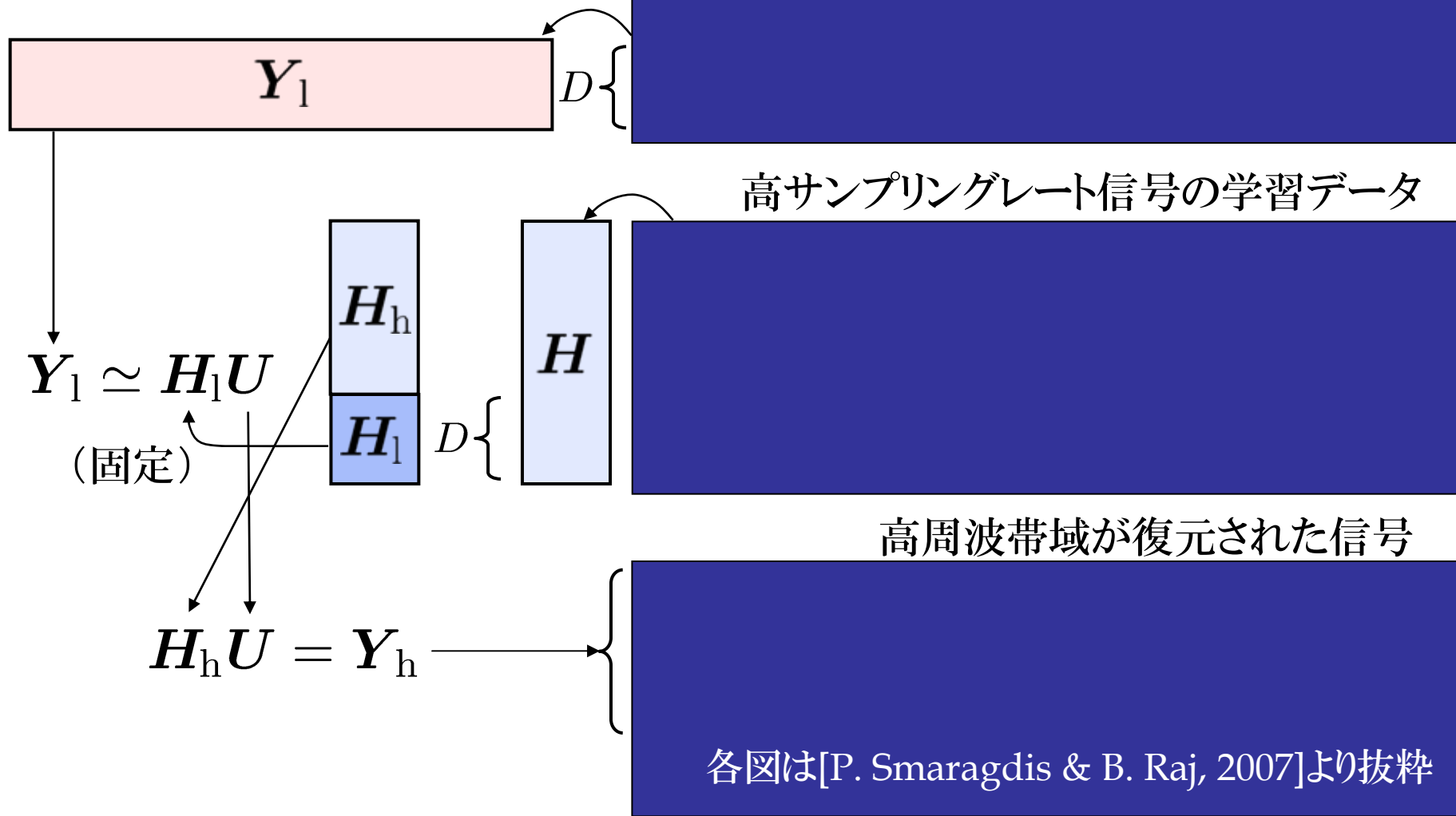
混合信号(テストデータ)

$$Y \simeq \underbrace{\begin{bmatrix} H_{\text{speech}} & H_{\text{chime}} \end{bmatrix}}_{\text{固定}} \begin{bmatrix} U_{\text{speech}} \\ U_{\text{chime}} \end{bmatrix}$$

何に使えるのか？(3/3)

• 音の「超解像」

[P. Smaragdis & B. Raj, 2007]



NMFの基本問題

- T 個の観測データ(非負値ベクトル) y_1, \dots, y_T
- I 個の非負値基底ベクトル h_1, \dots, h_I の非負結合でどの観測データも良く表現できる基底セットを求めたい

$$y_t \simeq \sum_{i=1}^I h_i U_{i,t}, \quad U_{i,t} \geq 0 \quad (t = 1, \dots, T)$$

➔ $I \leq \Omega$ のとき, 各基底ベクトルはいくつかの観測データで共起するパターン(パーツと呼ぶ)を表現するよう誘導される

$$\left. \begin{array}{l} Y := (y_1, \dots, y_T) \\ H := (h_1, \dots, h_I) \\ U := (U_{i,t})_{I \times T} \end{array} \right\} \quad Y \simeq HU \quad H_{\omega,i} \geq 0, U_{i,t} \geq 0$$

minimize $D(Y, HU)$

subject to $H_{\omega,i} \geq 0, U_{i,t} \geq 0$

• 想定する状況

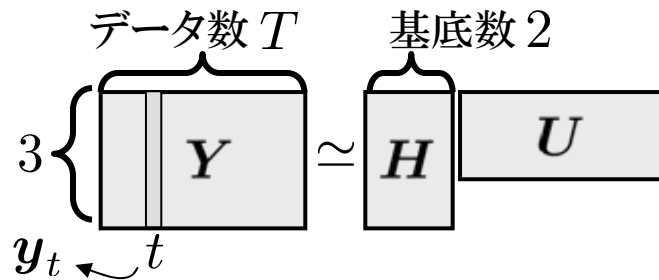
- データは非負値
- 高々 I 種類のパーツだけですべての観測データは構成されている
- パーツの和で観測データをモデル化

どんな性質があるのか？

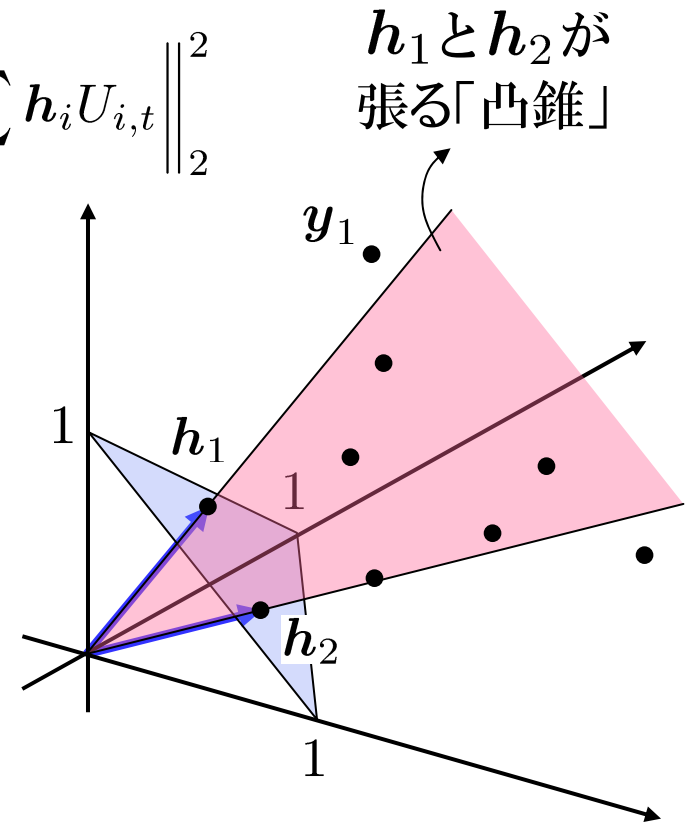
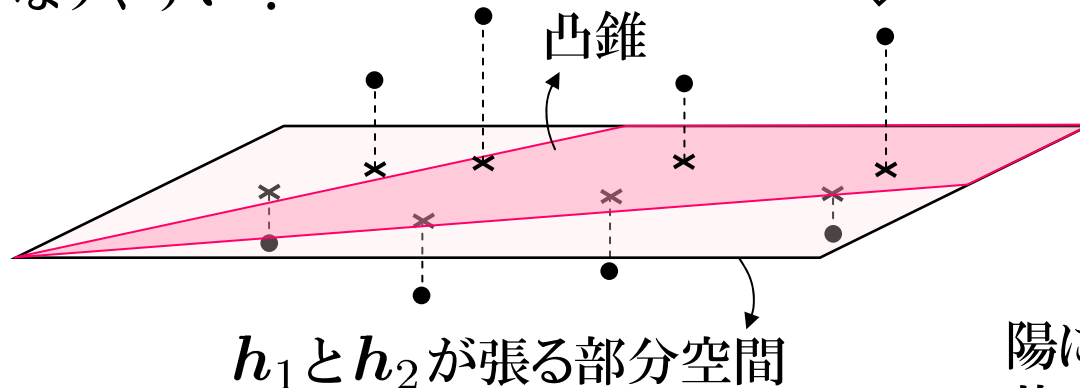
- Frobeniusノルム規準のNMF

$$\text{minimize } \|Y - HU\|_F^2 = \sum_t \left\| y_t - \sum_i h_i U_{i,t} \right\|_2^2$$

$$\text{subject to } H_{\omega,i} \geq 0, U_{i,t} \geq 0$$



斜交基底だから係数はスパースになりやすい？



陽に制約を入れていないのに
基底が直交化される傾向に？

どうやって求めるのか

- NMFにおける代表的な最適化規準
 - Frobeniusノルム
 - Iダイバージェンス(一般化KLダイバージェンス)
- 押さえておくべき基本原理
 - 補助関数法
 - 凸不等式(Jensenの不等式)

NMFにおける代表的な最適化規準

- Frobeniusノルム規準

$$\begin{aligned} \text{minimize} \quad & \|Y - HU\|_F^2 = \sum_{\omega,t} \left| Y_{\omega,t} - \sum_i H_{\omega,i} U_{i,t} \right|^2 \\ \text{subject to} \quad & H_{\omega,i} \geq 0, U_{i,t} \geq 0 \end{aligned}$$

なんとかしたい部分

- Iダイバージェンス(一般化KLダイバージェンス)

$$\begin{aligned} \text{minimize} \quad & \sum_{\omega,t} \left(Y_{\omega,t} \log \frac{Y_{\omega,t}}{\sum_i H_{\omega,i} U_{i,t}} - Y_{\omega,t} + \sum_i H_{\omega,i} U_{i,t} \right) \\ \text{subject to} \quad & H_{\omega,i} \geq 0, U_{i,t} \geq 0 \end{aligned}$$

なんとかしたい部分

いずれも $Y = HU$ のとき 0 になる

押さえておくべき基本原理(1/2)

• 補助関数法 [J. de Leeuw, 1994]

- $D(\theta) = \min_{\bar{\theta}} G(\theta, \bar{\theta})$ を満たす $G(\theta, \bar{\theta})$ を補助関数と定義

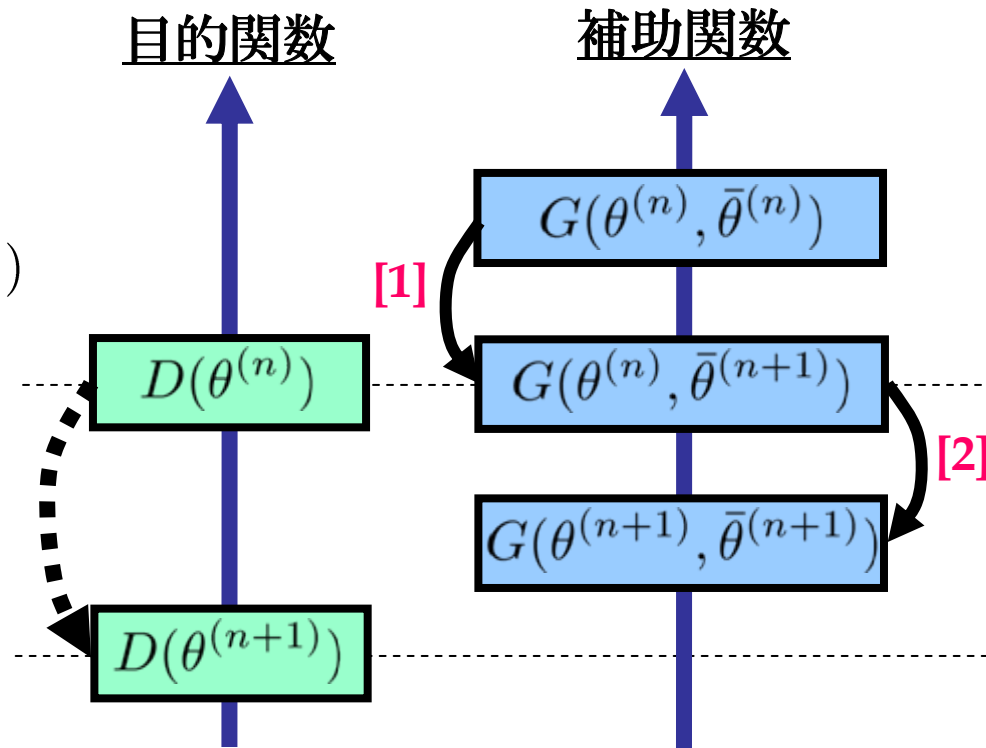
- 反復アルゴリズム

[1] $\bar{\theta}^{(n+1)} = \operatorname{argmin}_{\bar{\theta}} G(\theta^{(n)}, \bar{\theta})$

[2] $\theta^{(n+1)} = \operatorname{argmin}_{\theta} G(\theta, \bar{\theta}^{(n+1)})$

- 収束性

$$\begin{aligned} D(\theta^{(n)}) &= G(\theta^{(n)}, \bar{\theta}^{(n+1)}) \\ &\geq G(\theta^{(n+1)}, \bar{\theta}^{(n+1)}) \\ &\geq D(\theta^{(n+1)}) \end{aligned}$$

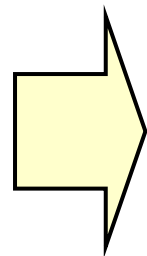


目的関数を直接最小化するのが難しいなら、
とりあえずその上限関数を作ってみよう！

押さえておくべき基本原理(2/2)

• Jensenの不等式

- $\phi(\cdot)$: 凸関数
- $\lambda_i \geq 0, \sum_i \lambda_i = 1$

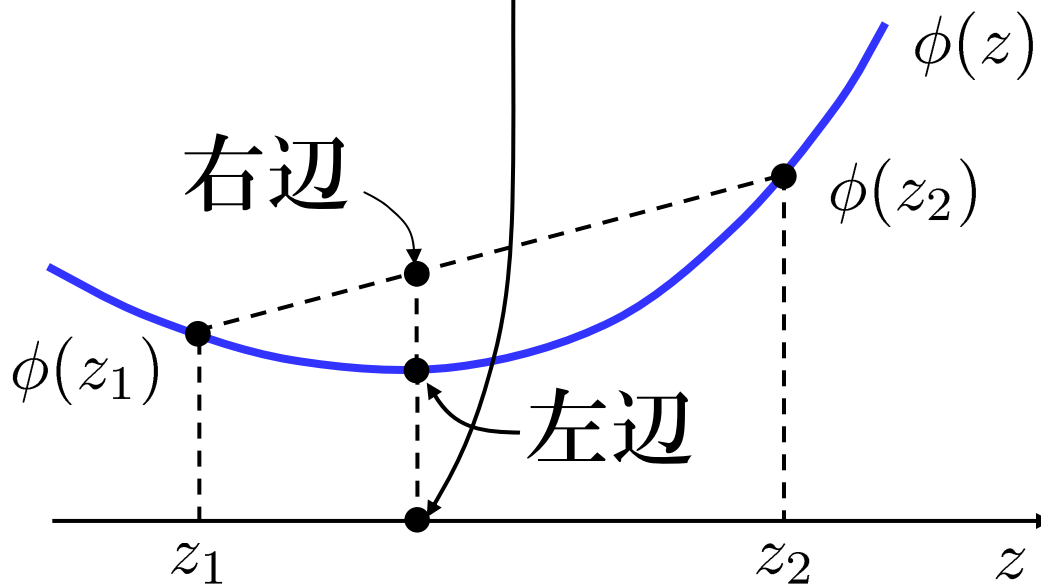


$$\phi\left(\sum_i \lambda_i z_i\right) \leq \sum_i \lambda_i \phi(z_i)$$

例えば、

$\phi(z) = \log z$ の場合:

$$\begin{aligned} & \log\left(\sum_i \lambda_i z_i\right) \\ & \leq \sum_i \lambda_i \log z_i \end{aligned}$$



Frobeniusノルム規準のNMFアルゴリズム

- 目的関数

$$\begin{aligned}\|Y - HU\|_F^2 &= \sum_{\omega,t} \left| Y_{\omega,t} - \sum_i H_{\omega,i} U_{i,t} \right|^2 \\ &= \sum_{\omega,t} \left(|Y_{\omega,t}|^2 - 2Y_{\omega,t} \sum_i H_{\omega,i} U_{i,t} + \underbrace{\left| \sum_i H_{\omega,i} U_{i,t} \right|^2}_{\text{Jensen's inequality}} \right)\end{aligned}$$

- 下線部に対してJensenの不等式を立ててみる

$$\begin{aligned}\left(\sum_i z_i \right)^2 &= \left(\sum_i \lambda_i \frac{z_i}{\lambda_i} \right)^2 \leq \sum_i \lambda_i \left(\frac{z_i}{\lambda_i} \right)^2 = \sum_i \frac{z_i^2}{\lambda_i} \\ &\Rightarrow \left| \sum_i H_{\omega,i} U_{i,t} \right|^2 \leq \sum_i \frac{H_{\omega,i}^2 U_{i,t}^2}{\lambda_{i,\omega,t}}\end{aligned}$$

- 補助関数が完成 $H_{\omega,i}$ または $U_{i,t}$ ごとの二次関数の和になっている

$$\|Y - HU\|_F^2 \leq \sum_{\omega,t} \left(|Y_{\omega,t}|^2 - 2Y_{\omega,t} \sum_i H_{\omega,i} U_{i,t} + \sum_i \frac{H_{\omega,i}^2 U_{i,t}^2}{\lambda_{i,\omega,t}} \right)$$

Frobeniusノルム規準のNMFアルゴリズム

- 補助関数が完成したらあとはステップ1とステップ2を導出すればOK!

$$G := \sum_{\omega,t} \left(|Y_{\omega,t}|^2 - 2Y_{\omega,t} \sum_i H_{\omega,i} U_{i,t} + \sum_i \frac{H_{\omega,i}^2 U_{i,t}^2}{\lambda_{i,\omega,t}} \right)$$

[1] $\lambda_{i,\omega,t} \leftarrow \operatorname{argmin}_{\lambda_{i,\omega,t}} G = \frac{H_{\omega,i} U_{i,t}}{\sum_{i'} H_{\omega,i'} U_{i',t}}$ 代入

[2] $H_{\omega,i} \leftarrow \operatorname{argmin}_{H_{\omega,i}} G = \frac{\sum_t Y_{\omega,t} U_{i,t}}{\sum_t \frac{U_{i,t}^2}{\lambda_{i,\omega,t}}}$

$U_{i,t} \leftarrow \operatorname{argmin}_{U_{i,t}} G = \frac{\sum_{\omega} Y_{\omega,t} H_{\omega,i}}{\sum_{\omega} \frac{H_{\omega,i}^2}{\lambda_{i,\omega,t}}}$

$$H_{\omega,i} \leftarrow H_{\omega,i} \frac{\sum_t Y_{\omega,t} U_{i,t}}{\sum_t U_{i,t} \sum_{i'} H_{\omega,i'} U_{i',t}}$$

$$U_{i,t} \leftarrow U_{i,t} \frac{\sum_{\omega} Y_{\omega,t} H_{\omega,i}}{\sum_{\omega} H_{\omega,i} \sum_{i'} H_{\omega,i'} U_{i',t}}$$

Iダイバージェンス規準のNMFアルゴリズム

- 目的関数

$$\begin{aligned} D &:= \sum_{\omega,t} \left(Y_{\omega,t} \log \frac{Y_{\omega,t}}{\sum_i H_{\omega,i} U_{i,t}} - Y_{\omega,t} + \sum_i H_{\omega,i} U_{i,t} \right) \\ &= \sum_{\omega,t} \left(Y_{\omega,t} \log Y_{\omega,t} - Y_{\omega,t} \log \underbrace{\sum_i H_{\omega,i} U_{i,t}}_{\text{下線部}} - Y_{\omega,t} + \sum_i H_{\omega,i} U_{i,t} \right) \end{aligned}$$

- 下線部に対してJensenの不等式を立ててみる

$$\log \left(\sum_i z_i \right) = \log \left(\sum_i \lambda_i \frac{z_i}{\lambda_i} \right) \leq \sum_i \lambda_i \log \left(\frac{z_i}{\lambda_i} \right)$$

$$\Rightarrow \log \sum_i H_{\omega,i} U_{i,t} \leq \sum_i \lambda_{i,\omega,t} \log \frac{H_{\omega,i} U_{i,t}}{\lambda_{i,\omega,t}}$$

- 補助関数が完成

$$D \leq \sum_{\omega,t} \left(Y_{\omega,t} \log Y_{\omega,t} - Y_{\omega,t} \sum_i \lambda_{i,\omega,t} \log \frac{H_{\omega,i} U_{i,t}}{\lambda_{i,\omega,t}} - Y_{\omega,t} + \sum_i H_{\omega,i} U_{i,t} \right)$$

Iダイバージェンス規準のNMFアルゴリズム

- 補助関数が完成したらあとはステップ1とステップ2を導出すればOK!

$$G := \sum_{\omega,t} \left(Y_{\omega,t} \log Y_{\omega,t} - Y_{\omega,t} \sum_i \lambda_{i,\omega,t} \log \frac{H_{\omega,i} U_{i,t}}{\lambda_{i,\omega,t}} - Y_{\omega,t} + \sum_i H_{\omega,i} U_{i,t} \right)$$

$$[1] \lambda_{i,\omega,t} \leftarrow \operatorname{argmin}_{\lambda_{i,\omega,t}} G = \frac{H_{\omega,i} U_{i,t}}{\sum_{i'} H_{\omega,i'} U_{i',t}} \quad \text{代入}$$

$$[2] H_{\omega,i} \leftarrow \operatorname{argmin}_{H_{\omega,i}} G = \frac{\sum_t Y_{\omega,t} \lambda_{i,\omega,t}}{\sum_t U_{i,t}}$$

$$U_{i,t} \leftarrow \operatorname{argmin}_{U_{i,t}} G = \frac{\sum_{\omega} Y_{\omega,t} \lambda_{i,\omega,t}}{\sum_{\omega} H_{\omega,i}}$$

$$H_{\omega,i} \leftarrow H_{\omega,i} \frac{\sum_t \frac{Y_{\omega,t} U_{i,t}}{\sum_{i'} H_{\omega,i'} U_{i',t}}}{\sum_t U_{i,t}}$$

$$U_{i,t} \leftarrow U_{i,t} \frac{\sum_{\omega} \frac{Y_{\omega,t} H_{\omega,i}}{\sum_{i'} H_{\omega,i'} U_{i',t}}}{\sum_{\omega} H_{\omega,i}}$$

統計モデルとしての解釈

- NMFは以下を仮定した最尤推定問題と等価

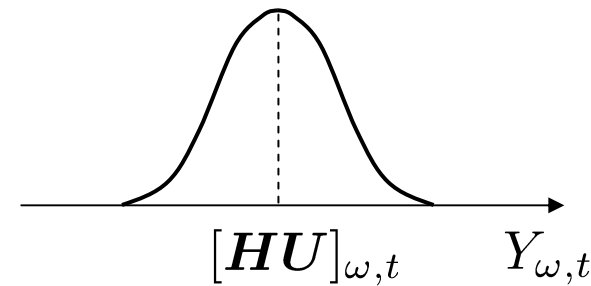
– Frobeniusノルム規準

$$Y_{\omega,t} \sim \mathcal{N}([HU]_{\omega,t}, \sigma^2)$$

$$P(Y|H, U) = \prod_{\omega,t} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2} (Y_{\omega,t} - [HU]_{\omega,t})^2}$$

$$\rightarrow -\log P(Y|HU) \stackrel{c}{=} \frac{1}{2\sigma^2} \sum_{\omega,t} (Y_{\omega,t} - [HU]_{\omega,t})^2$$

正規分布



– Iダイバージェンス規準

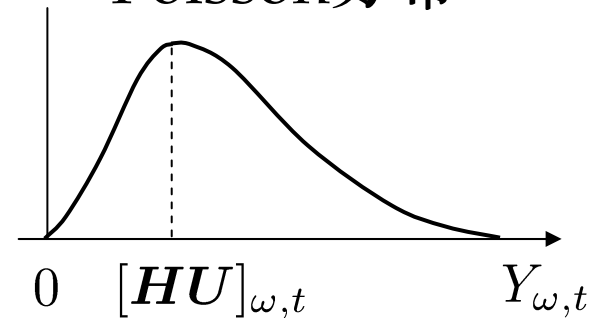
$$Y_{\omega,t} \sim \text{Pois}([HU]_{\omega,t})$$

$$P(Y|H, U) = \prod_{\omega,t} \frac{[HU]_{\omega,t}^{Y_{\omega,t}} e^{-[HU]_{\omega,t}}}{Y_{\omega,t}!}$$

$$\rightarrow -\log P(Y|HU) \stackrel{c}{=}$$

$$\sum_{\omega,t} (-Y_{\omega,t} \log [HU]_{\omega,t} + [HU]_{\omega,t})$$

Poisson分布



NMFアルゴリズムの一般化

- Bregmanダイバージェンス [L.M. Bregman, 1967]

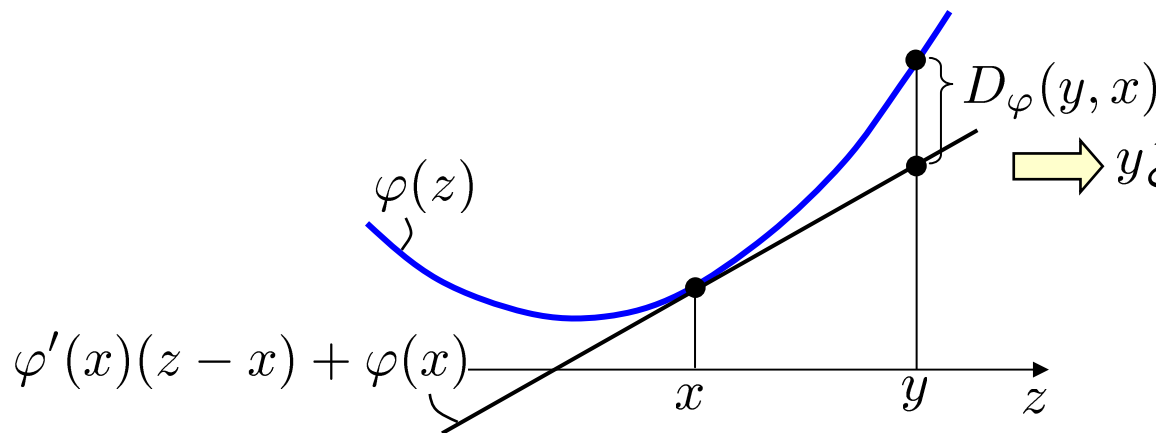
$$D_{\varphi}(y, x) := \varphi(y) - \varphi(x) - \varphi'(x)(y - x)$$

$\varphi : \mathbb{R} \rightarrow \mathbb{R}$ は任意の微分可能な凸関数

→ 指数分布族と対応

[A. Banerjee et al., 2005]

$D_{\varphi}(x, y)$ の最小化は最尤推定問題として解釈可能



⇒ y と x が近いほど小さい

ただし距離の公理は満たさない(対称性が成り立たない)

例) 二乗誤差, Iダイバージェンス, 板倉齋藤距離

- ◆ $\varphi(z) = z^2 \longrightarrow D_{\varphi}(y, x) = y^2 - x^2 - 2x(y - x) = (y - x)^2$

- ◆ $\varphi(z) = z \log z \longrightarrow D_{\varphi}(y, x) = y \log y - x \log x - (1 + \log x)(y - x) = y \log \frac{y}{x} - (y - x)$

- ◆ $\varphi(z) = -\log z \longrightarrow D_{\varphi}(y, x) = -\log y + \log x + \frac{1}{x}(y - x) = \frac{y}{x} - \log \frac{y}{x} - 1$

Bregmanダイバージェンスと対応する距離尺度

y : 観測値 x : NMF型モデル

$$\operatorname{argmin}_x \begin{cases} D_\varphi(y, x) \\ D_\varphi(x, y) \end{cases}$$

[I.S. Dhillon & S. Sra, 2005]

	$D_\varphi(y, x)$	$D_\varphi(x, y)$
$\varphi(z) = z^2$	二乗誤差 $(y - x)^2$	二乗誤差 $(x - y)^2$
$\varphi(z) = z \log z$	Iダイバージェンス $y \log \frac{y}{x} - (y - x)$	双対Iダイバージェンス $x \log \frac{x}{y} - (x - y)$
$\varphi(z) = -\log z$	板倉斎藤距離 $\frac{y}{x} - \log \frac{y}{x} - 1$	双対板倉斎藤距離 $\frac{x}{y} - \log \frac{x}{y} - 1$

[H. Kameoka et al., 2006][C. Févotte et al., 2008]

Bregmanダイバージェンスと対応する尤度関数

y : 観測値 x : NMF型モデル

$$\operatorname{argmin}_x \begin{cases} D_\varphi(y, x) \\ D_\varphi(x, y) \end{cases} \Leftrightarrow \operatorname{argmax}_x P(y|x)$$

	$D_\varphi(y, x)$	$D_\varphi(x, y)$
$\varphi(z) = z^2$	Gauss分布 $y x \sim \mathcal{N}(x, \sigma^2)$	Gauss分布 $y x \sim \mathcal{N}(x, \sigma^2)$
$\varphi(z) = z \log z$	Poisson分布 $y x \sim \text{Pois}(x)$	Gamma分布 $y x \sim \text{Gamma}(1, x + 1)$
$\varphi(z) = -\log z$	指数分布 $y x \sim \text{Exp}(1/x)$	逆Gamma分布 $y x \sim \text{InvGamma}(0, x)$

NMFアルゴリズムの(できるだけ)一般化

- β ダイバージェンス [S. Eguchi & Y. Kano, 2001]

$$D_{\beta}(y, x) := \frac{y^{\beta}}{\beta(\beta-1)} + \frac{x^{\beta}}{\beta} - \frac{yx^{\beta-1}}{\beta-1}$$

Bregmanダイバージェンスにおいて

$$\varphi(z) = \frac{1}{\beta(\beta+1)} (z^{\beta+1} - (\beta+1)z + \beta) \text{ と置いたもの}$$

例) 二乗誤差, Iダイバージェンス, 板倉齋藤距離

$$\blacklozenge \beta = 2 \quad D_{\beta}(y, x) = \frac{y^2}{2} + \frac{x^2}{2} - yx = \frac{1}{2}(y-x)^2$$

$$\blacklozenge \beta \rightarrow 1 \quad D_{\beta}(y, x) \rightarrow y \log \frac{y}{x} + x \quad \because \lim_{\beta \rightarrow 1} \frac{yx^{\beta-1}}{\beta-1} = \lim_{\beta \rightarrow 1} yx^{\beta-1} \log x = y \log x$$

$$\blacklozenge \beta \rightarrow 0 \quad D_{\beta}(y, x) \rightarrow \frac{y}{x} - \log \frac{y}{x} \quad \because \lim_{\beta \rightarrow 0} \frac{x^{\beta}}{\beta} = \lim_{\beta \rightarrow 0} x^{\beta} \log x = \log x$$

NMFアルゴリズム: β ダイバージェンス最小化

[M. Nakano et al., 2010]

$$\text{minimize} \quad \sum_{\omega, t} D_{\beta} \left(Y_{\omega, t}, \sum_i H_{\omega, i} U_{i, t} \right)$$

$$\text{subject to} \quad H_{\omega, i} \geq 0, U_{i, t} \geq 0 \quad (1 \leq \omega \leq \Omega, 1 \leq t \leq T)$$

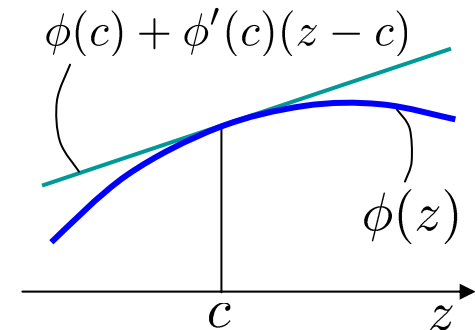
- 目的関数

$$\theta := \{H, U\}$$

$$D(\theta) = \sum_{\omega, t} \left[\frac{Y_{\omega, t}^{\beta}}{\beta(\beta-1)} + \frac{1}{\beta} \underbrace{\left(\sum_i H_{\omega, i} U_{i, t} \right)^{\beta}}_{\text{凸/凹}} - \frac{Y_{\omega, t}}{\beta-1} \underbrace{\left(\sum_i H_{\omega, i} U_{i, t} \right)^{\beta-1}}_{\text{凸/凹}} \right]$$

- 補助関数の設計方針

- 凸不等式 $\phi \left(\sum_i \lambda_i z_i \right) \leq \sum_i \lambda_i \phi(z_i) \quad \begin{cases} \phi : \text{凸関数} \\ \sum_i \lambda_i = 1 \end{cases}$
- 接線不等式 $\phi(z) \leq \phi(c) + \phi'(c)(z - c) \quad \phi : \text{凹関数}$



➡ 凸関数項の上限を凸不等式, 凹関数項の上限を接線不等式を使って設計し, 補助関数を構成 [H. Kameoka et al., 2006]

NMFアルゴリズム： β ダイバージェンス最小化

[M. Nakano et al., 2010]

• 補助関数

$$\frac{1}{\beta} \left(\underbrace{\sum_i H_{\omega,i} U_{i,t}}_{\geq 0} \right) \leq \begin{cases} \text{凸} \left\{ \frac{1}{\beta} \sum_i \lambda_{i,\omega,t} \left(\frac{H_{\omega,i} U_{i,t}}{\lambda_{i,\omega,t}} \right)^\beta \right. & (\beta \geq 1) \\ \text{凹} \left\{ \frac{1}{\beta} c_{\omega,t}^\beta + c_{\omega,t}^{\beta-1} \left(\sum_i H_{\omega,i} U_{i,t} - c_{\omega,t} \right) \right. & (\beta \leq 1, \beta \neq 0) \end{cases}$$

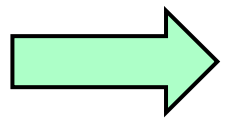
$$-\frac{1}{\beta-1} \left(\sum_i H_{\omega,i} U_{i,t} \right)^{\beta-1} \leq \begin{cases} \text{凸} \left\{ -\frac{1}{\beta-1} \sum_i \lambda_{i,\omega,t} \left(\frac{H_{\omega,i} U_{i,t}}{\lambda_{i,\omega,t}} \right)^{\beta-1} \right. & (\beta \leq 2, \beta \neq 1) \\ \text{凹} \left\{ -\frac{1}{\beta-1} c_{\omega,t}^{\beta-1} - c_{\omega,t}^{\beta-2} \left(\sum_i H_{\omega,i} U_{i,t} - c_{\omega,t} \right) \right. & (\beta \geq 2) \end{cases}$$

(例) $0 < \beta < 1$ のケース:

ただし, $\sum_i \lambda_{i,\omega,t} = 1$

$$G(\theta, \bar{\theta}) = \sum_{\omega,t} \left[\frac{Y_{\omega,t}^\beta}{\beta(\beta-1)} + \frac{1}{\beta} c_{\omega,t}^\beta + c_{\omega,t}^{\beta-1} \left(\sum_i H_{\omega,i} U_{i,t} - c_{\omega,t} \right) \right. \\ \left. - \frac{Y_{\omega,t}}{\beta-1} \sum_i \lambda_{i,\omega,t} \left(\frac{H_{\omega,i} U_{i,t}}{\lambda_{i,\omega,t}} \right)^{\beta-1} \right]$$

$\underbrace{\{c_{\omega,t}\}, \{\lambda_{i,\omega,t}\}}_{\text{subject to } \sum_i \lambda_{i,\omega,t} = 1}$



$$D(\theta) = \min_{\bar{\theta}} G(\theta, \bar{\theta}) \quad \begin{cases} \lambda_{i,\omega,t}^{(n+1)} = \operatorname{argmin}_{\lambda_{i,\omega,t}} G(\theta^{(n)}, \bar{\theta}) = \frac{H_{\omega,i}^{(n)}}{\sum_{i'} H_{\omega,i'}^{(n)} U_{i',t}^{(n)}} \\ c_{\omega,t}^{(n+1)} = \operatorname{argmin}_{c_{\omega,t}} G(\theta^{(n)}, \bar{\theta}) = \sum_{i'} H_{\omega,i'}^{(n)} U_{i',t}^{(n)} \end{cases}$$

NMFアルゴリズム： β ダイバージェンス最小化

[M. Nakano et al., 2010]

- 補助関数 (例) $0 < \beta < 1$ のケース:

$$G(\theta, \bar{\theta}) = \sum_{\omega, t} \left[\frac{Y_{\omega, t}^\beta}{\beta(\beta - 1)} + \frac{1}{\beta} c_{\omega, t}^\beta + c_{\omega, t}^{\beta-1} \left(\sum_i H_{\omega, i} U_{i, t} - c_{\omega, t} \right) - \frac{Y_{\omega, t}}{\beta - 1} \sum_i \lambda_{i, \omega, t} \left(\frac{H_{\omega, i} U_{i, t}}{\lambda_{i, \omega, t}} \right)^{\beta-1} \right]$$

$$\begin{aligned} H_{\omega, i}^{(n+1)} &= \underset{H_{\omega, i}}{\operatorname{argmin}} G(\theta, \bar{\theta}^{(n+1)}) \\ U_{i, t}^{(n+1)} &= \underset{U_{i, t}}{\operatorname{argmin}} G(\theta, \bar{\theta}^{(n+1)}) \end{aligned}$$

$$\begin{cases} \frac{\partial G(\theta, \bar{\theta})}{\partial H_{\omega, i}} = \sum_t \left[c_{\omega, t}^{(n+1)\beta-1} U_{i, t}^{(n)} - \frac{Y_{\omega, t} U_{i, t}^{(n)\beta-1}}{\lambda_{i, \omega, t}^{(n+1)\beta-2}} H_{\omega, i}^{\beta-2} \right] = 0 \\ \frac{\partial G(\theta, \bar{\theta})}{\partial U_{i, t}} = \sum_\omega \left[c_{\omega, t}^{(n+1)\beta-1} H_{\omega, i}^{(n)} - \frac{Y_{\omega, t} H_{\omega, i}^{(n)\beta-1}}{\lambda_{i, \omega, t}^{(n+1)\beta-2}} U_{i, t}^{\beta-2} \right] = 0 \end{cases}$$

$$\begin{aligned} H_{\omega, i}^{(n+1)} &= \left(\frac{\sum_t c_{\omega, t}^{(n+1)\beta-1} U_{i, t}^{(n)}}{\sum_t Y_{\omega, t} U_{i, t}^{(n)\beta-1} / \lambda_{i, \omega, t}^{(n+1)\beta-2}} \right)^{\beta-2} \xrightarrow{\beta \rightarrow 1} H_{\omega, i}^{(n+1)} = \frac{\sum_t Y_{\omega, t} \lambda_{i, \omega, t}^{(n+1)}}{\sum_t U_{i, t}^{(n)}} \\ U_{i, t}^{(n+1)} &= \left(\frac{\sum_\omega c_{\omega, t}^{(n+1)\beta-1} H_{\omega, i}^{(n)}}{\sum_\omega Y_{\omega, t} H_{\omega, i}^{(n)\beta-1} / \lambda_{i, \omega, t}^{(n+1)\beta-2}} \right)^{\beta-2} \xrightarrow{\beta \rightarrow 1} U_{i, t}^{(n+1)} = \frac{\sum_\omega Y_{\omega, t} \lambda_{i, \omega, t}^{(n+1)}}{\sum_\omega H_{\omega, i}^{(n)}} \end{aligned}$$

NMF と pLSA [T. Hofmann, 1999] との関係

(probabilistic Latent Semantic Analysis)

- 各行列要素を確率だと考えてみよう

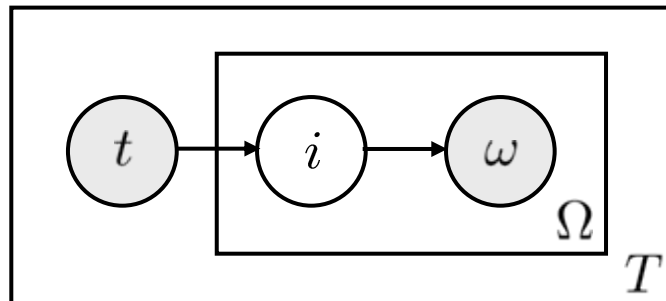
$$X = H U$$

$P(\omega|t)$ 文書 t に単語 ω が現われる確率

$P(\omega|i)$ トピック i のとき単語 ω が生成される確率

$P(i|t)$ 文書 t がトピック i のものである確率

$$\rightarrow P(\omega|t) = \sum_i P(\omega|i)P(i|t)$$



NMF と pLSA [T. Hofmann, 1999] との関係

(probabilistic Latent Semantic Analysis)

- pLSAでは各文書中の単語数 $N_{\omega,t}$ から $P(\omega|i), P(i|t)$ を推定

$$\log P(\omega|t)^{N_{\omega,t}} \rightarrow \text{maximize}$$

$$\begin{aligned} &\text{maximize} && \sum_{\omega,t} N_{\omega,t} \log P(\omega|t) \\ &\text{subject to} && P(\omega|i) \geq 0, P(i|t) \geq 0 \\ &&& \sum_{\omega} P(\omega|i) = 1, \sum_i P(i|t) = 1 \end{aligned}$$

$N_{\omega,t}$ と $P(\omega|t)$ の
Iダイバージェンス
最小化と同じ意味

NMF と Sparse coding [B.A. Olshausen, 1996] との関係

- T 個の観測データ(実数値ベクトル) y_1, \dots, y_T
- すべての観測データが I 個の実数値基底ベクトル a_1, \dots, a_I の "スパース" な線形結合で表される基底セットを求めたい

-exactなreconstruction

$$\begin{array}{ll} \text{minimize} & S(U) \\ \text{subject to} & Y = AU \end{array}$$

NMFでは副次効果的に
スパースになっていたのに
対し, sparse codingで
はスパース性コストにより
強制的にスパース化

-approximation

$$\begin{array}{ll} \text{minimize} & \|Y - AU\|_F^2 + \lambda S(U) \\ \text{subject to} & \|U\|_F^2 = 1 \end{array}$$

勾配法等

- スパース性コスト [B.A. Olshausen, 1996][G. Harpur, 2000]

- L_p ノルム $S(U) = \sum_{i,t} |U_{i,t}|^p$

$$S(U) = \sum_{i,t} \log(1 + U_{i,t}^2) \quad \text{尖度} \quad S(U) = \sum_i \left(\frac{\langle (U_{i,t} - \mu_i)^4 \rangle_t}{\sigma_i^2} - 3 \right)$$

NMFによる振幅スペクトログラムの分解表現の問題点

観測振幅スペクトルを振幅スペクトルパーツの和でモデル化

😞 振幅スペクトルの加法性を仮定した不正確なモデル

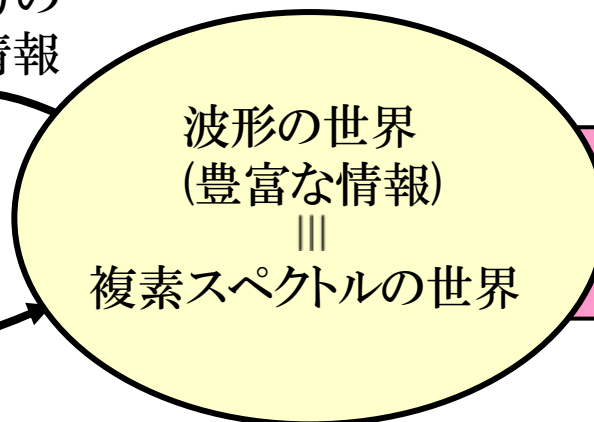
- 振幅スペクトルは本当は非加法的（加法的なのは波形！）
- ➔ モデル自体に誤差を含んでいる

😞 位相情報の放棄

- 推定したモデルが波形に対応しない
- 波形領域で実現される多くの信号処理法と組み合わせられない

NMFモデルが
仮定される世界

周波数成分の
含有率の情報



広い
応用展開

位相を捨てるため元に戻せない！

振幅スペクトルの非加法性について

◆音響信号 $y(u) \in \mathbb{R}$ \longrightarrow $y(u) = \sum_i x^i(u)$ (波形同士は加法的)

↓ 短時間フーリエ変換
(時間周波数分解) ...線形な変換

$Y_{\omega,t} = \int W(u)y(u+t)e^{j\omega u}du \in \mathbb{C} \longrightarrow Y_{\omega,t} = \sum_i X_{\omega,t}^i$ (もちろん加法的)

↓ 絶対値をとる ...非線形な変換

$|Y_{\omega,t}|$: 時刻 t に周波数 ω の成分が
どれほど含まれているか $\longrightarrow |Y_{\omega,t}| = \left| \sum_i X_{\omega,t}^i \right|$

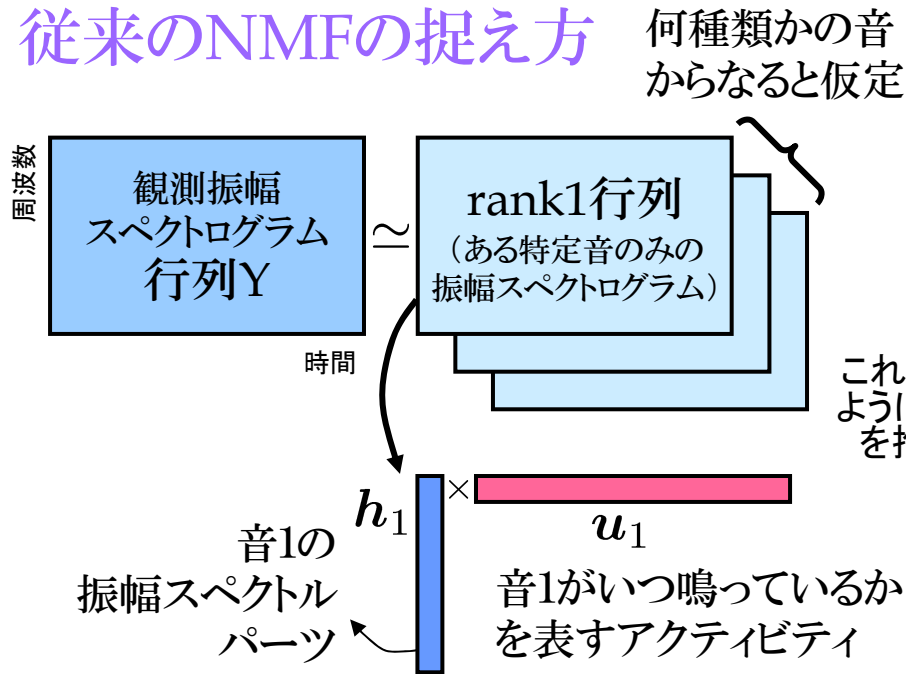
$Y \simeq HU$ ($H \succeq 0, U \succeq 0$) $\neq \sum_i |X_{\omega,t}^i|$ (振幅スペクトル
同士は**非加法的**)

NMFモデル

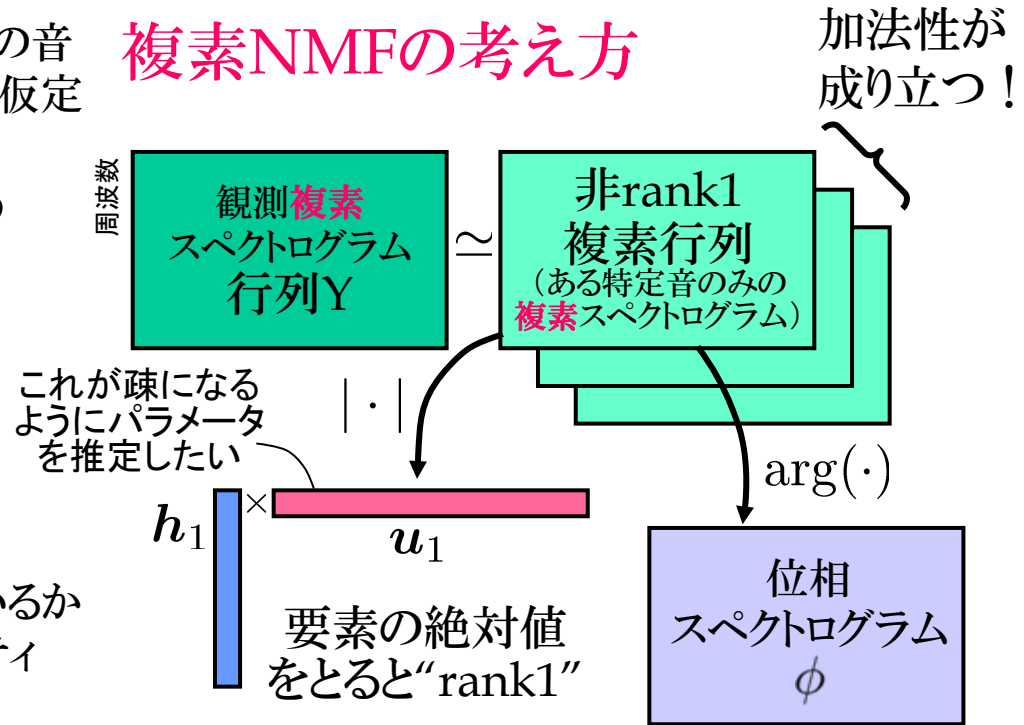
$|Y_{\omega,t}| = \sum_i \underbrace{H_{\omega}^i U_t^i}_{\text{振幅スペクトル?}}$

新モデル「複素NMF」の提案

従来のNMFの捉え方



複素NMFの考え方



◆複素NMFモデル

- 複素スペクトログラム $Y_{\omega,t} \in \mathbb{C}$ をモデル化

$$Y_{\omega,t} = \sum_i X_{\omega,t}^i \quad \swarrow \quad |X_{\omega,t}^i| = H_{\omega}^i U_t^i$$

行列積の形にならない！
(新しいクラスのスパース表現モデル)

$$= \sum_i \underline{|X_{\omega,t}^i|} e^{j\phi_{\omega,t}^i} = \sum_i \underline{H_{\omega}^i U_t^i} e^{j\phi_{\omega,t}^i}$$

$H_{\omega}^i \geq 0, U_t^i \geq 0$

複素NMFアルゴリズム

- 定義

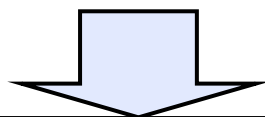
- 複素スペクトログラム $Y_{\omega,t} \in \mathbb{C}$
 - 振幅スペクトル基底 $\mathcal{H} \equiv \{H_{\omega}^i\}$
 - 位相スペクトログラム $\phi \equiv \{\phi_{\omega,t}^i\}$
 - ゲイン $\mathcal{U} \equiv \{U_t^i\}$
- $F_{\omega,t} = \sum_{i=1}^I H_{\omega}^i U_t^i e^{j\phi_{\omega,t}^i}$

音響信号モデル

- 最適化問題

$$Y_{\omega,t} = F_{\omega,t} + \epsilon_{\omega,t} \begin{cases} \epsilon_{\omega,t} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2) \\ U_t^i \sim \mathcal{GGD}(0, b, p) \quad (0 < p < 2) \end{cases}$$

↪ 一般化正規分布



minimize $f(\mathcal{H}, \mathcal{U}, \phi) := \sum_{\omega,t} |Y_{\omega,t} - F_{\omega,t}|^2 + 2\lambda \sum_{i,t} |U_t^i|^p$

subject to $\sum_{\omega} H_{\omega}^i = 1 \quad (i = 1, \dots, I)$

スパース正則化項

複素NMFアルゴリズム

• 補助関数法

$$f(\mathcal{H}, \mathcal{U}, \phi) = \sum_{\omega, t} |Y_{\omega, t} - F_{\omega, t}|^2 + 2\lambda \sum_{i, t} |U_t^i|^p$$

$$\leq \sum_{\omega, t} \sum_i \frac{|\bar{Y}_{\omega, t}^i - H_{\omega}^i U_t^i e^{j\phi_{\omega, t}^i}|^2}{\beta_{\omega, t}^i} + \lambda \sum_{i, t} (p|\bar{U}_t^i|^{p-2} U_t^{i2} + (2-p)|\bar{U}_t^i|^p)$$

subject to $\sum_i \bar{Y}_{\omega, t}^i = Y_{\omega, t}$

$0 < \beta_{\omega, t}^i < 1, \sum_i \beta_{\omega, t}^i = 1$
を満たす任意の定数

$=: g(\mathcal{H}, \mathcal{U}, \phi, \bar{Y}, \bar{U})$

ℓ 回目の反復計算後のパラメータ値: $\mathcal{H}^{(\ell)} = \{H_{\omega}^{i(\ell)}\}$, $\mathcal{U}^{(\ell)} = \{U_t^{i(\ell)}\}$, $\phi^{(\ell)} = \{\phi_{\omega, t}^{i(\ell)}\}$

Step 1) $\bar{Y}_{\omega, t}^{i(\ell+1)} \leftarrow H_{\omega}^{i(\ell)} U_t^{i(\ell)} e^{j\phi_{\omega, t}^{i(\ell)}} + \beta_{\omega, t}^i (Y_{\omega, t} - F_{\omega, t}^{(\ell)})$

Step 2) $\bar{U}_t^{i(\ell+1)} \leftarrow U_t^{i(\ell)}$

Step 2) $H_{\omega}^{i(\ell+1)} \leftarrow \frac{\sum_t U_t^{i(\ell)} \text{Re}[\bar{Y}_{\omega, t}^{i(\ell+1)*} e^{j\phi_{\omega, t}^{i(\ell)}}]}{\sum_t U_t^{i(\ell)2} / \beta_{\omega, t}^i}$ **必ず正值**

Step 2) $U_t^{i(\ell+1)} \leftarrow \frac{\sum_t H_{\omega}^{i(\ell)} \text{Re}[\bar{Y}_{\omega, t}^{i(\ell+1)*} e^{j\phi_{\omega, t}^{i(\ell)}}]}{\sum_t U_t^{i(\ell)2} / \beta_{\omega, t}^i + \lambda p |\bar{U}_t^{i(\ell+1)}|^{p-2}}$ **必ず正值**

Step 2) $e^{j\phi_{\omega, t}^{i(\ell+1)}} \leftarrow \bar{Y}_{\omega, t}^{i(\ell+1)} / |\bar{Y}_{\omega, t}^{i(\ell+1)}|$

$f(\mathcal{H}^{(\ell)}, \mathcal{U}^{(\ell)}, \phi^{(\ell)})$

$= g(\mathcal{H}^{(\ell)}, \mathcal{U}^{(\ell)}, \phi^{(\ell)}, \bar{Y}^{(\ell+1)}, \bar{U}^{(\ell+1)})$

$\geq g(\mathcal{H}^{(\ell+1)}, \mathcal{U}^{(\ell+1)}, \phi^{(\ell+1)}, \bar{Y}^{(\ell+1)}, \bar{U}^{(\ell+1)})$

$\geq f(\mathcal{H}^{(\ell+1)}, \mathcal{U}^{(\ell+1)}, \phi^{(\ell+1)})$

f は増加しない!

NMFと等価となる条件

複素NMFはNMFを包含

Step 1) $\bar{Y}_{\omega,t}^i \leftarrow H_{\omega}^i U_t^i e^{j\phi_{\omega,t}^i} + \beta_{\omega,t}^i (Y_{\omega,t} - F_{\omega,t})$

$$\bar{U}_t^i \leftarrow U_t^i$$

Step 2) $H_{\omega}^i \leftarrow \frac{\sum_t U_t^i \operatorname{Re}[\bar{Y}_{\omega,t}^{i*} e^{j\phi_{\omega,t}^i}]/\beta_{\omega,t}^i}{\sum_t U_t^{i2}/\beta_{\omega,t}^i}$

$$U_t^i \leftarrow \frac{\sum_{\omega} H_{\omega}^i \operatorname{Re}[\bar{Y}_{\omega,t}^{i*} e^{j\phi_{\omega,t}^i}]/\beta_{\omega,t}^i}{\sum_t U_t^{i2}/\beta_{\omega,t}^i + \lambda p |\bar{U}_t^i|^{p-2}}$$

$$e^{j\phi_{\omega,t}^i} \leftarrow \bar{Y}_{\omega,t}^i / |\bar{Y}_{\omega,t}^i|$$

[条件1]

$\phi_{\omega,t}^i$ を $\arg(Y_{\omega,t})$ に
初期設定する

$\phi_{\omega,t}^i = \arg(Y_{\omega,t})$ は
Step1, Step2に対して
不動点になっている

[条件2]

Step 3) $\beta_{\omega,t}^i \leftarrow \frac{H_{\omega}^i U_t^i}{\sum_k H_{\omega}^k U_t^k}$ を実行

$\beta_{\omega,t}^i : 0 < \beta_{\omega,t}^i < 1, \sum_i \beta_{\omega,t}^i = 1$
を満たす任意の定数

$$H_{\omega}^i \leftarrow H_{\omega}^i \frac{\sum_t U_t^i |Y_{\omega,t}|}{\sum_t U_t^i \sum_k H_{\omega}^k U_t^k}$$

$$U_t^i \leftarrow U_t^i \frac{\sum_{\omega} H_{\omega}^i |Y_{\omega,t}|}{\sum_{\omega} H_{\omega}^i \sum_k H_{\omega}^k U_t^k + \lambda p |U_t^i|^{p-2}}$$

Lee & Seungの二乗誤差規準のNMFアルゴリズム[Lee & Seung, 2000]と等価!

NMFによる モノラル信号分離 デモンストレーション

Complex NMF

by H. Kameoka

Open

サンプリング 16000 Hz

フレーム長 64 ms

トータル時間 23 s

Factorize

スペクトルパーツ数 30

エンベロープパーツ数 5

Save

表示するスペクトルパーツ

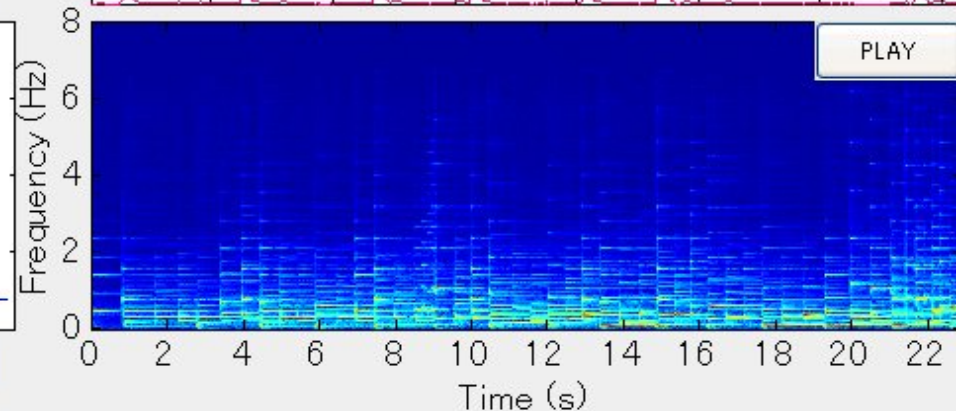
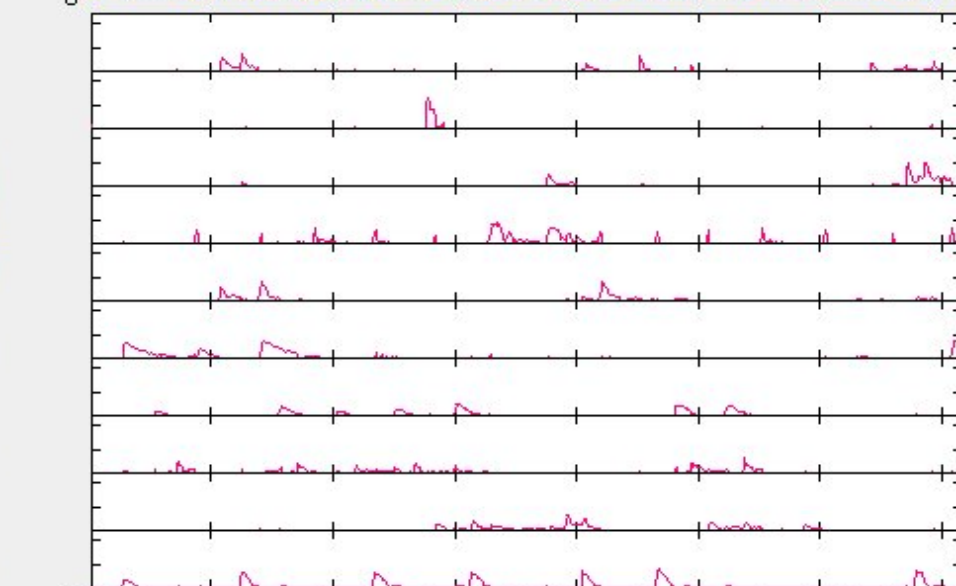
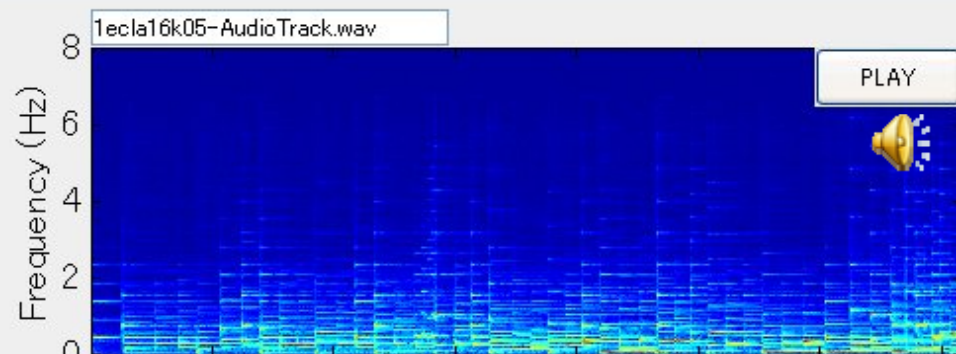
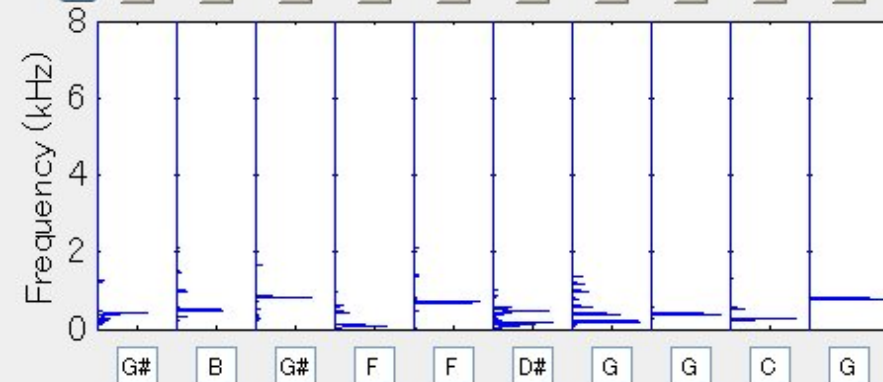
No. 1 - 10

トランスポーズ

0	+	+	+	+	+	+	+	+	+	+
m	-	-	-	-	-	-	-	-	-	-

ボリュームコントロール

0										
M										



Complex NMF

by H. Kameoka

Open	サンプリング	16000 Hz	フレーム長	64 ms
	トータル時間	23 s		
Factorize	スペクトルパーツ数	30	Save	
	エンベロープパーツ数	5		

表示するスペクトルパーツ

No. 1 - 10

ト

0 m

「ソ」以外は音量オフ

0 0

0 0

+

+

-

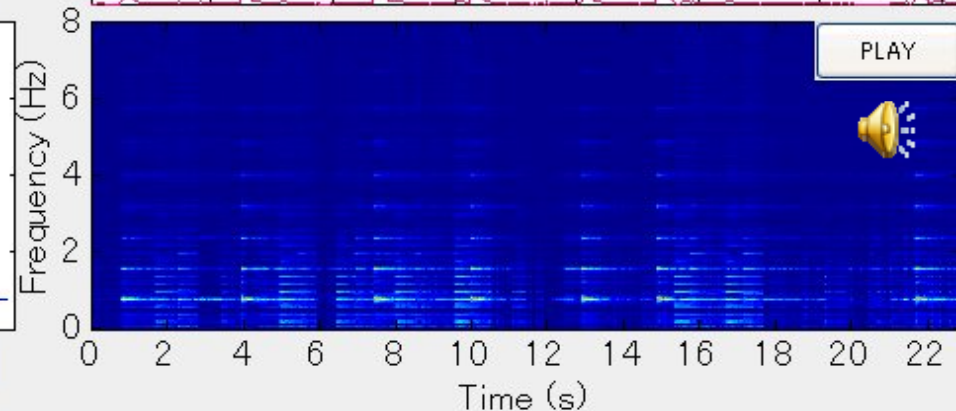
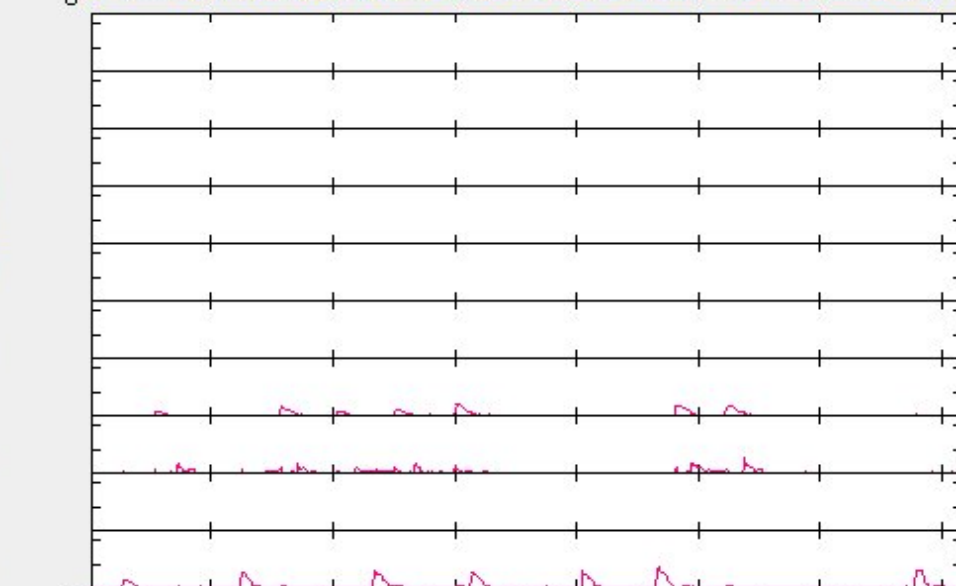
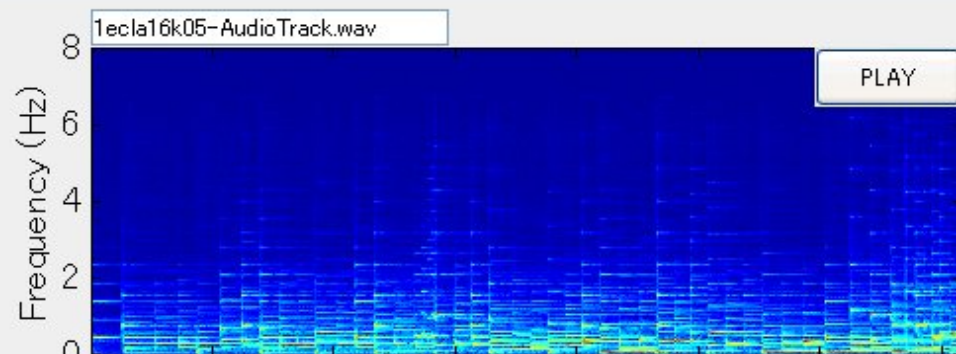
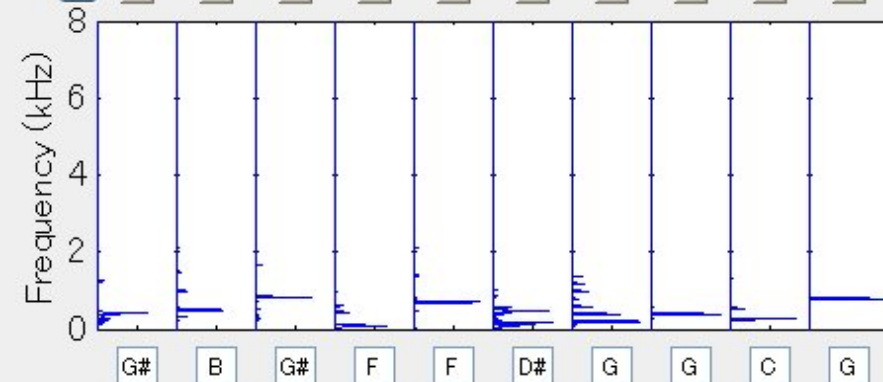
-

ボリュームコントロール

0 M

8

0 0 0 0 0 0 0 0 0 0



Complex NMF

by H. Kameoka

Open	サンプリング	16000 Hz	フレーム長	64 ms
	トータル時間	23 s		
Factorize	スペクトルパーツ数	30	Save	
	エンベロープパーツ数	5		

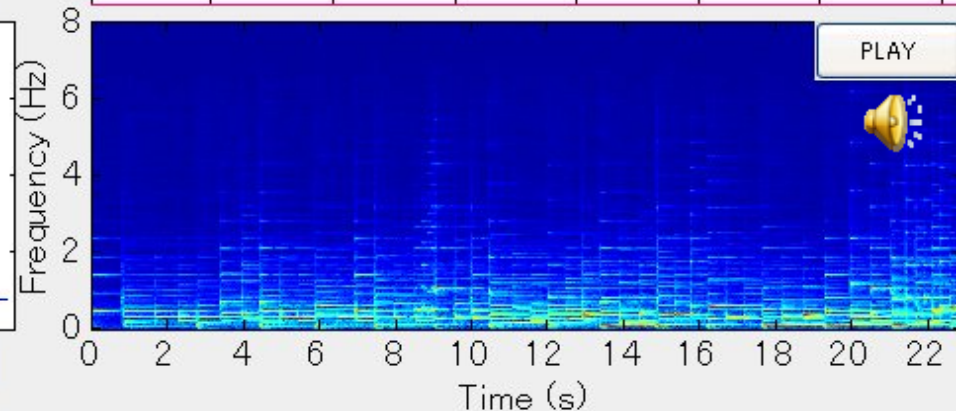
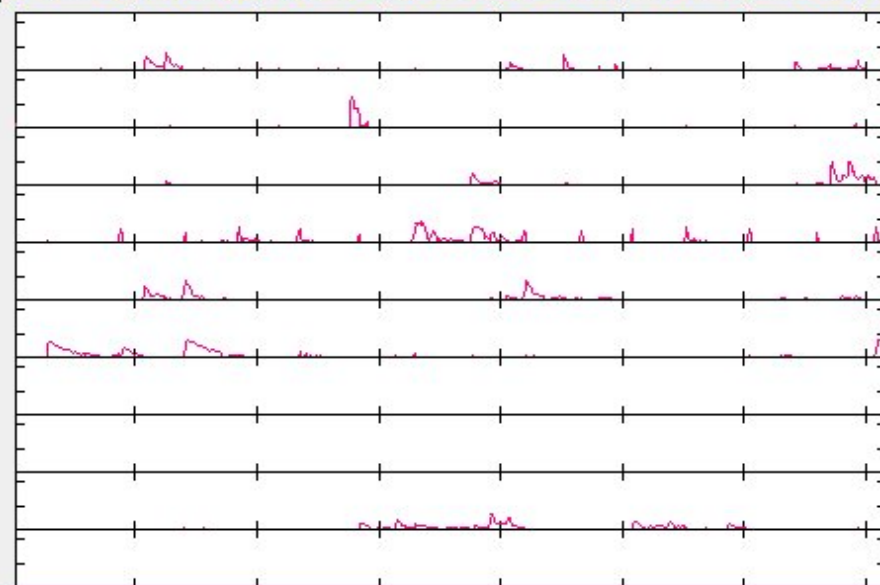
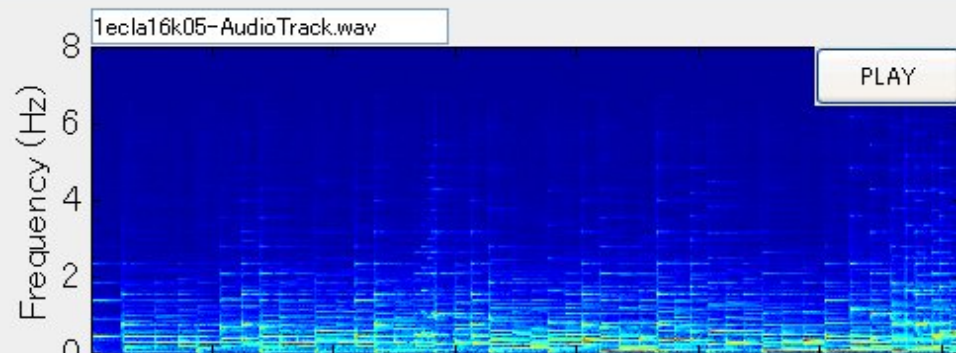
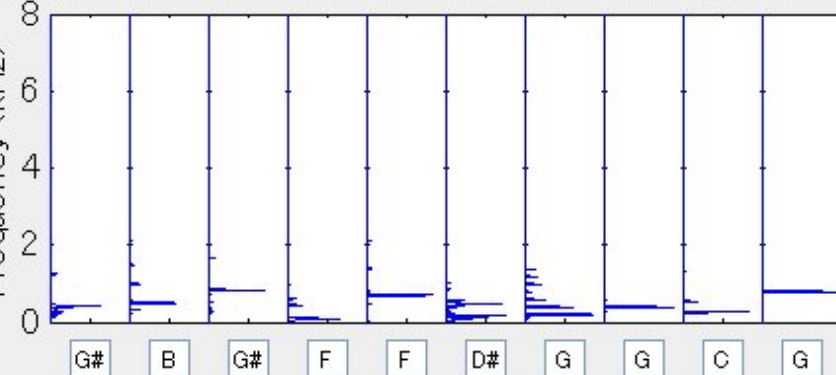
表示するスペクトルパーツ

No. 1 - 10

「ソ」だけ音量オフ

ボリュームコントロール

0	M																		
---	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



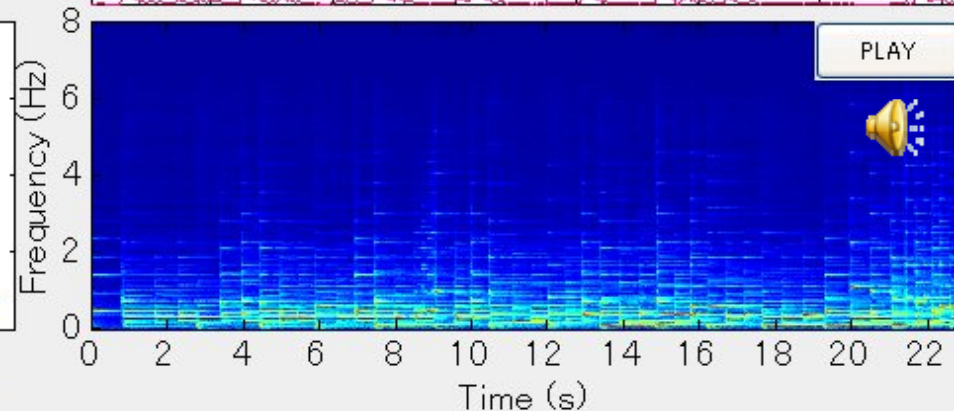
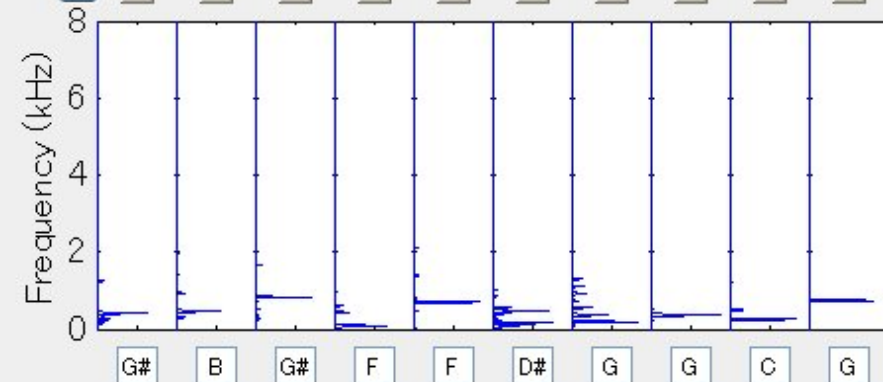
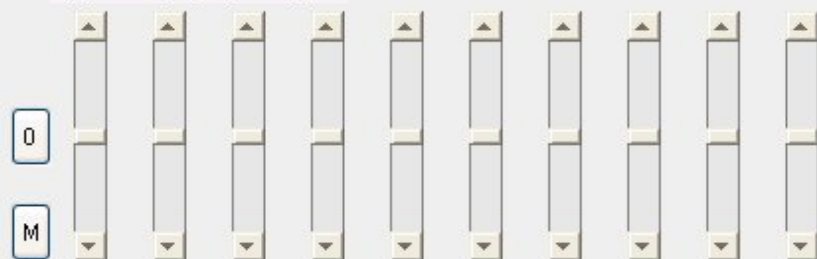
一部の基底関数に対してのみスペクトル伸縮変形を施し、混合信号を再構成（スペクトル伸縮はピッチトランスポーズに相当。例えばトランスポーズ「-1」は、半音下げという意味。）

No. 1-10

トランスポ



ボリュームコントロール



back.wav

PLAY

PLAY

まとめ

1. 非負値行列因子分解(NMF)とは

- 何に使えるのか(音響信号処理を題材として)
- どのような性質があるのか
- どのように求めるのか
- 統計モデルとしての解釈とアルゴリズムの一般化
- 他の技術とどう関係しているのか

2. 複素NMF [H. Kameoka et al., 2008]

- NMFにおける問題点
- 解決のためのアイデア
- NMFとどう関係しているのか

参考文献

- [1] D. P. W. Ellis and J. Arroyo, "Eigenrhythms: Drum pattern basis sets for classification and generation," *Proc. ISMIR 2004*, pp. 554-559, 2004.
- [2] D. D. Lee and H. S. Seung, "Learning the parts of objects with nonnegative matrix factorization," *Nature*, vol. 401, pp. 788-791, 1999.
- [3] P. Paatero, U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics* **5**: 111-126, 1994.
- [3] P. Smaragdis and J.C. Brown, "Non-negative matrix factorization for music transcription," *Proc. WASPAA 2003*, pp. 177-180, 2003.
- [4] D.D. Lee and H.S. Seung, "Algorithms for nonnegative matrix factorization," *Proc. NIPS 2000*, pp. 556-562, 2000.
- [5] P. Smaragdis, B. Raj, M.V. Shashanka, "Supervised and semi-supervised separation of sounds from single-channel mixtures," *Proc. 7th International Conference on Independent Component Analysis and Signal Separation*, 2007.
- [6] P. Smaragdis and B. Raj, "Example-driven bandwidth expansion," *Proc. WASPAA 2007*, in CD-ROM, 2007.
- [7] L.M. Bregman, "The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming," *USSR Computational Mathematics and Mathematical Physics* **7**: pp. 200-217, 1967.
- [8] A. Banerjee, S. Merugu, I.S. Dhillon and J. Ghosh, "Clustering with Bregman divergences," *Journal of Machine Learning Research* **6**, pp. 1705-1749, 2005.
- [9] I.S. Dhillon and S. Sra, "Generalized nonnegative matrix approximations with Bregman divergences," *Proc. NIPS 2005*, pp. 283-290, 2005.

参考文献

- [10] J. de Leeuw. Block-relaxation methods in statistics. In H.H. Bock, W. Lenski, and M.M. Richter, editors, *Information Systems and Data Analysis*, Berlin, 1994. Springer Verlag.
- [11] 亀岡, 後藤, 嵯峨山, “スペクトル制御エンベロープによる混合音中の周期および非周期成分の選択的イコライザ,” *情報研報*, 2006-MUS-66-13, pp. 77-84, 2006.
- [12] C. Févotte, N. Bertin and J.-L. Durrieu, “Nonnegative matrix factorization with the Itakura-Saito divergence with application to music analysis,” Technical Report TELECOM ParisTech 2008D006, 2008.
- [13] S. Eguchi and Y. Kano, “Robustifying maximum likelihood estimation,” Technical report, Institute of Statistical Mathematics, Research Memo. 802, 2001.
- [14] M. Nakano, H. Kameoka, J. Le Roux, Y. Kitano, N. Ono and S. Sagayama, “Convergence-guaranteed multiplicative algorithms for nonnegative matrix factorization with beta-divergence,” *Proc. MLSP2010*, to appear.
- [15] B.A. Olshausen, and D.J. Field, “Emergence of simple-cell receptive field properties by learning a sparse code for natural images,” *Nature*, vol. 381, pp. 607–609, 1996.
- [16] G. Harpur, R. Prager, “Experiments with low-entropy neural networks,” in R. Baddeley, P. Hancock, P. Földiák (eds.) *Information Theory and the Brain*, Cambridge University Press, pp. 84-100, 2000.
- [17] T. Hofmann, “Probabilistic latent semantic analysis,” *Proc. UAI*, pp. 289-296, 1999.
- [18] 亀岡, 小野, 柏野, 嵯峨山, “複素NMF: 新しいスパース信号分解表現と基底系学習アルゴリズム,” *日本音響学会2008年秋季研究発表会講演論文集*, 2-8-13, pp. 657-660, 2008.