

第6回

音声音響信号処理

(非負値行列因子分解)

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

10/3 守谷先生担当

10/17 守谷先生担当

10/24 線形予測分析と自己回帰モデル

10/31 拡散性雑音に頑健なマイクロホンアレイ信号処理

11/7 (休講)

11/14 時間周波数解析

→ 11/21 **非負値行列因子分解**

11/28

12/5

12/12

12/19 (休講)

1/16

1/23

1/30

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。

※講義用ホームページのURL:

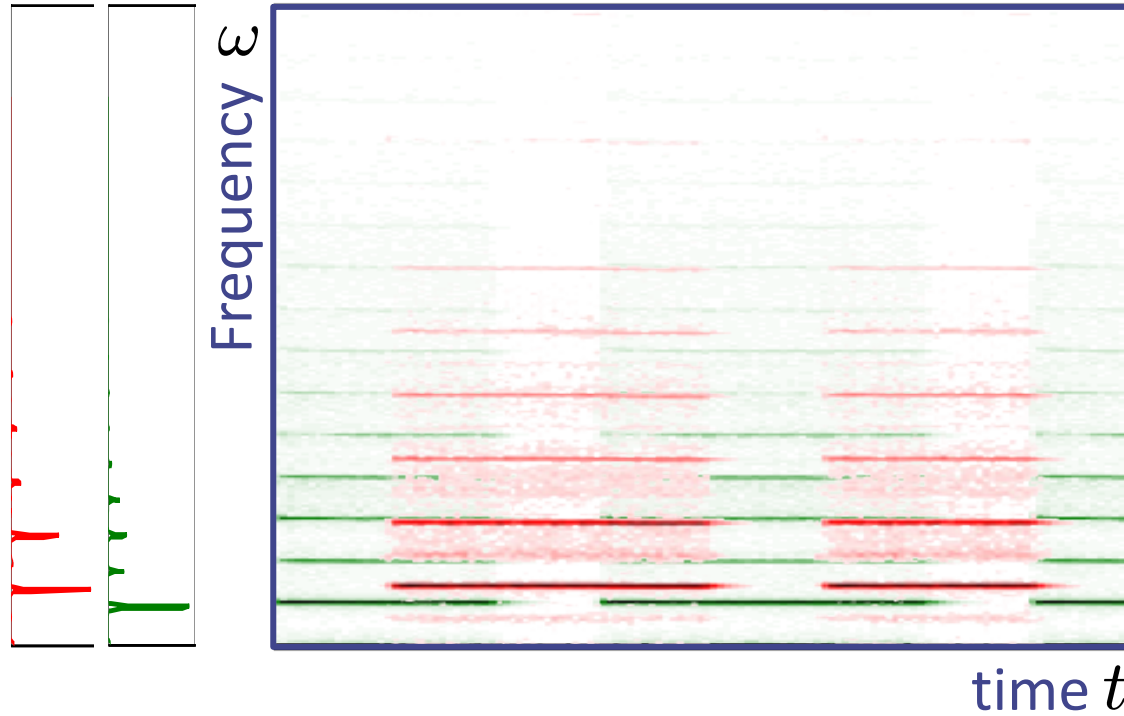
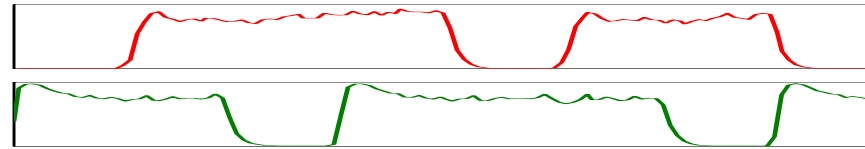
<http://hil.t.u-tokyo.ac.jp/~kameoka/SAP/>

本日の話題

◆非負値行列因子分解

- データ行列を基底行列と係数行列の積に分解する
多変量解析手法の一つ
 - ◆ 非負制約を置く所がポイント
 - ◆ スパース性という概念と密接に関係
- 画像処理分野で登場した技術
 - ◆ 顔画像データから目・鼻・口などの特徴的パーツを抽出することが目的
- 近年, モノラル音源分離の手法としても注目
 - ◆ スペクトログラムをデータ行列と見なして, 行列の積に分解
 - ◆ ICASSP2011(音響信号処理分野で最大規模の国際会議)では21セッション中に33件もの関連発表が!
 - ◆ 国内でも徐々に関連研究が盛んになりつつある

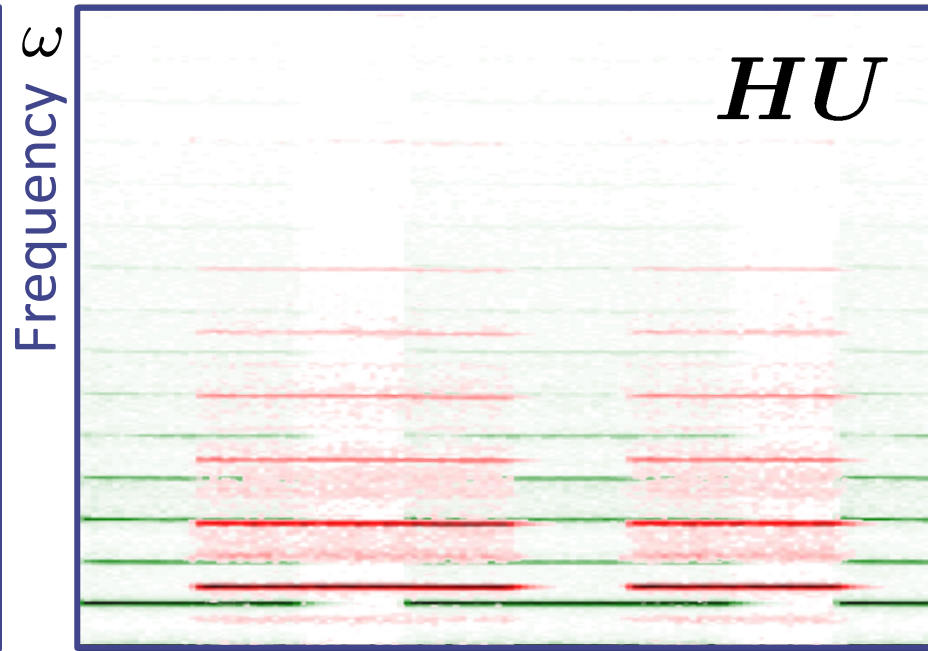
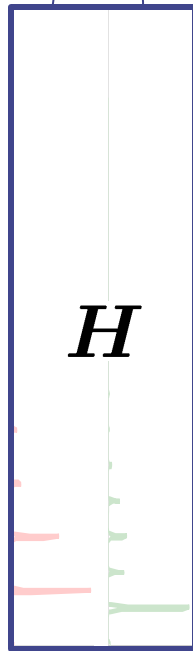
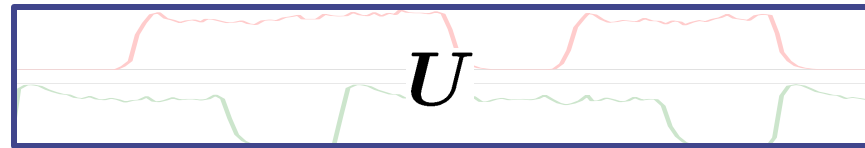
「行列の積」としてのスペクトログラム



「行列の積」としてのスペクトログラム

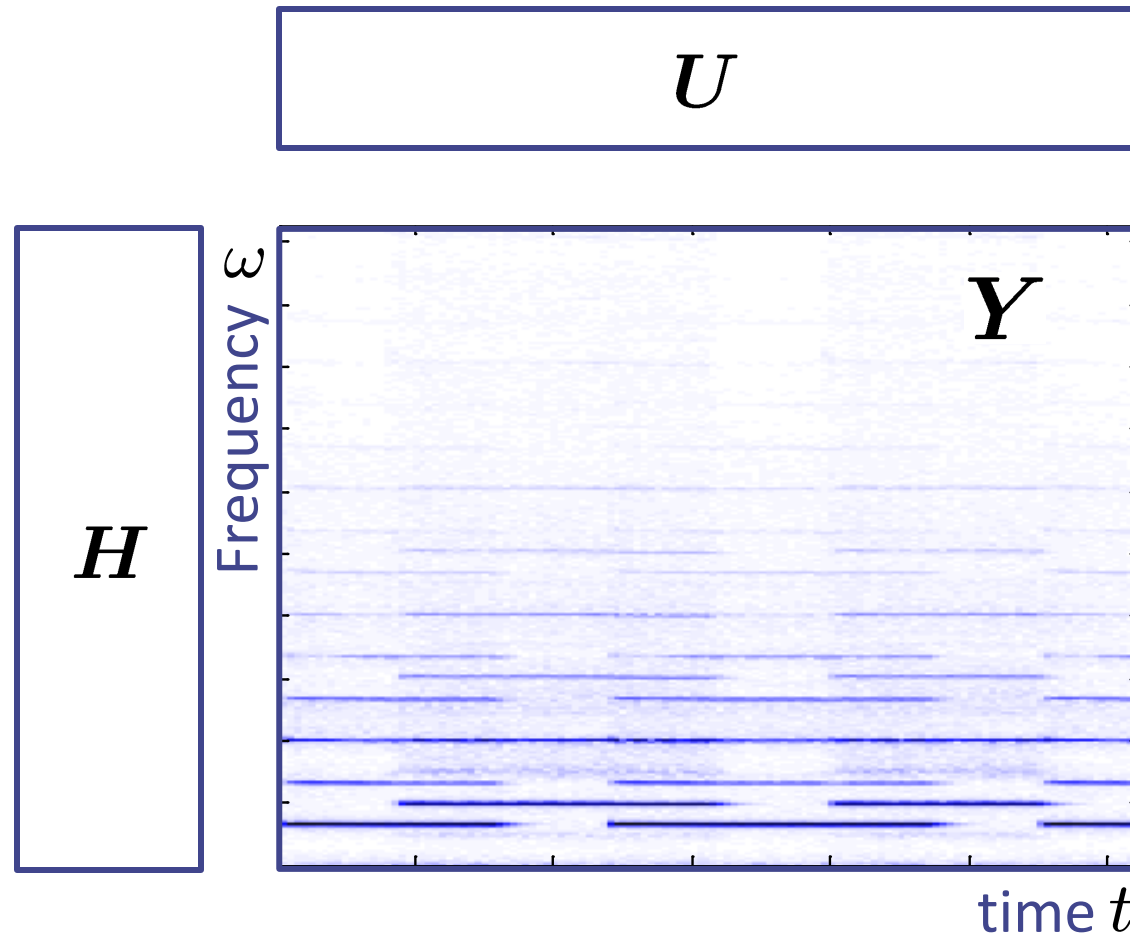
各基底の「アクティベーション」

「基底スペクトル」



time t

「行列の積」としてのスペクトログラム



目次

1. 非負値行列因子分解(NMF)

- 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

2. 複素NMF [Kameoka2008]

- NMFによるスペクトログラム分解法が根源的に抱える問題点
- これを解決する新モデル

3. NMFの改良・拡張モデル

- スパースNMF, NMFD, NMF2D, ソースフィルタNMF, アクティベーションの連続性規準入りNMF, ハーモニックNMF, 板倉齋藤距離規準NMF, 複素NMF, 状態遷移NMF, ノンパラメトリックベイズNMF, etc...

目次

1. 非負値行列因子分解(NMF)

- 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

2. 複素NMF [Kameoka2008]

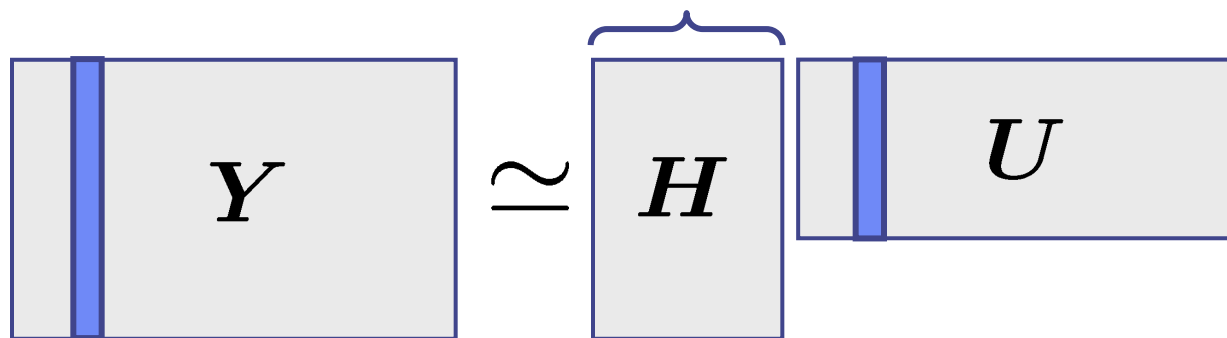
- NMFによるスペクトログラム分解法が根源的に抱える問題点
- これを解決する新モデル

3. NMFの改良・拡張モデル

- スパースNMF, NMFD, NMF2D, ソースフィルタNMF, アクティベーションの連続性規準入りNMF, ハーモニックNMF, 板倉齋藤距離規準NMF, 複素NMF, 状態遷移NMF, ノンパラメトリックベイズNMF, etc...

非負値行列因子分解 (Nonnegative Matrix Factorization) とは

◆非負値行列を2つの非負値行列の積で表現



◆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

- 次元圧縮

H の各列: 基底ベクトル U の各列: 結合係数

NMFが生まれた背景

- ◆ 画像処理分野で生まれた技術 [Lee1999]
 - 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 [Paatero1994]

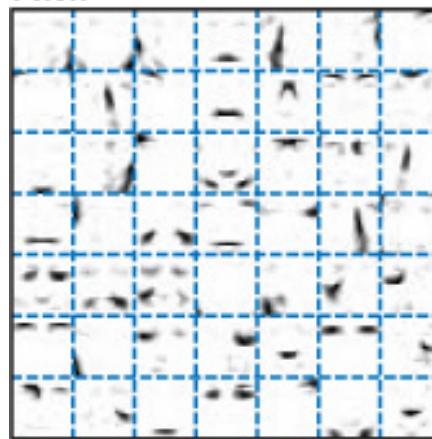
- ◆ 音のスペクトルを画像と見なして適用(後述)
→ 音声分離・自動採譜等
[Smaragdis2003]以降極めて多数...

- ◆ 効率的な反復アルゴリズム [Lee2000]

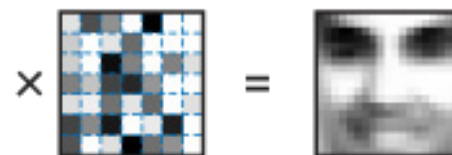
$$H_{\omega,k} \leftarrow H_{\omega,k} \frac{[YU^T]_{\omega,k}}{[HUU^T]_{\omega,k}}$$

$$U_{k,t} \leftarrow U_{k,t} \frac{[H^T Y]_{k,t}}{[H^T HU]_{k,t}}$$

NMF



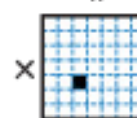
Original



VQ

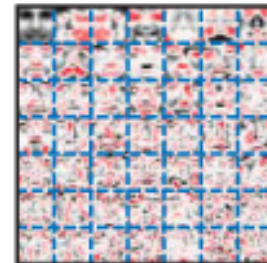


||



x

PCA



||



x

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(でないとは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

$\times 1$



$\times 1$



$\times (-2)$

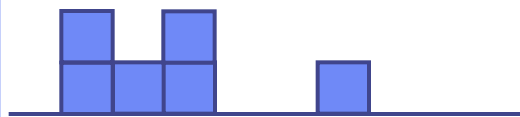
+)



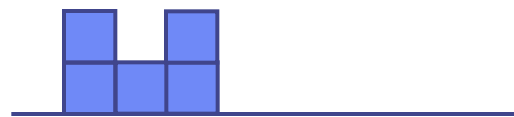
係数行列の非負性について

◆ 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

×1



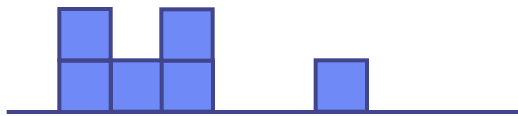
+)



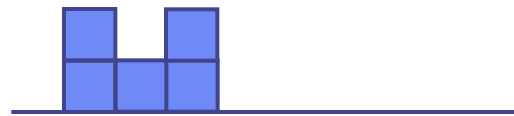
係数行列の非負性について

◆ 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

$\times 1$

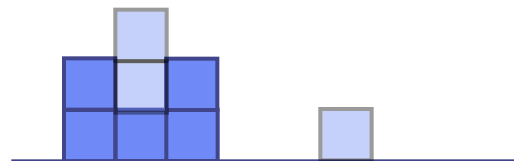


($\times 1$?)



(引き算できない
んだった...)

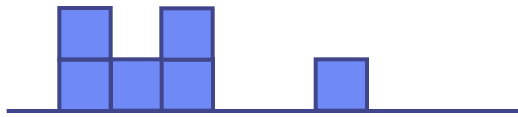
+)



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

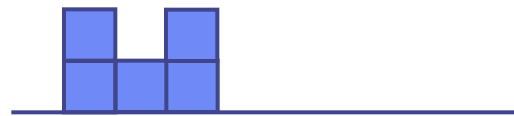
観測データ



基底ベクトル



+)



係数

×1

×0

×0

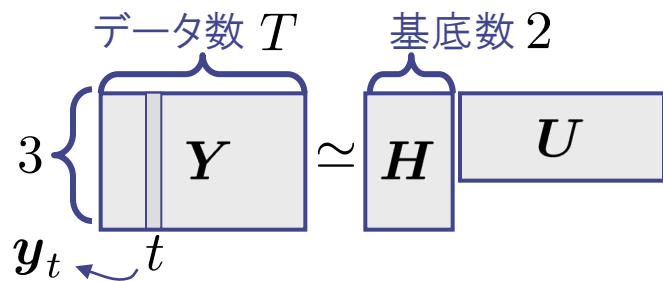
スパース

別の見方から...

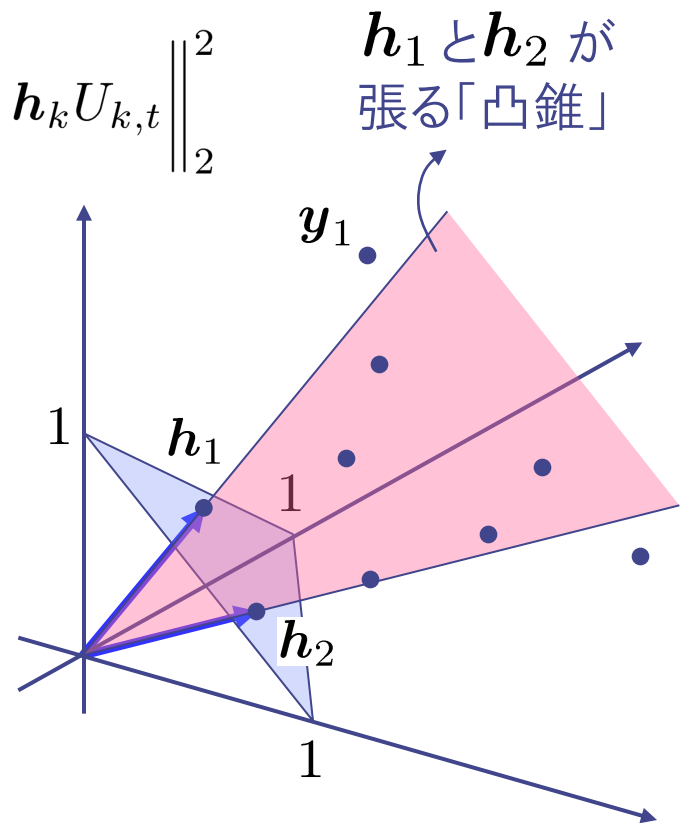
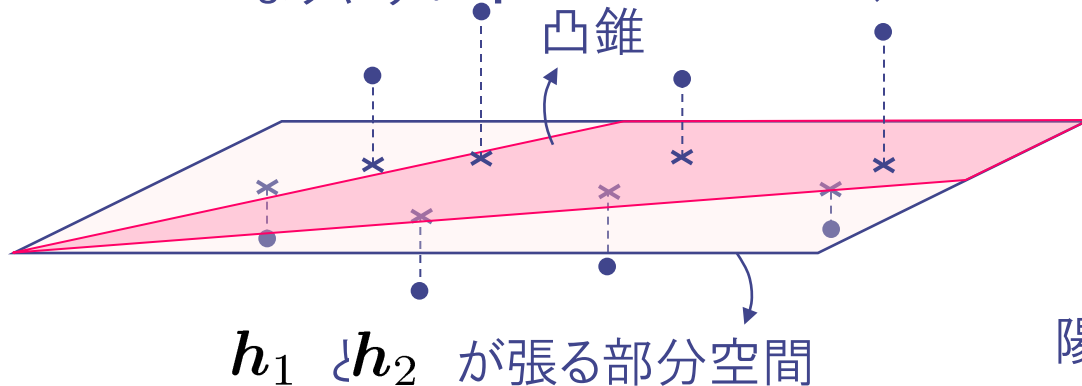
◆ Frobeniusノルム規準のNMF

$$\text{minimize } \|Y - HU\|_F^2 = \sum_t \left\| \mathbf{y}_t - \sum_k \mathbf{h}_k U_{k,t} \right\|_2^2$$

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



斜交基底だから係数はスパースになりやすい?



\mathbf{h}_1 と \mathbf{h}_2 のなす角を大きくした方がお得!

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

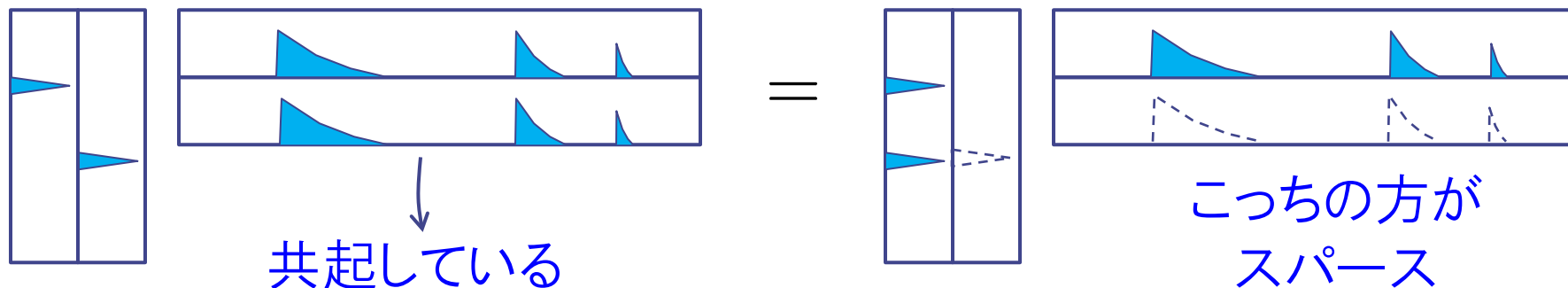
スパース性がもたらす効果


$$Y \approx H U$$

これがスパースになるよう誘導される

◆ HがどうなればUはスパースになる？

- 共起するパーツがあればひとまとめにした方がスパースに！



- 共起頻度が高いパーツのペアをひとまとめにした方がスパースに！
(上のひとまとめ操作では  を3つ分減らせている)

頻出するひとまとまりのパーツが

各基底ベクトルとして得られる傾向になる

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

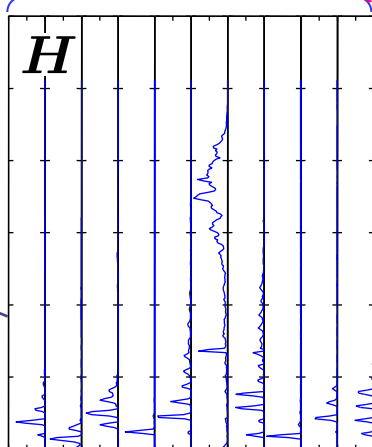
$|Y_{\omega,t}|$: 時刻 t に周波数 ω の成分が
どれほど含まれているか

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

各基底の
アクティビティ

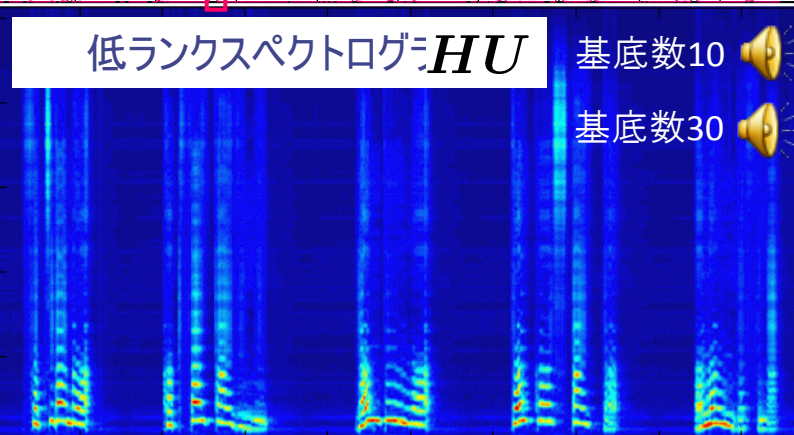
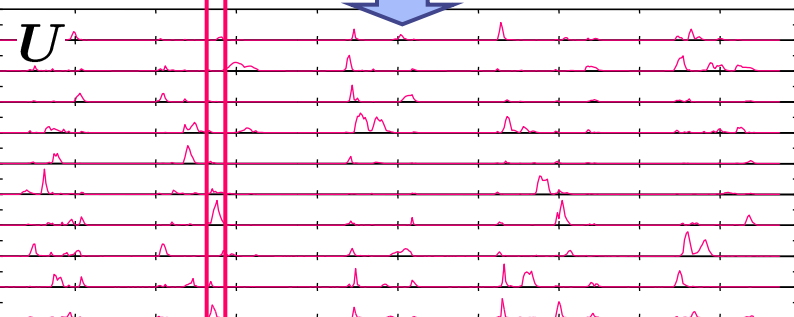
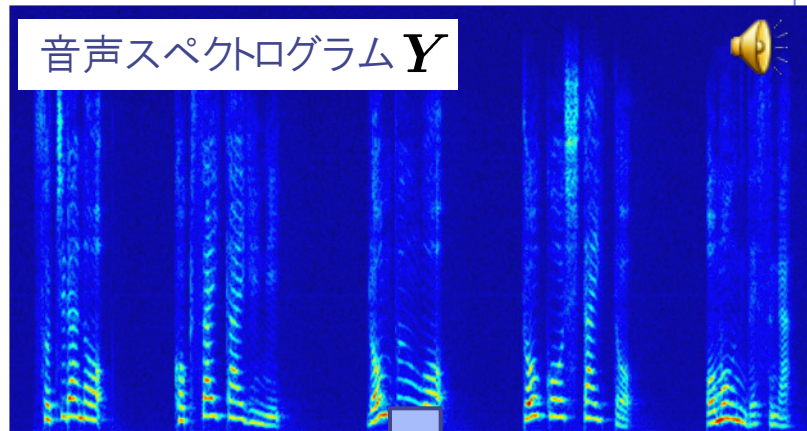
基底スペクトル

周波数 \uparrow



頻出するスペクトル
パーツが獲得される

周波数 \uparrow

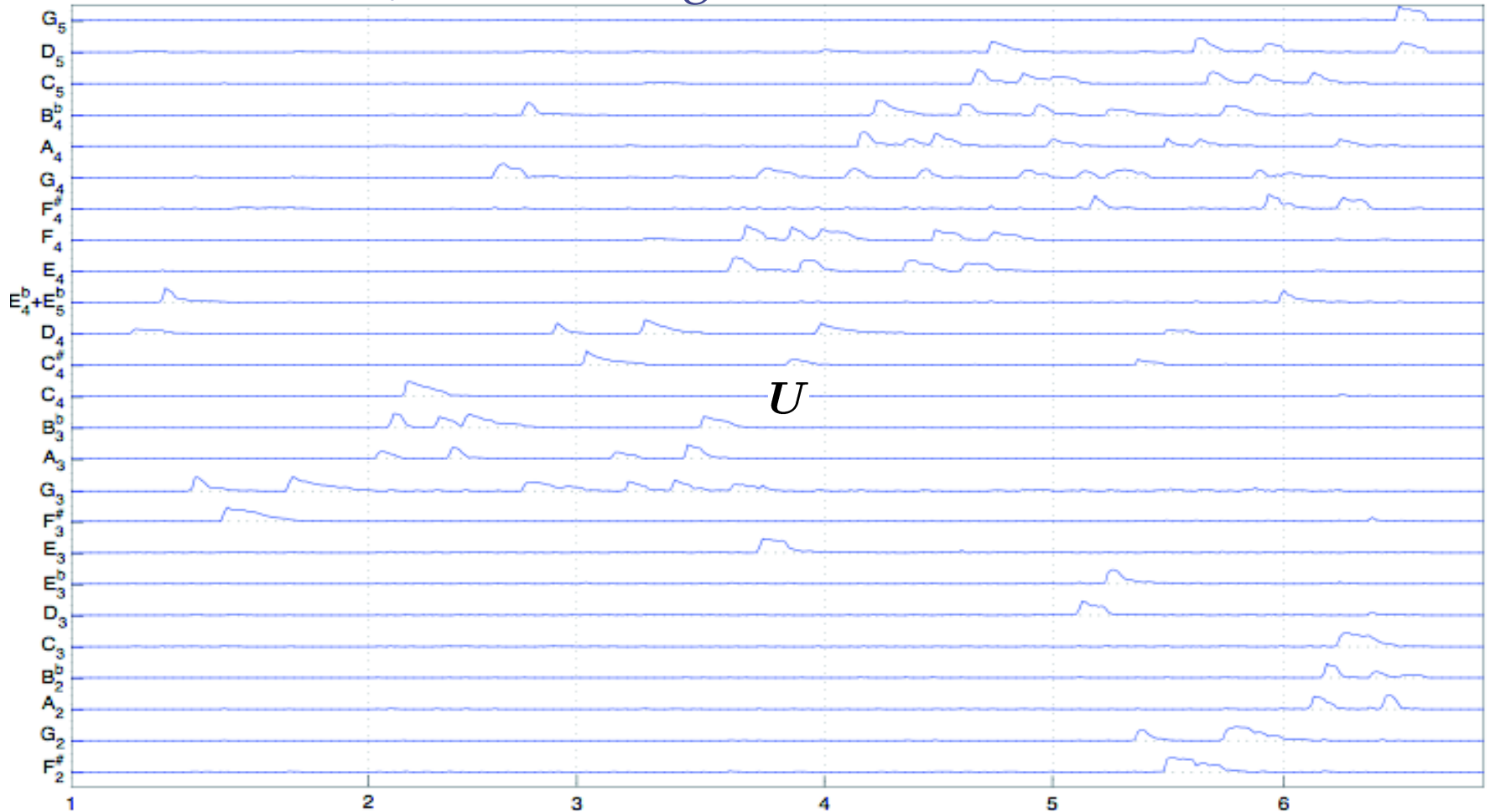


時刻 \rightarrow

何に使えるのか？(1/3)

◆ 自動採譜 [P. Smaragdís et al., 2003]

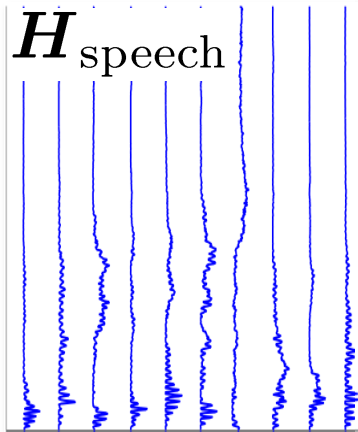
J.S. Bach: Fuge #16 in G minor 🗣️



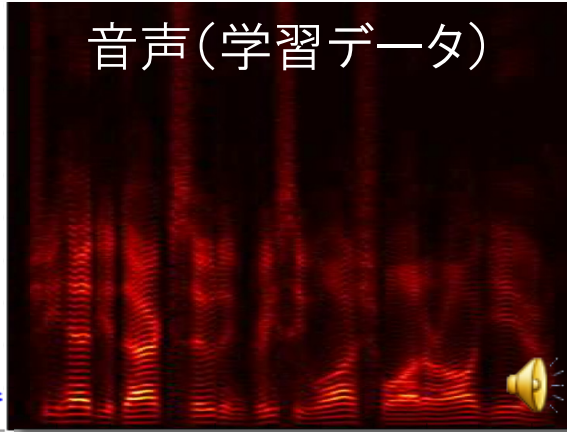
何に使えるのか？(2/3)

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

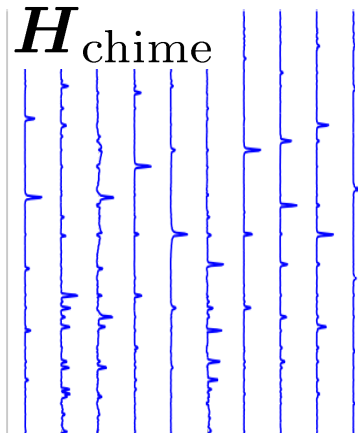
H_{speech}



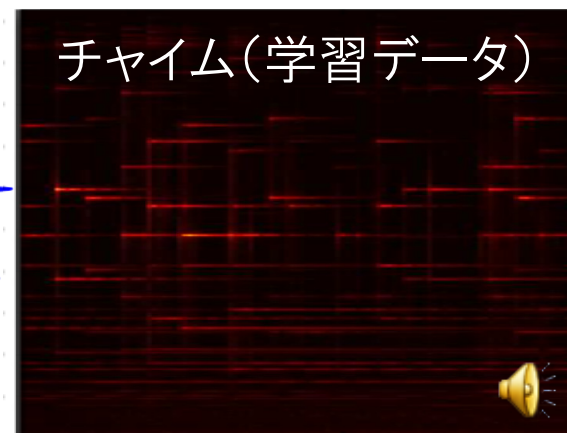
音声(学習データ)



H_{chime}

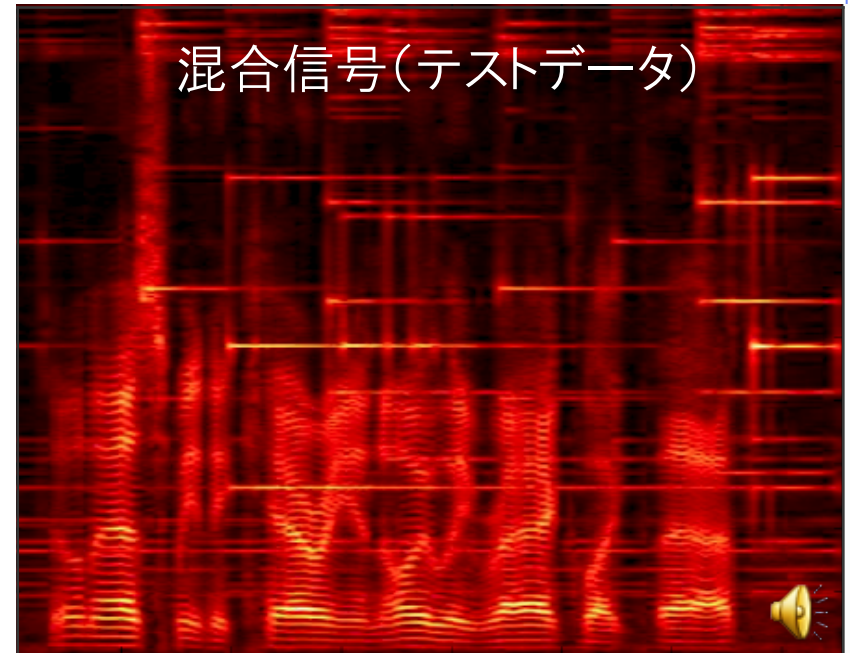


チャイム(学習データ)



Y

混合信号(テストデータ)



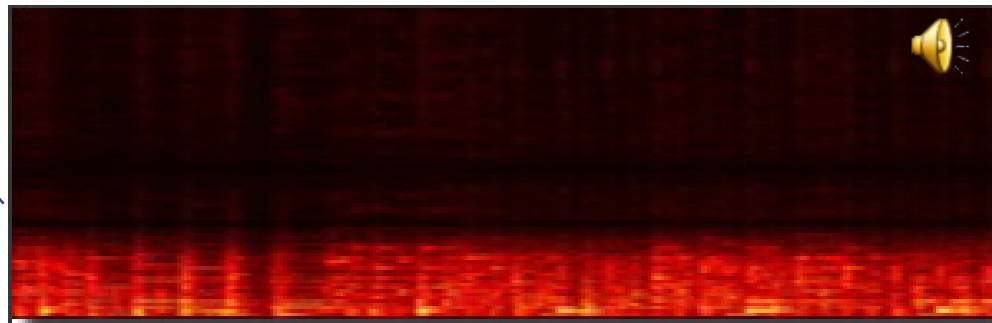
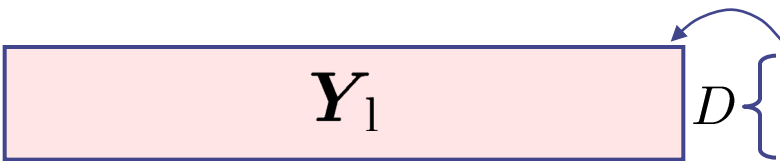
$$Y \simeq \underbrace{\begin{bmatrix} H_{\text{speech}} & H_{\text{chime}} \end{bmatrix}}_{\text{固定}} \begin{bmatrix} U_{\text{speech}} \\ U_{\text{chime}} \end{bmatrix}$$

何に使えるのか？(3/3)

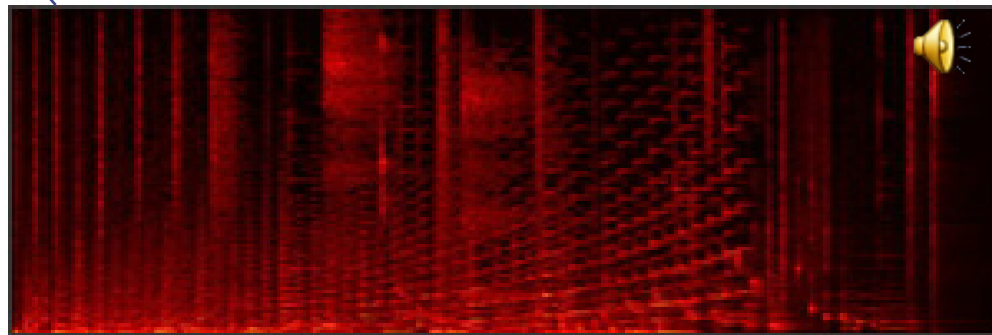
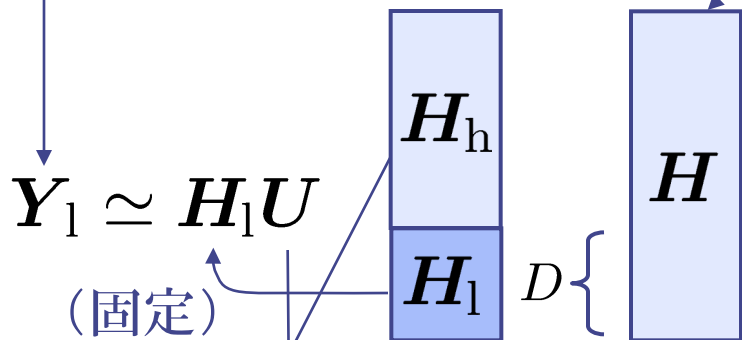
◆音の「超解像」

[P. Smaragdis & B. Raj, 2007]

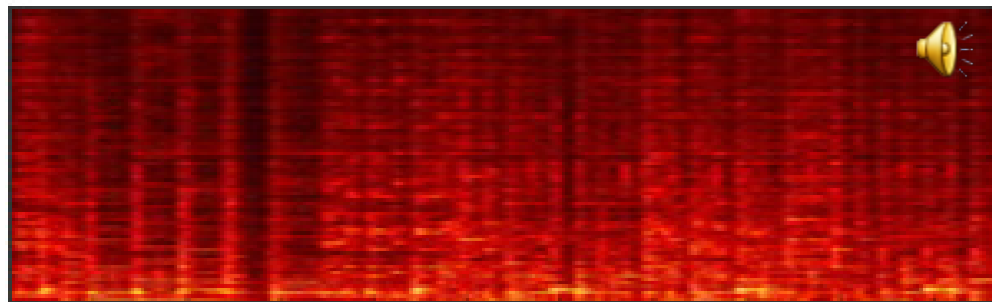
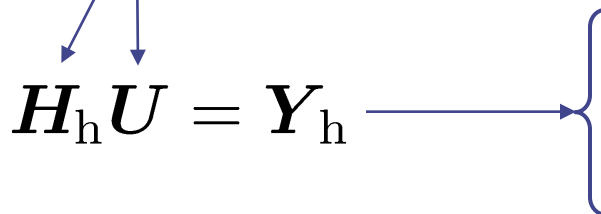
低サンプリングレートの信号



高サンプリングレート信号の学習データ



高周波帯域が復元された信号



NMFの基本問題

- ◆ T 個の観測データ(非負値ベクトル) $\mathbf{y}_1, \dots, \mathbf{y}_T$
- ◆ K 個の非負値基底ベクトル $\mathbf{h}_1, \dots, \mathbf{h}_K$ 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

$$\mathbf{y}_t \simeq \sum_{k=1}^K \mathbf{h}_k U_{k,t}, \quad U_{k,t} \geq 0 \quad (t = 1, \dots, T)$$

$$\left. \begin{array}{l} \mathbf{Y} := (\mathbf{y}_1, \dots, \mathbf{y}_T) \\ \mathbf{H} := (\mathbf{h}_1, \dots, \mathbf{h}_K) \\ \mathbf{U} := (U_{k,t})_{K \times T} \end{array} \right\} \mathbf{Y} \simeq \mathbf{H}\mathbf{U} \quad H_{\omega,k} \geq 0, U_{k,t} \geq 0$$

\mathbf{Y} と $\mathbf{H}\mathbf{U}$ の乖離度を表す規準

$$\begin{array}{ll} \text{minimize} & \mathcal{D}(\mathbf{Y} | \mathbf{H}\mathbf{U}) \\ \text{subject to} & H_{\omega,k} \geq 0, U_{k,t} \geq 0 \end{array}$$

どうやって求めるのか

◆NMFにおける代表的な最適化規準

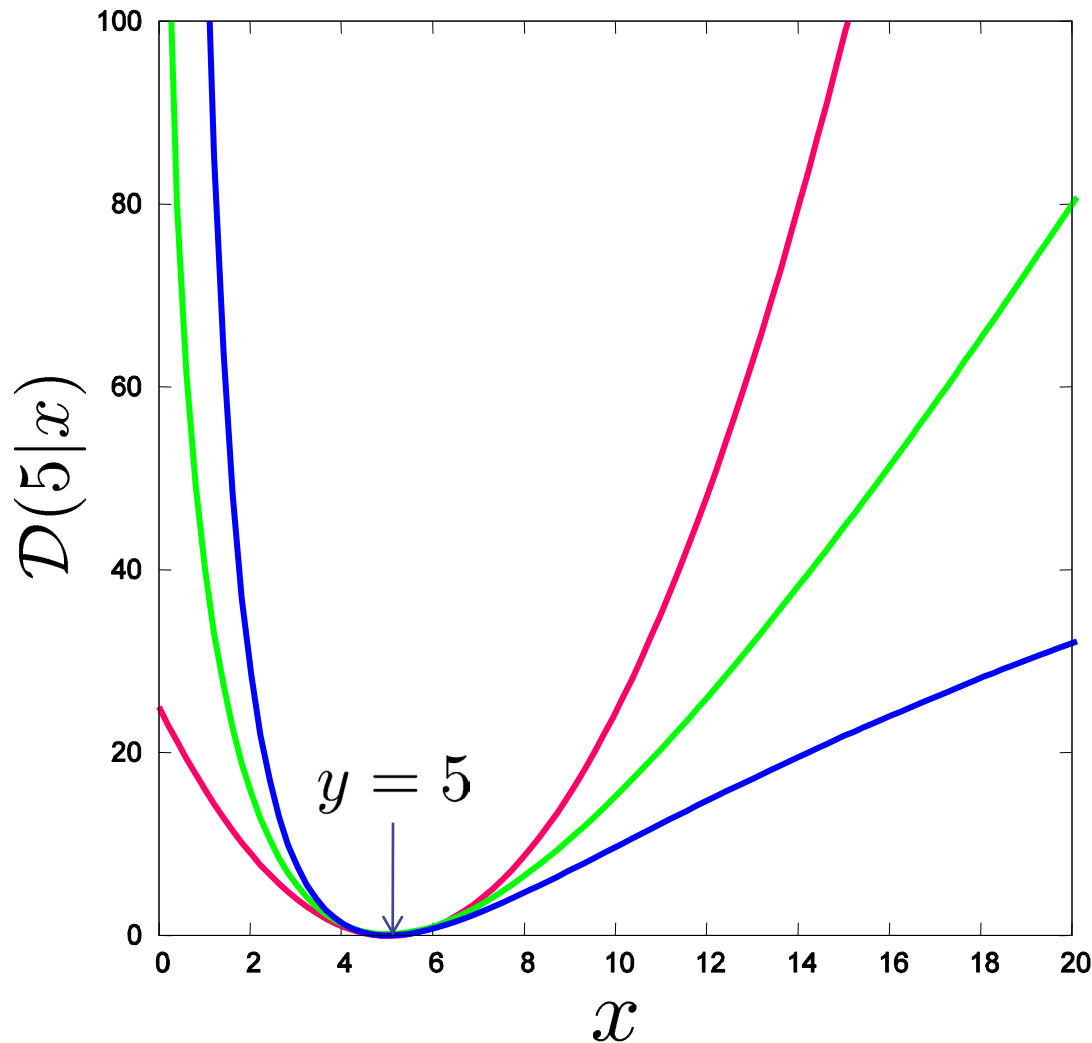
- Frobeniusノルム(二乗誤差)
- lダイバージェンス(一般化KLダイバージェンス)

◆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)

二乗誤差と I ダイバージェンス

- $D(y|x)$: x の y からの近さの度合い



— 二乗誤差

$$D(x|y) = (y - x)^2$$

— Iダイバージェンス

$$D(x|y) = y \log \frac{y}{x} - (y - x)$$

— 板倉齋藤距離

$$D(x|y) = \frac{y}{x} - \log \frac{y}{x} - 1$$

NMFにおける代表的な最適化規準

◆ Frobeniusノルム規準

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

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)

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

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

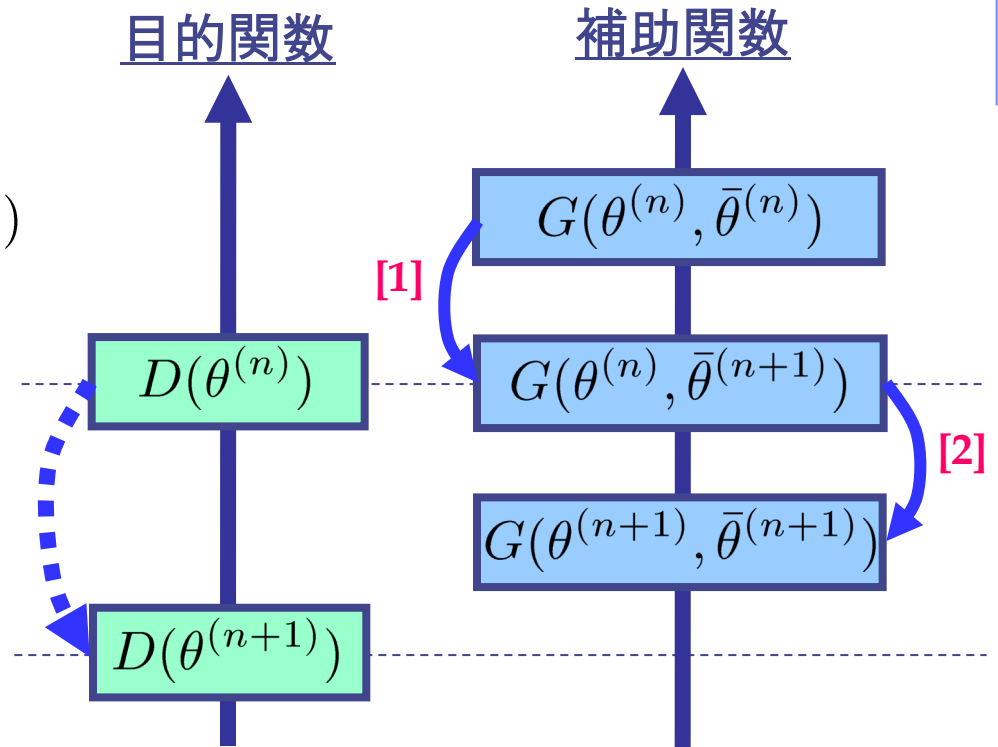
- $D(\theta) = \min_{\bar{\theta}} G(\theta, \bar{\theta})$ を満たす $G(\theta, \bar{\theta})$ を補助関数と定義
- 反復アルゴリズム

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

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

■ 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

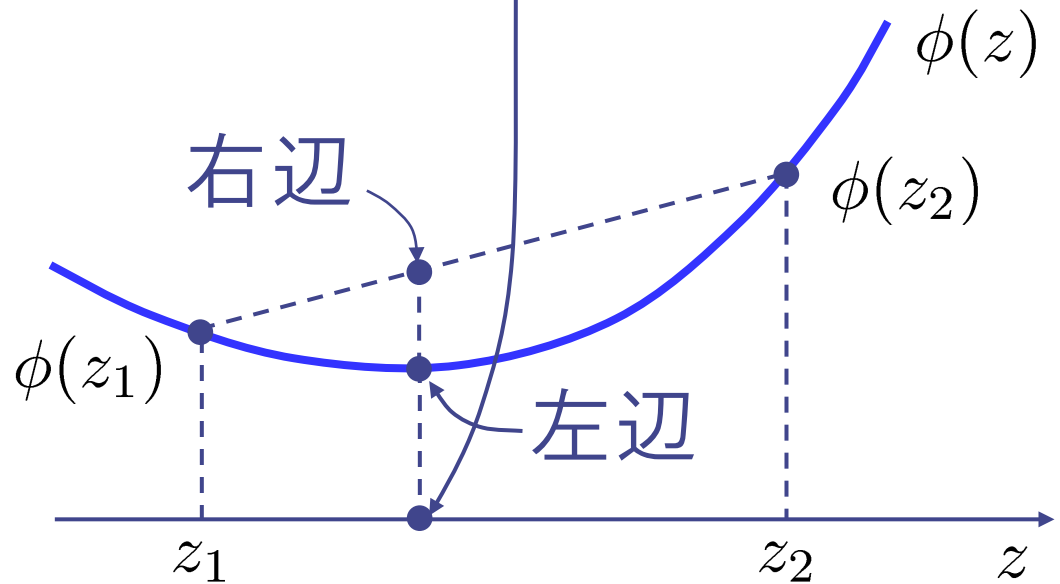
■ $\lambda_i \geq 0, \sum_i \lambda_i = 1$

$$\Rightarrow \phi\left(\underbrace{\sum_i \lambda_i z_i}_{\text{平均}}$$

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

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

- 下線部に対してJensenの不等式を立ててみる

$$\left(\sum_i z_i \right)^2 = \left(\sum_i \lambda_i \frac{z_i}{\lambda_i} \right)^2 \leq \sum_i \lambda_i \left(\frac{z_i}{\lambda_i} \right)^2 = \sum_i \frac{z_i^2}{\lambda_i}$$

適用 $\Rightarrow \left| \sum_k H_{\omega,k} U_{k,t} \right|^2 \leq \sum_k \frac{H_{\omega,k}^2 U_{k,t}^2}{\lambda_{k,\omega,t}}$

- 補助関数が完成

$H_{\omega,k}$ または $U_{k,t}$ ごとの二次関数の和になっている

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

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!

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

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

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

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

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

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

Iダイバージェンス規準のNMFアルゴリズム

- 目的関数

$$\begin{aligned}\mathcal{D} &:= \sum_{\omega,t} \left(Y_{\omega,t} \log \frac{Y_{\omega,t}}{\sum_k H_{\omega,k} U_{k,t}} - Y_{\omega,t} + \sum_k H_{\omega,k} U_{k,t} \right) \\ &= \sum_{\omega,t} \left(Y_{\omega,t} \log Y_{\omega,t} - Y_{\omega,t} \log \sum_k H_{\omega,k} U_{k,t} - Y_{\omega,t} + \sum_k H_{\omega,k} U_{k,t} \right)\end{aligned}$$

- 下線部に対してJensenの不等式を立ててみる

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

適用 $\rightarrow -\log \sum_k H_{\omega,k} U_{k,t} \leq -\sum_k \lambda_{k,\omega,t} \log \frac{H_{\omega,k} U_{k,t}}{\lambda_{k,\omega,t}}$

- 補助関数が完成

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

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!

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

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

$$[2] H_{\omega,k} \leftarrow \operatorname{argmin}_{H_{\omega,k}} G = \frac{\sum_t Y_{\omega,t} \lambda_{k,\omega,t}}{\sum_t U_{k,t}} \quad H_{\omega,k} \leftarrow H_{\omega,k} \frac{\sum_t \frac{Y_{\omega,t} U_{k,t}}{\sum_{k'} H_{\omega,k'} U_{k',t}}}{\sum_t U_{k,t}}$$

$$U_{k,t} \leftarrow \operatorname{argmin}_{U_{k,t}} G = \frac{\sum_{\omega} Y_{\omega,t} \lambda_{k,\omega,t}}{\sum_{\omega} H_{\omega,k}} \quad U_{k,t} \leftarrow U_{k,t} \frac{\sum_{\omega} \frac{Y_{\omega,t} H_{\omega,k}}{\sum_{k'} H_{\omega,k'} U_{k',t}}}{\sum_{\omega} H_{\omega,k}}$$

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

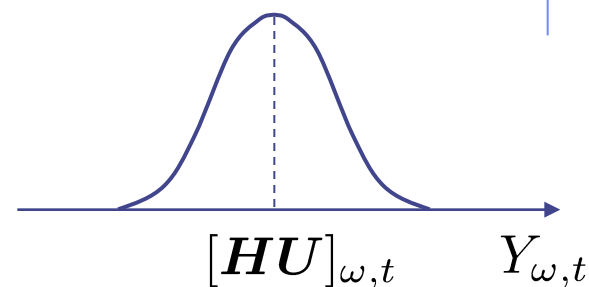
■ Frobeniusノルム規準

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

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

正規分布

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



■ 1ダイバージェンス規準

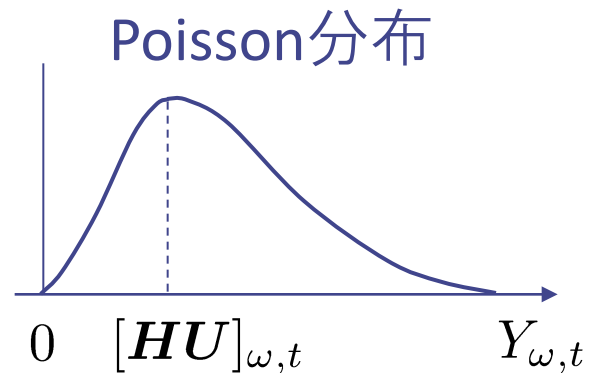
$$Y_{\omega,t} \sim \text{Poisson}([HU]_{\omega,t})$$

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

Poisson分布

$$\rightarrow -\log P(\mathbf{Y}|\mathbf{H}\mathbf{U}) \stackrel{c}{=} \sum_{\omega,t} (-Y_{\omega,t} \log [HU]_{\omega,t} + [HU]_{\omega,t})$$

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



目次

1. 非負値行列因子分解(NMF)

- 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

2. 複素NMF [Kameoka2008]

- NMFによるスペクトログラム分解法が根源的に抱える問題点
- これを解決する新モデル

3. NMFの改良・拡張モデル

- スパースNMF, NMFD, NMF2D, ソースフィルタNMF, アクティベーションの連続性規準入りNMF, ハーモニックNMF, 板倉齋藤距離規準NMF, 複素NMF, 状態遷移NMF, ノンパラメトリックベイズNMF, etc...

目次

1. 非負値行列因子分解(NMF)

- 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

2. 複素NMF [Kameoka2008]

- NMFによるスペクトログラム分解法が根源的に抱える問題点
- これを解決する新モデル

3. NMFの改良・拡張モデル

- スパースNMF, NMFD, NMF2D, ソースフィルタNMF, アクティベーションの連続性規準入りNMF, ハーモニックNMF, 板倉齋藤距離規準NMF, 複素NMF, 状態遷移NMF, ノンパラメトリックベイズNMF, etc...

NMFによるスペクトログラム分解法の根源的な問題点

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

振幅スペクトル
の世界

波形の世界
(豊富な情報)

≡
複素スペクトル
の世界

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...線形な変換
(時間周波数分解)

$Y_{\omega,t} = \int W(u)y(u+t)e^{j\omega u}du \in \mathbb{C}$ \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 $|Y_{\omega,t}| = \left| \sum_i X_{\omega,t}^i \right|$

$\mathbf{Y} \simeq \mathbf{H}\mathbf{U}$ ($\mathbf{H} \succeq 0, \mathbf{U} \succeq 0$)

$\neq \sum_i |X_{\omega,t}^i|$ (振幅スペクトル
同士は非加法的)

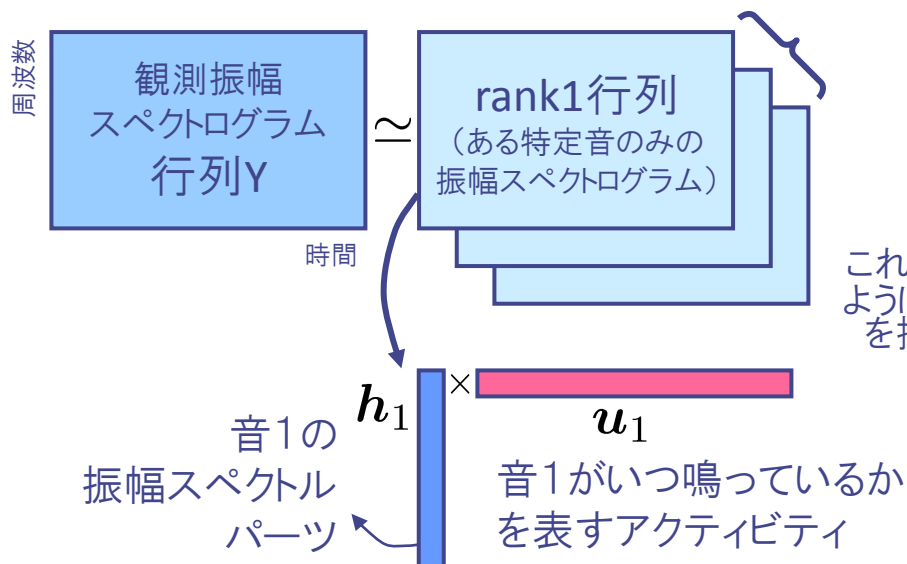
NMFモデル

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

新モデル「複素NMF」 [Kameoka2008]

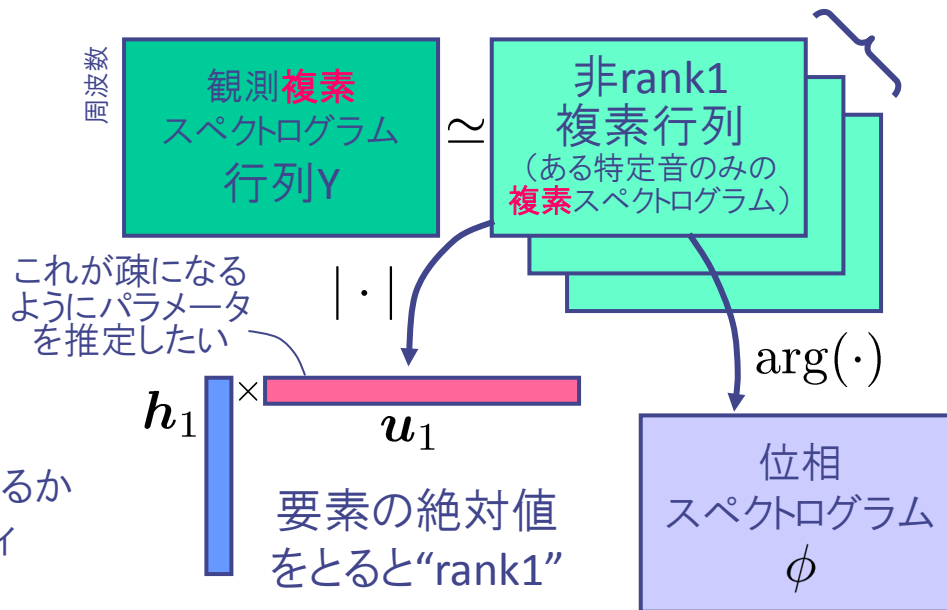
従来のNMFの捉え方

何種類かの音からなると仮定



複素NMFの考え方

加法性が成り立つ！



◆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

$$\begin{aligned}
 Y_{\omega,t} &= \sum_i X_{\omega,t}^i \\
 &= \sum_i |X_{\omega,t}^i| e^{j\phi_{\omega,t}^i} = \sum_i \underline{H_{\omega}^i U_t^i} e^{j\phi_{\omega,t}^i}
 \end{aligned}$$

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

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)

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 $\mathcal{U} \equiv \{U_t^i\}$

音響信号モデル

$$F_{\omega,t} = \sum_{i=1}^I H_{\omega}^i U_t^i e^{j\phi_{\omega,t}^i}$$

◆ 最適化問題

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

一般化正規分布

$$\begin{aligned} \text{minimize} \quad & f(\mathcal{H}, \mathcal{U}, \phi) := \sum_{\omega,t} |Y_{\omega,t} - F_{\omega,t}|^2 + 2\lambda \sum_{i,t} |U_t^i|^p \\ \text{subject to} \quad & \sum_{\omega} H_{\omega}^i = 1 \quad (i = 1, \dots, I) \end{aligned}$$

スパース正則化項

複素NMFアルゴリズム

◆ 補助関数法

$$\begin{aligned}
 f(\mathcal{H}, \mathcal{U}, \phi) &= \sum_{\omega, t} |Y_{\omega, t} - F_{\omega, t}|^2 + 2\lambda \sum_{i, t} |U_t^i|^p \\
 &\leq \sum_{\omega, t} \sum_i \frac{|\bar{Y}_{\omega, t}^i - H_{\omega}^i U_t^i e^{j\phi_{\omega, t}^i}|^2}{\beta_{\omega, t}^i} + \lambda \sum_{i, t} (p|\bar{U}_t^i|^{p-2} U_t^{i2} + (2-p)|\bar{U}_t^i|^p) \\
 &\text{subject to } \sum_i \bar{Y}_{\omega, t}^i = Y_{\omega, t}
 \end{aligned}$$

$0 < \beta_{\omega, t}^i < 1, \sum_i \beta_{\omega, t}^i = 1$
 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)	$\bar{Y}_{\omega, t}^{i(\ell+1)} \leftarrow H_{\omega}^{i(\ell)} U_t^{i(\ell)} e^{j\phi_{\omega, t}^{i(\ell)}} + \beta_{\omega, t}^i (Y_{\omega, t} - F_{\omega, t}^{(\ell)})$	$f(\mathcal{H}^{(\ell)}, \mathcal{U}^{(\ell)}, \phi^{(\ell)})$ $= g(\mathcal{H}^{(\ell)}, \mathcal{U}^{(\ell)}, \phi^{(\ell)}, \bar{Y}^{(\ell+1)}, \bar{U}^{(\ell+1)})$ $\geq g(\mathcal{H}^{(\ell+1)}, \mathcal{U}^{(\ell+1)}, \phi^{(\ell+1)}, \bar{Y}^{(\ell+1)}, \bar{U}^{(\ell+1)})$ $\geq f(\mathcal{H}^{(\ell+1)}, \mathcal{U}^{(\ell+1)}, \phi^{(\ell+1)})$
	$\bar{U}_t^{i(\ell+1)} \leftarrow U_t^{i(\ell)}$	
Step 2)	$H_{\omega}^{i(\ell+1)} \leftarrow \frac{\sum_t U_t^{i(\ell)} \operatorname{Re}[\bar{Y}_{\omega, t}^{i(\ell+1)*} e^{j\phi_{\omega, t}^{i(\ell)}}]}{\sum_t U_t^{i(\ell)2} / \beta_{\omega, t}^i}$	
	$U_t^{i(\ell+1)} \leftarrow \frac{\sum_t H_{\omega}^{i(\ell)} \operatorname{Re}[\bar{Y}_{\omega, t}^{i(\ell+1)*} e^{j\phi_{\omega, t}^{i(\ell)}}]}{\sum_t U_t^{i(\ell)2} / \beta_{\omega, t}^i + \lambda p \bar{U}_t^{i(\ell+1)} ^{p-2}}$	
	$e^{j\phi_{\omega, t}^{i(\ell+1)}} \leftarrow \bar{Y}_{\omega, t}^{i(\ell+1)} / \bar{Y}_{\omega, t}^{i(\ell+1)} $	

必ず正值

必ず正值

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

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

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

[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

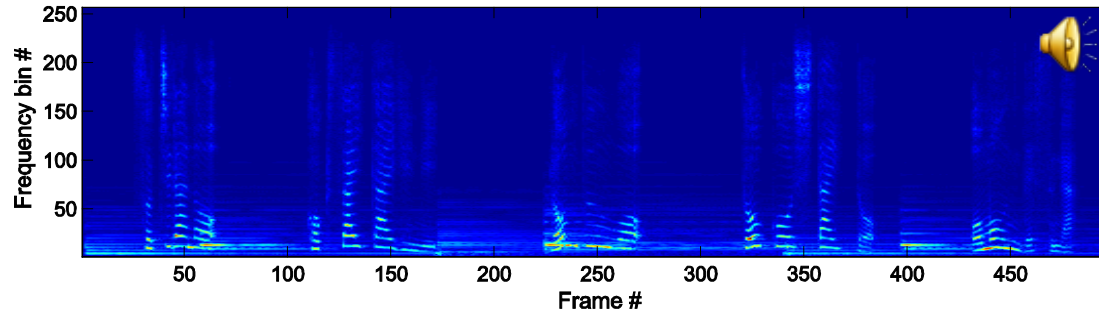
$$H_{\omega}^i \leftarrow H_{\omega}^i \frac{\sum_t U_t^i |Y_{\omega,t}|}{\sum_t U_t^i \sum_k H_{\omega}^k U_t^k}$$

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

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

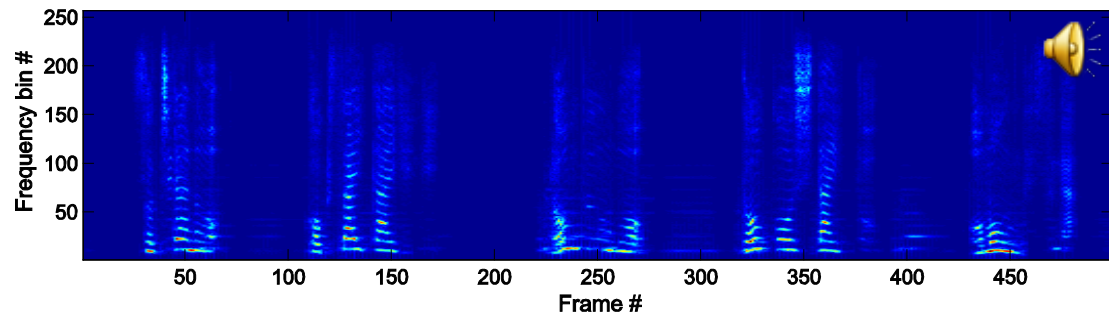
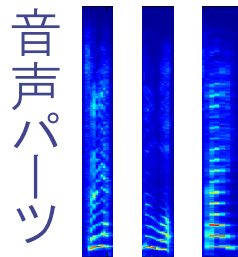
複素NMFによる音声と音楽の混合信号の分離

混合信号(音声+音楽)

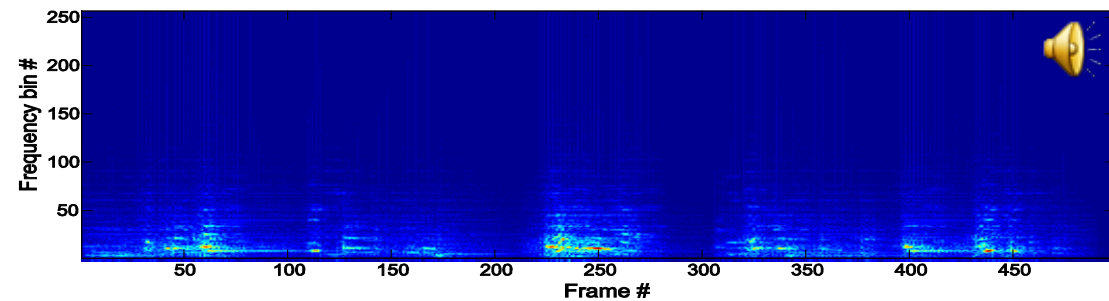
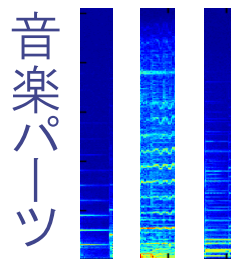


| \rightarrow

抽出音声信号



+



目次

1. 非負値行列因子分解(NMF)

- 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

2. 複素NMF [Kameoka2008]

- NMFによるスペクトログラム分解法が根源的に抱える問題点
- これを解決する新モデル

3. NMFの改良・拡張モデル

- スパースNMF, NMFD, NMF2D, ソースフィルタNMF, アクティベーションの連続性規準入りNMF, ハーモニックNMF, 板倉齋藤距離規準NMF, 複素NMF, 状態遷移NMF, ノンパラメトリックベイズNMF, etc...

目次

1. 非負値行列因子分解(NMF)

- 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

2. 複素NMF [Kameoka2008]

- NMFによるスペクトログラム分解法が根源的に抱える問題点
- これを解決する新モデル

3. NMFの改良・拡張モデル

- スパースNMF, NMFD, NMF2D, ソースフィルタNMF, アクティベーションの連続性規準入りNMF, ハーモニックNMF, 板倉齋藤距離規準NMF, 複素NMF, 状態遷移NMF, ノンパラメトリックベイズNMF, etc...

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

NMFの改良・拡張モデル

2003 2004 2005 2006 2007 2008 2009 2010

Smaragdis

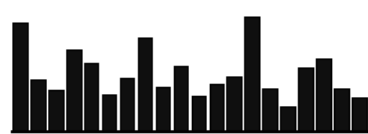
Hoyer

スパースNMF [Hoyer2004]

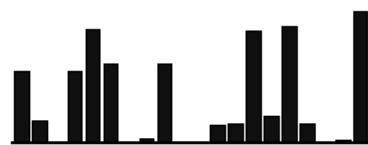
◆ アクティベーションを強制的にスパース化

$$\begin{aligned} & \text{minimize} \quad \sum_{\omega,t} \mathcal{D}\left(Y_{\omega,t} \mid \sum_i H_{\omega,i} U_{i,t}\right) \\ & \text{subject to} \quad \frac{\sqrt{T} - \sum_{k,t} |U_{k,t}| / \sqrt{\sum_{k,t} U_{k,t}^2}}{\sqrt{T} - 1} = S_U \end{aligned}$$

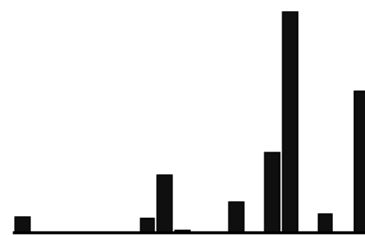
スパース性尺度(1のときスパース, 0のとき非スパース)



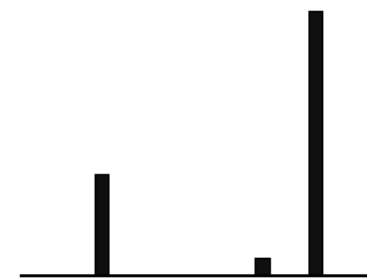
0.1



0.4



0.7



0.9

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Hoyer

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Hoyer

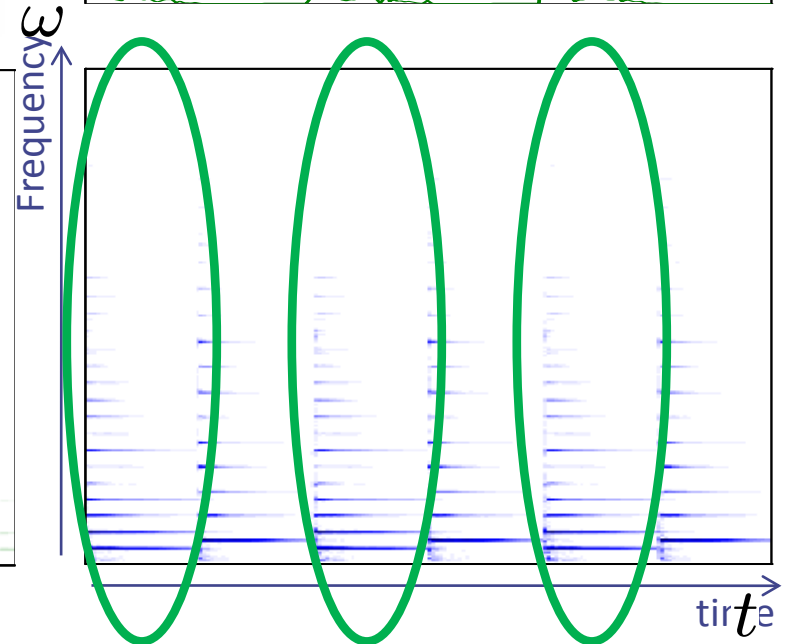
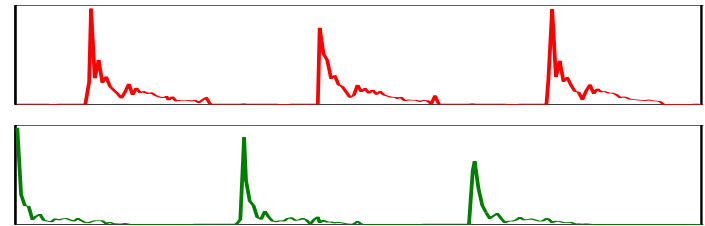
Nonnegative Matrix Factor Deconvolution [Smaragdis2004]

- ◆ スペクトログラム素片パーツを時間方向に畳み込んで
楽音スペクトログラムを表現

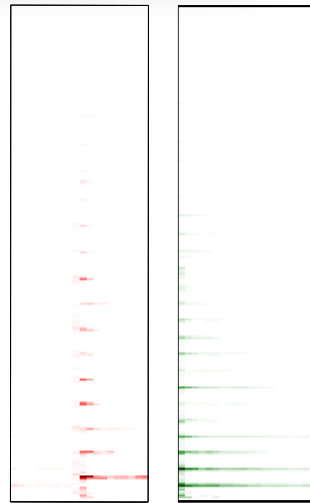
$$Y_{\omega,t} \simeq \sum_{k,\phi} H_{\omega,\phi,k} U_{k,t-\phi}$$

スペクトログラム素片 オンセット

オンセット



スペクトログラム
素片パーツ



NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Hoyer

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Dhillon

Hoyer

Bregmanダイバージェンス規準NMF [Dhillon2005]

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

$$\mathcal{D}_\varphi(y|x) := \varphi(y) - \varphi(x) - \varphi'(x)(y - x) \quad (\varphi \text{ は任意の微分可能な凸関数})$$

⇒ y と x が近いほど小さい

$$\text{◆ minimize } \sum_{\omega,t} \mathcal{D}_\varphi \left(\sum_k H_{\omega,k} U_{k,t} \mid Y_{\omega,t} \right)$$

$$\varphi(z) = z^2 \quad \text{◆ : 二乗誤差誤差}$$

$$\varphi(z) = z \log z \quad \text{◆ : I ダイバージェンス}$$

$$\varphi(z) = -\log z \quad \text{◆ : 板倉齋藤距離}$$

1. $\varphi'(xy) = \varphi'(x)\varphi'(y)/\varphi'(1)$ を満たすクラス ($\varphi(z) = z^2$ はこのタイプ)

$$H_{\omega,k} \leftarrow H_{\omega,k} \cdot \varphi'^{-1} \left(\frac{\varphi'(1) \sum_t \varphi'(Y_{\omega,t}) U_{k,t}}{\sum_t \varphi' \left(\sum_k H_{\omega,k} U_{k,t} \right) U_{k,t}} \right)$$

Uの更新式も同様に
解析的に求まる

2. $\varphi'(xy) = \varphi'(x) + \varphi'(y) + \varphi'(1)$ を満たすクラス ($\varphi(z) = z \log z, -\log z$ はこのタイプ)

$$H_{\omega,k} \leftarrow H_{\omega,k} \cdot \varphi'^{-1} \left(\frac{\sum_t \left[\varphi'(Y_{\omega,t}) - \varphi' \left(\sum_k H_{\omega,k} U_{k,t} \right) - \varphi'(1) \right] U_{k,t}}{\sum_t U_{k,t}} \right)$$

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Dhillon

Hoyer

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

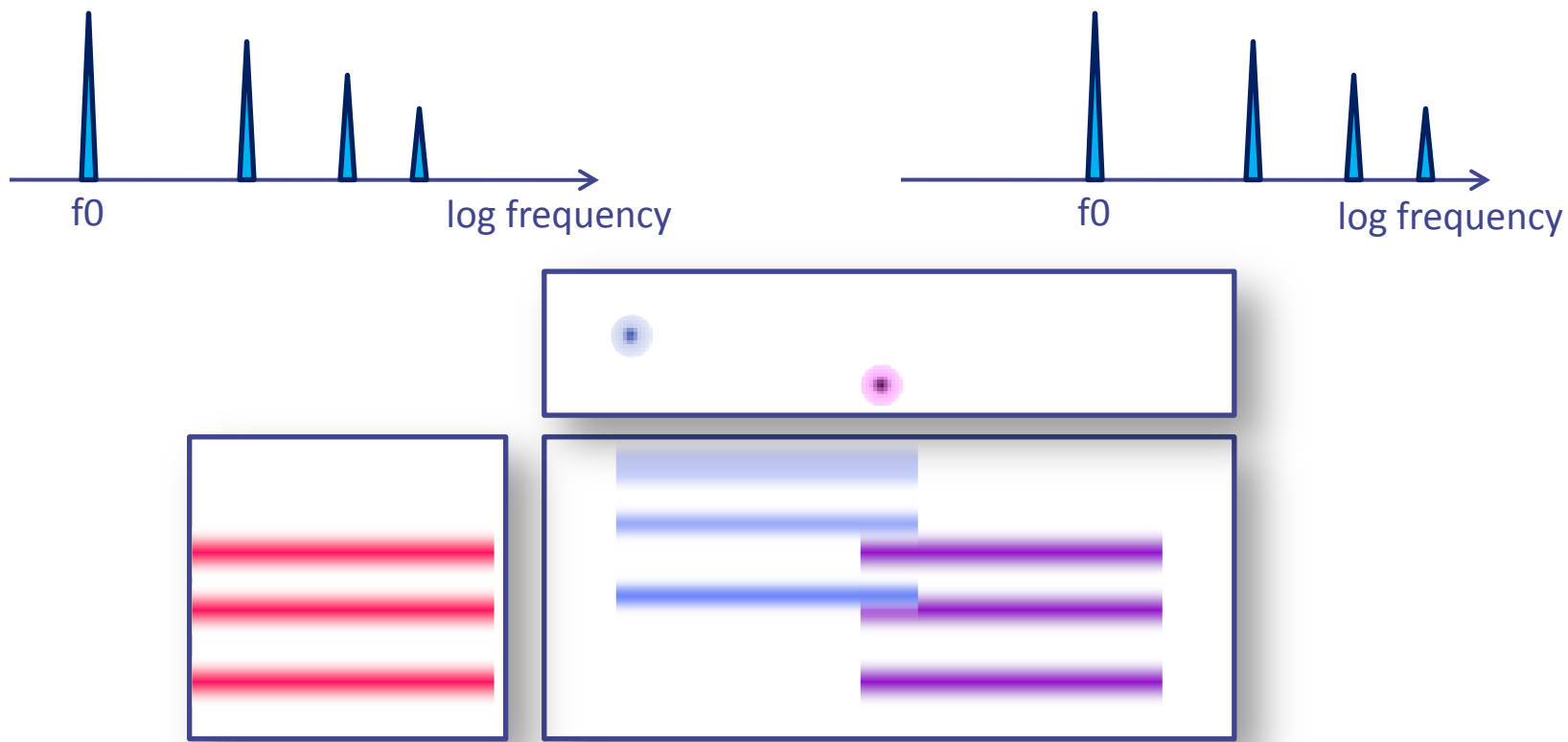
Dhillon

Hoyer

Schmidt

Nonnegative Matrix Factor 2D Deconvolution [Schmidt2006]

- ◆ 楽音の調波構造が対数周波数軸上でシフト不変であると仮定 (ピッチが変わっても倍音パワー比は不変という仮定)
- ◆ スペクトログラムパーツの対数周波数-時間平面での2次元畳み込みで楽音スペクトログラムを表現



NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Dhillon

Hoyer

Schmidt

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Virtanen

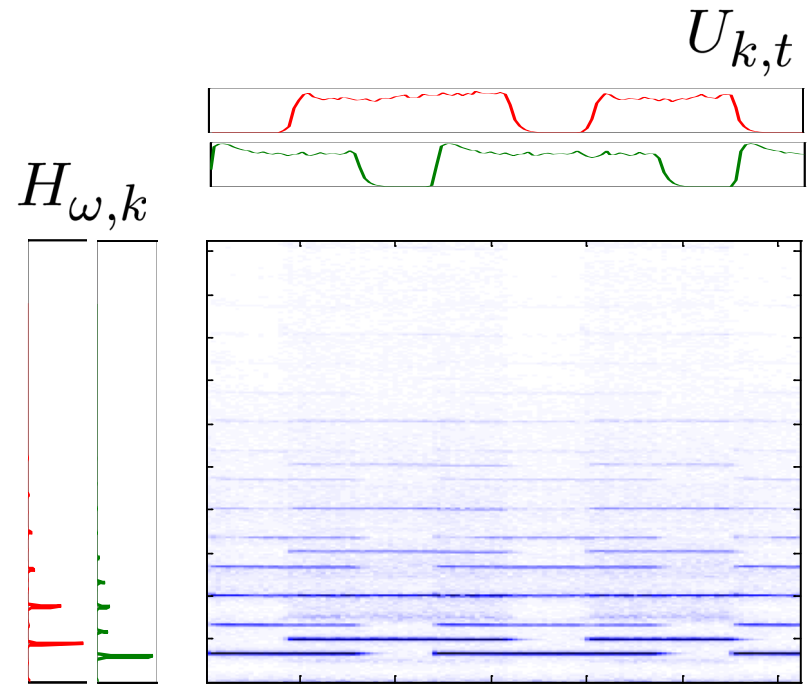
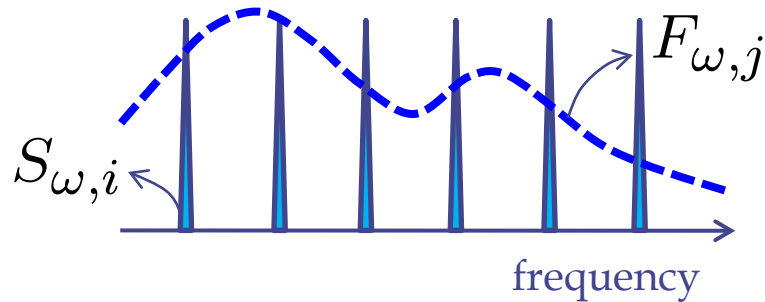
Dhillon

Hoyer

Schmidt

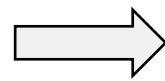
ソースフィルタモデルに基づくNMF [Virtanen2006]

- ◆ $\{i, j\}$ 番目の基底スペクトルを i 番目のソーススペクトルと j 番目のフィルタの積で表現



従来NMF:

$$\sum_k H_{\omega,k} U_{k,t}$$



ソースフィルタNMF:

$$\sum_{i,j} \underline{S_{\omega,i} F_{\omega,j}} U_{i,j,t}$$

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Virtanen

Dhillon

Hoyer

Schmidt

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Virtanen

Dhillon

Hoyer

Kompass

Schmidt

βダイバージェンス規準NMF [Kompas2007]

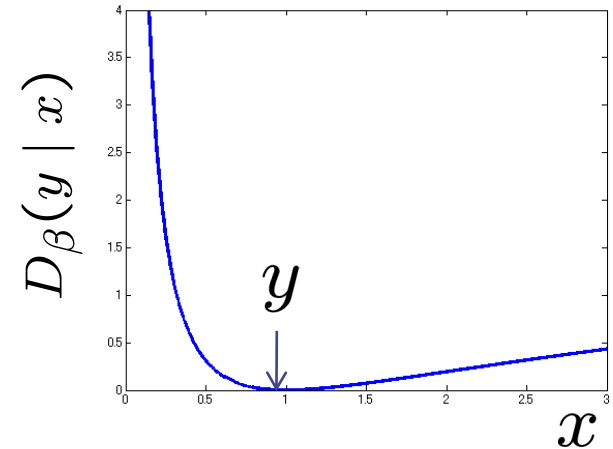
◆βダイバージェンス[Eguchi2001]を局所最小化するアルゴリズム

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

β = 0 : 板倉齋藤距離

β = 1 : Iダイバージェンス

β = 2 : 二乗誤差



$$\mathbf{H} \leftarrow \mathbf{H} \cdot \frac{(\mathbf{Y} \cdot (\mathbf{H}\mathbf{U})^{\beta-2})\mathbf{U}^{\mathbf{T}}}{(\mathbf{H}\mathbf{U})^{\beta-1}\mathbf{U}^{\mathbf{T}}} \quad \mathbf{U} \leftarrow \mathbf{U} \cdot \frac{\mathbf{H}^{\mathbf{T}}(\mathbf{Y} \cdot (\mathbf{H}\mathbf{U})^{\beta-2})}{\mathbf{H}^{\mathbf{T}}(\mathbf{H}\mathbf{U})^{\beta-1}}$$

$1 \leq \beta \leq 2$ の場合しか収束性が保証されない

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Virtanen

Dhillon

Hoyer

Kompass

Schmidt

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Virtanen

Dhillon

Virtanen

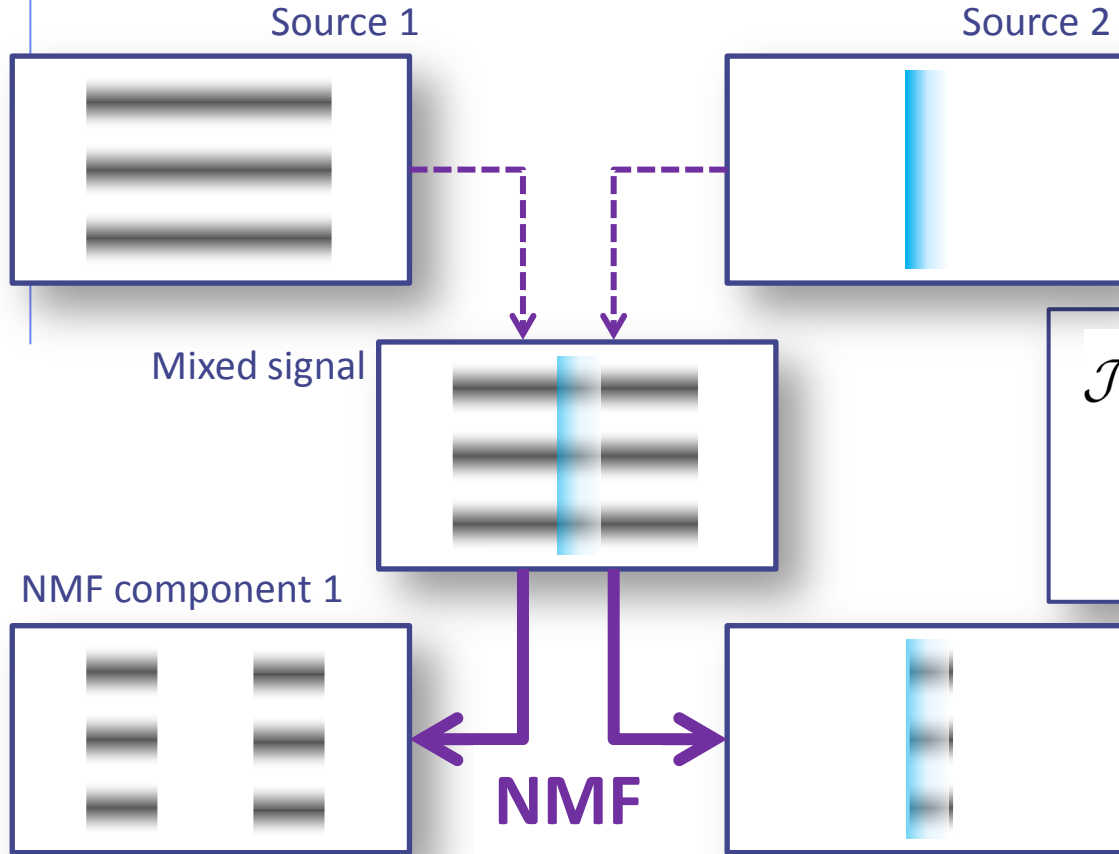
Hoyer

Kompass

Schmidt

アクティベーションの連続性規準入りNMF [Virtanen2007]

◆ 楽器音のパワーエンベロープは滑らかという仮定



そこで、アクティベーションの連続性規準を導入...

$$\mathcal{J} = \sum_{\omega,t} \mathcal{D}\left(Y_{\omega,t} \mid \sum_k H_{\omega,k} U_{k,t}\right) + \sum_{k,t} (U_{k,t} - U_{k,t-1})^2$$

- アクティベーションが滑らかなほど小さい値をとるコスト関数
- この項を含めて最適化を行うことで、左図のようなことが起こりにくなる

音源のストリームが分断されてしまっている！

別音源のストリームが混入してしまっている！

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Virtanen

Dhillon

Virtanen

Hoyer

Kompass

Schmidt

NMFの改良・拡張モデル

2003 2004 2005 2006 2007 2008 2009 2010

Smaragdis

Raczynski

Smaragdis

Virtanen

Dhillon

Virtanen

Hoyer

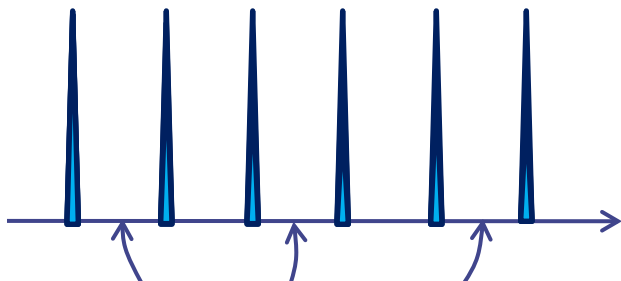
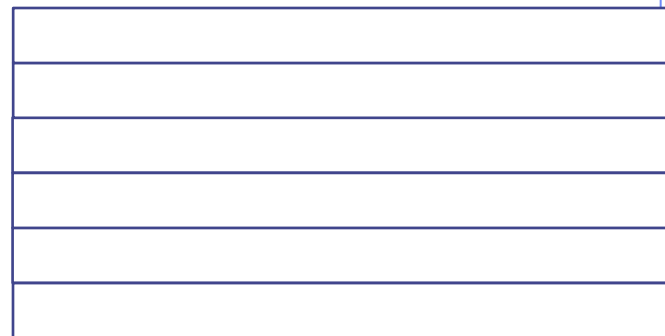
Kompass

Schmidt

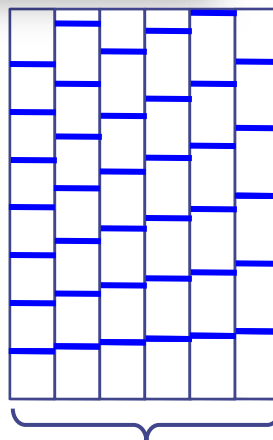
ハーモニックNMF [Raczynski2007]

- ◆ 調波構造形の基底スペクトル
- ◆ アクティベーションの無相関規準

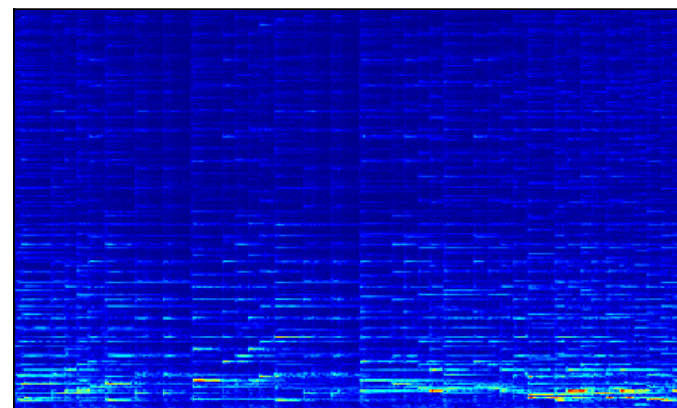
$$\mathcal{J} = \sum_{\omega,t} \mathcal{D} \left(Y_{\omega,t} \left| \sum_k H_{\omega,k} U_{k,t} \right. \right) + \sum_{k \neq j} \sum_t (U_{k,t} - U_{j,t})^2$$



基本周波数・高調波周波数
以外の成分を0に初期設定



音名に対応する分だけ基底スペクトルを用意



NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Dhillon

Hoyer

Virtanen

Schmidt

Raczynski

Virtanen

Kompass

NMFの改良・拡張モデル

2003 2004 2005 2006 2007 2008 2009 2010

Smaragdis

Raczynski

Smaragdis

Virtanen

Vincent

Dhillon

Virtanen

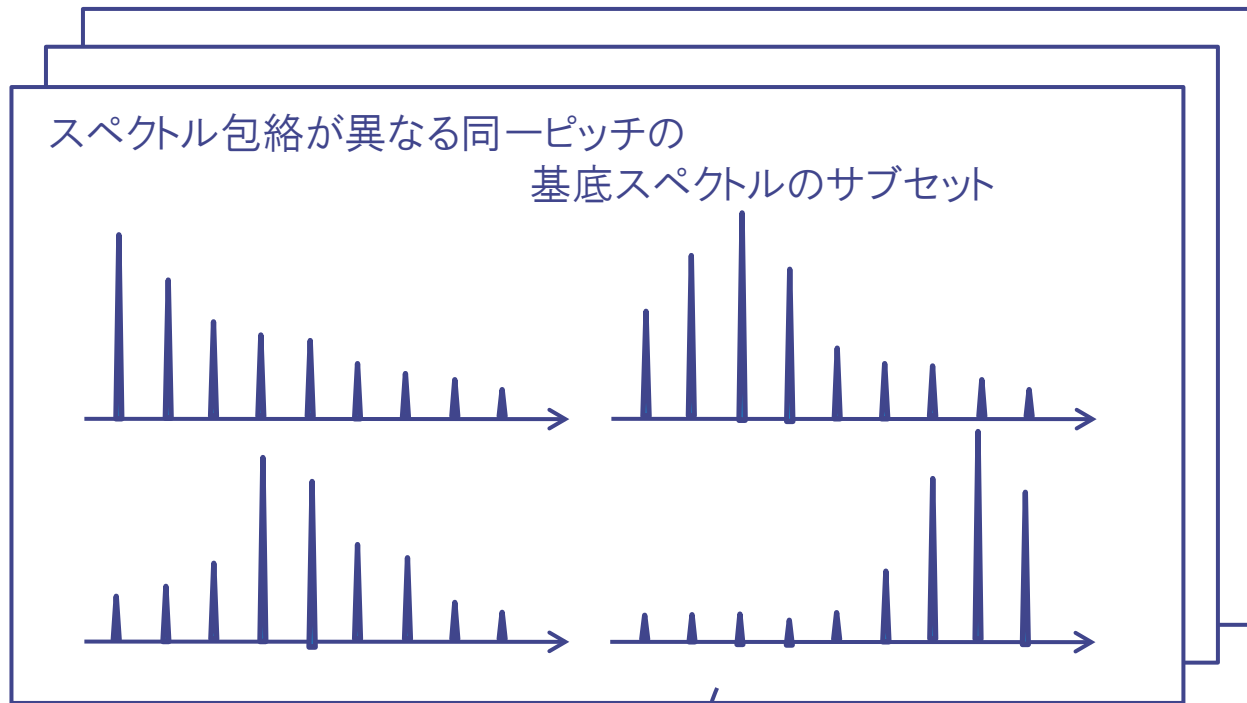
Hoyer

Kompass

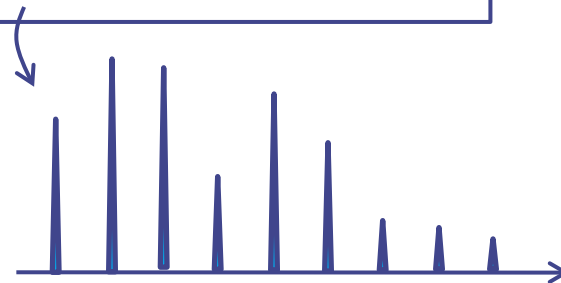
Schmidt

ハーモニックNMF [Vincent2008]

◆調波構造形の基底スペクトル



これらの重ね合わせでさまざまな
スペクトル包絡形状を表現



NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Dhillon

Hoyer

Virtanen

Schmidt

Raczynski

Virtanen

Kompass

Vincent

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Dhillon

Hoyer

Schmidt

Virtanen

Virtanen

Kompass

Raczynski

Vincent

Cemgil

ベイジアンNMF [Cemgil2008]

◆ Iダイバージェンス規準NMFを生成モデルの観点から解釈

$$\begin{aligned} \mathcal{J} &= \sum_{\omega,t} \mathcal{D}\left(Y_{\omega,t} \mid \sum_k H_{\omega,k} U_{k,t}\right) \\ &= \sum_{\omega,t} \left(Y_{\omega,t} \log \frac{Y_{\omega,t}}{\sum_k H_{\omega,k} U_{k,t}} - Y_{\omega,t} + \sum_k H_{\omega,k} U_{k,t} \right) \end{aligned}$$



$$Y_{\omega,t} \sim \text{Poisson}\left(\sum_k H_{\omega,k} U_{k,t}\right)$$

$$\text{Poisson}(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$



$$\begin{aligned} C_{\omega,t,k} &\sim \text{Poisson}(H_{\omega,k} U_{k,t}) \\ Y_{\omega,t} &= \sum_k C_{\omega,t,k} \end{aligned}$$

$$H_{\omega,k} \sim \text{Gamma}(a, b)$$

$$U_{k,t} \sim \text{Gamma}(c, d)$$

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Dhillon

Hoyer

Schmidt

Virtanen

Virtanen

Kompass

Raczynski

Vincent

Cemgil

NMFの改良・拡張モデル

2003

2004

2005

2006

2007

2008

2009

2010

Smaragdis

Smaragdis

Dhillon

Hoyer

Schmidt

Raczynski

Virtanen

Virtanen

Kompass

Vincent

Cemgil

Févotte

板倉齋藤距離規準NMF [Févotte2008]

- ◆ 観測信号をガウス性信号の重ね合わせとしてモデル化
- ◆ パワースペクトル密度がNMF型の構造をもつと仮定

$$C_{\omega,t,k} \sim \mathcal{N}_{\mathbb{C}}(0, H_{\omega,k} U_{k,t})$$

$$Y_{\omega,t} = \sum_k C_{\omega,t,k}$$

ここでの $Y_{\omega,t}$ と $C_{\omega,t,k}$ は
複素スペクトル成分(複素数値)
であることに注意。

$$Y_{\omega,t} \sim \mathcal{N}_{\mathbb{C}}\left(0, \sum_k H_{\omega,k} U_{k,t}\right)$$

$C_{\omega,t,k}$ と $C_{\omega,t,k'}$ の独立性を仮定

$$\begin{aligned} \mathcal{J} &= \sum_{\omega,t} \mathcal{D}\left(Y_{\omega,t} \mid \sum_k H_{\omega,k} U_{k,t}\right) \\ &= \sum_{\omega,t} \left(\frac{Y_{\omega,t}}{\sum_k H_{\omega,k} U_{k,t}} - \log \frac{Y_{\omega,t}}{\sum_k H_{\omega,k} U_{k,t}} - 1 \right) \end{aligned}$$

NMFの改良・拡張モデル

2003 2004 2005 2006 2007 2008 2009 2010

Smaragdis

Raczynski

Smaragdis

Virtanen

Vincent

Dhillon

Virtanen

Hoyer

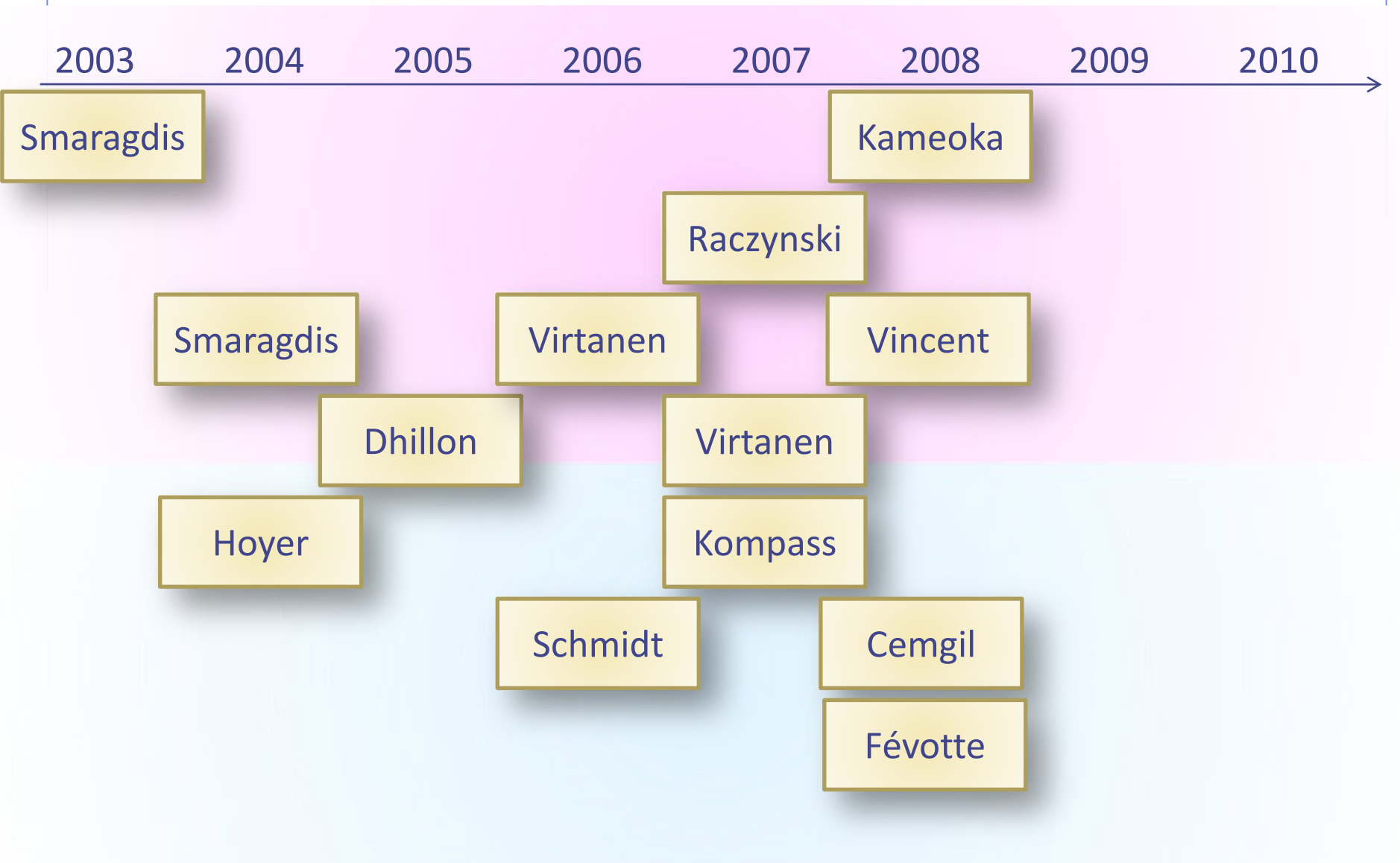
Kompass

Schmidt

Cemgil

Févotte

NMFの改良・拡張モデル



複素NMF [Kameoka2008]

- ◆ パワースペクトルも振幅スペクトルも本来は非加法的
- ◆ 複素スペクトル領域モデルでNMFと同様のスペクトルパーツ獲得機能が実現できないか？

$$\begin{aligned} x(u) &= \sum_k c^{(k)}(u) \\ &\quad \downarrow \text{(STFT)} \\ X_{\omega,t} &= \sum_k C_{\omega,t,k} \\ &\quad \downarrow \text{(magnitude)} \\ |X_{\omega,t}| &= \left| \sum_k C_{\omega,t,k} \right| \\ &\neq \sum_k |C_{\omega,t,k}| \end{aligned}$$

NMFにおけるモデルが厳密には不適切であることを示している

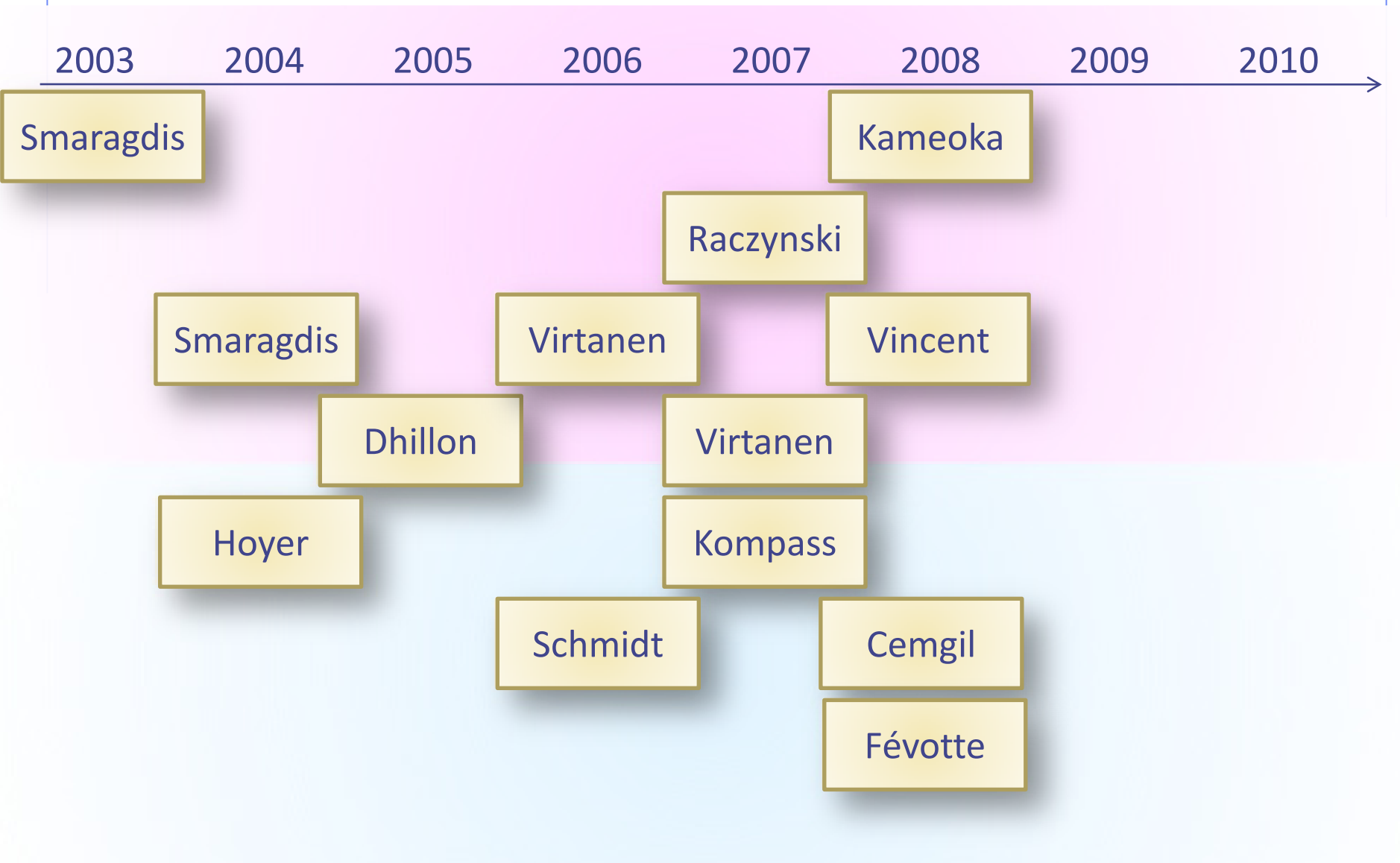
要素信号の複素スペクトログラム $C_{\omega,t,k}$ の振幅成分が“rank1”構造をもつように混合信号をモデル化

$$\begin{aligned} X_{\omega,t} &= \sum_k |C_{\omega,t,k}| e^{j\phi_{\omega,t,k}} \\ &= \sum_k H_{\omega,k} U_{k,t} e^{j\phi_{\omega,t,k}} \end{aligned}$$

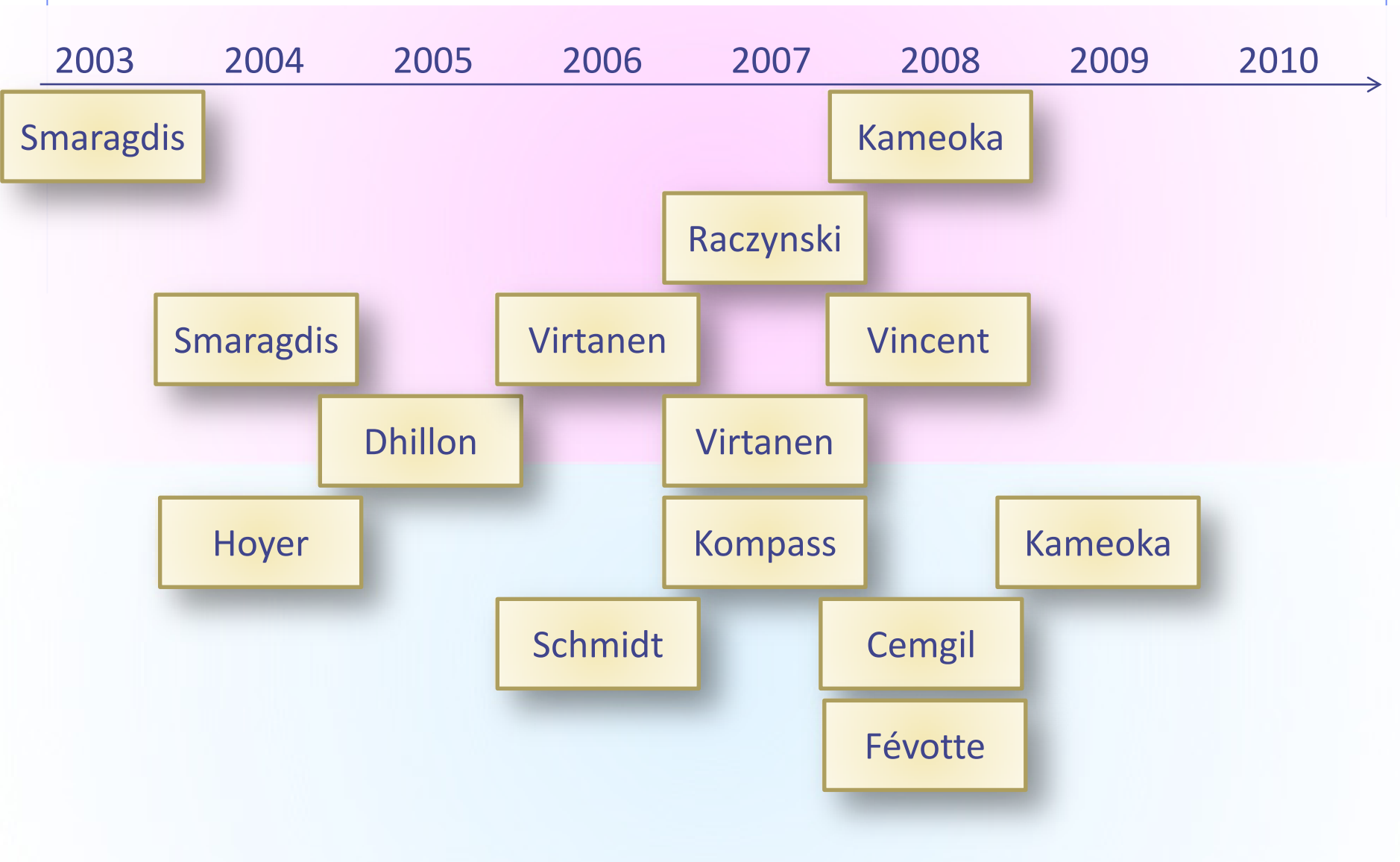
$$\sum_{\omega,t} |Y_{\omega,t} - X_{\omega,t}|^2 \rightarrow \text{minimum}$$

ここでの $Y_{\omega,t}$ と $C_{\omega,t,k}$ も複素スペクトル成分(複素数値)であることに注意。

NMFの改良・拡張モデル



NMFの改良・拡張モデル



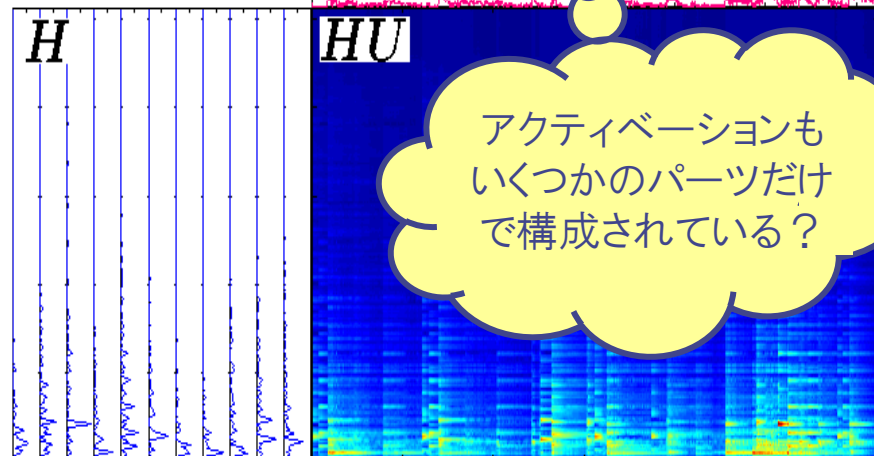
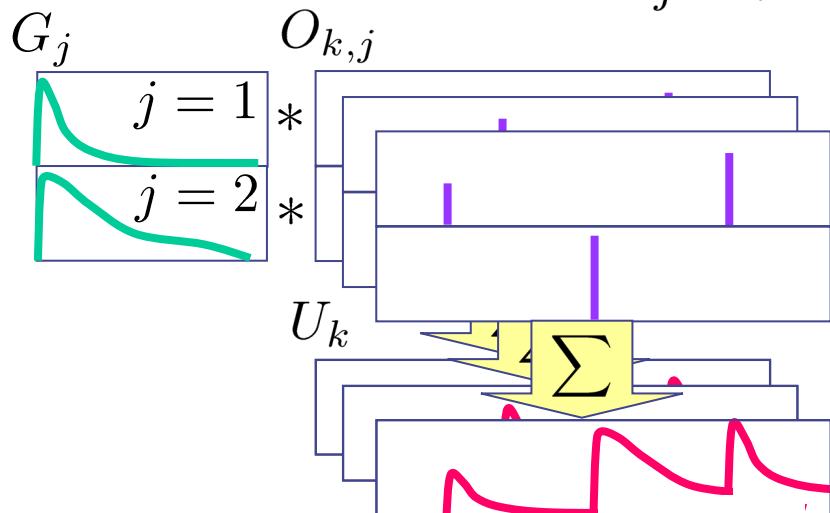
MusicFactorizer [Kameoka2009]

- ◆ 音楽信号は限られた種類の音価の音だけで構成される
→ NMFのアクティベーションもまたパーツベース表現ができる？
- ◆ アクティベーションをエンベロープパーツの畳み込み混合でモデル化

$$Y_{\omega,t} \simeq \sum_k H_{\omega,k} U_{k,t}$$

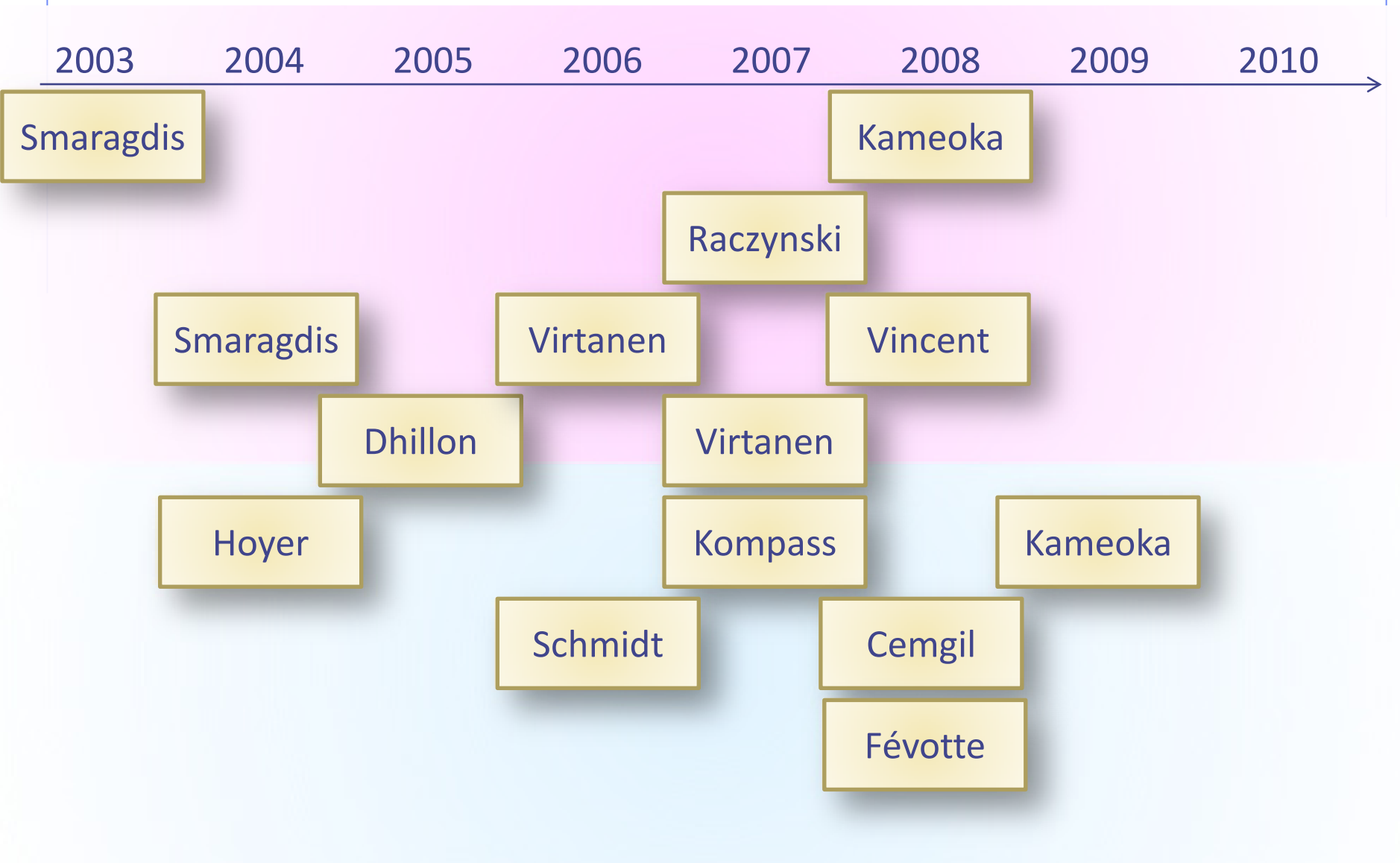
エンベロープパーツ
のアクティベーション
(ノートのオンセット
↑ に対応)

「畳み込み混合」 $U_{k,t} = \sum_j \sum_{\tau} G_{j,\tau} O_{k,j,t-\tau}$

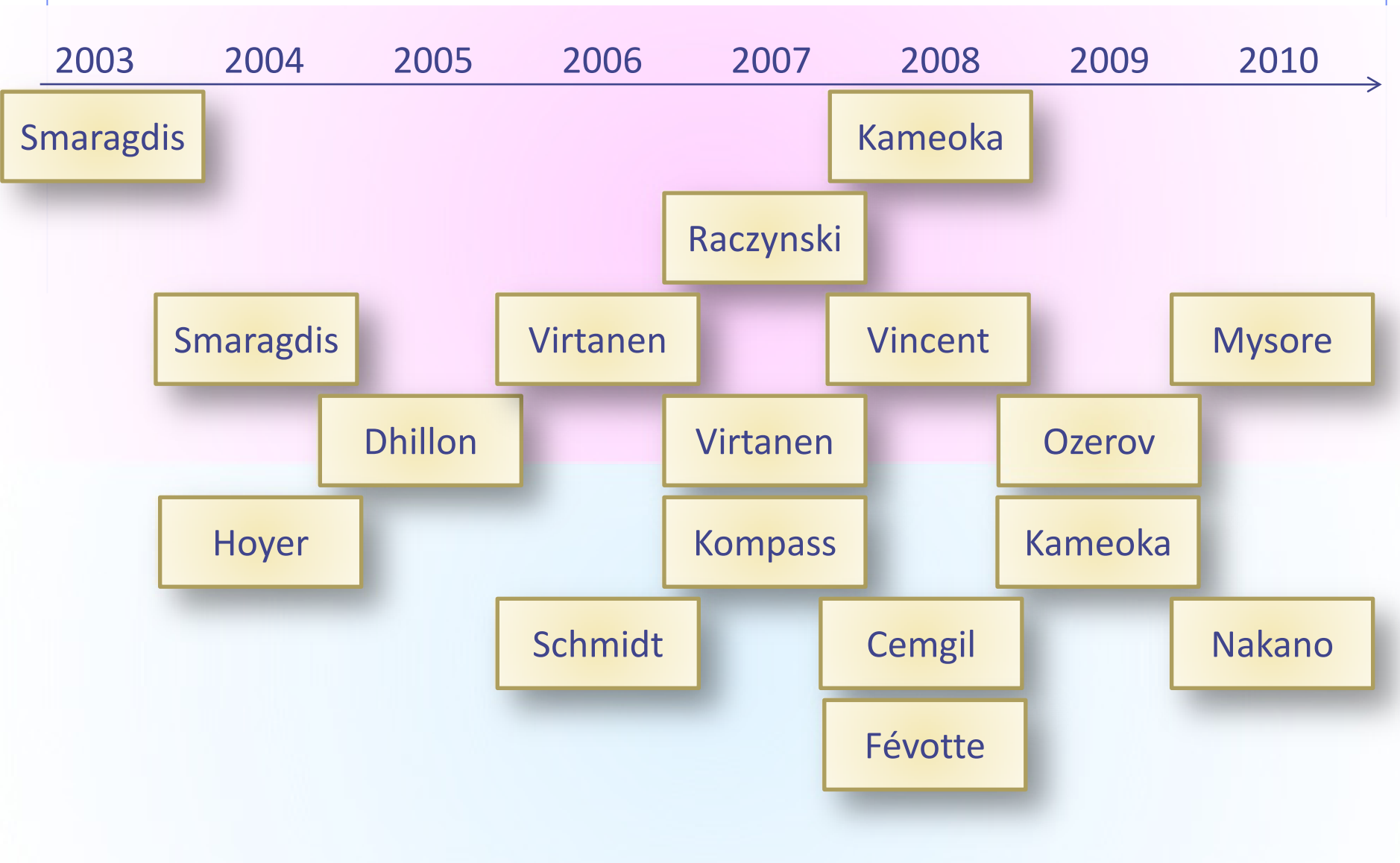


アクティベーションも
いくつかのパーツだけ
で構成されている？

NMFの改良・拡張モデル



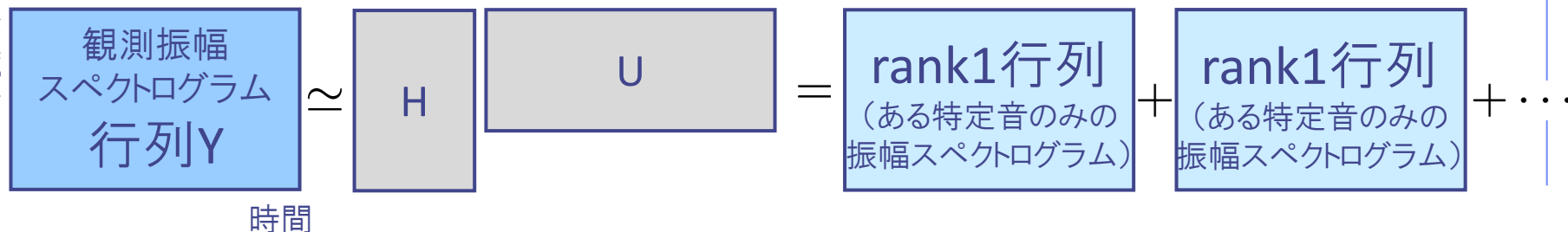
NMFの改良・拡張モデル



状態遷移NMF [Ozerov2009,Nakano2010,Mysore2010]

◆ NMF: スペクトログラムをrank1行列の和に分解していることに相当

周波数



- 楽音のスペクトログラムは実際にはrank1にならない
(ピアノ音にattack/decay/sustainなどの複数状態があるように,
通常, 楽音スペクトルの形は発音中に多様に時間変化する)

◆ 各楽音スペクトログラムをrank1行列でモデル化する代わりに
非負値を出力する隠れマルコフモデル(HMM)でモデル化

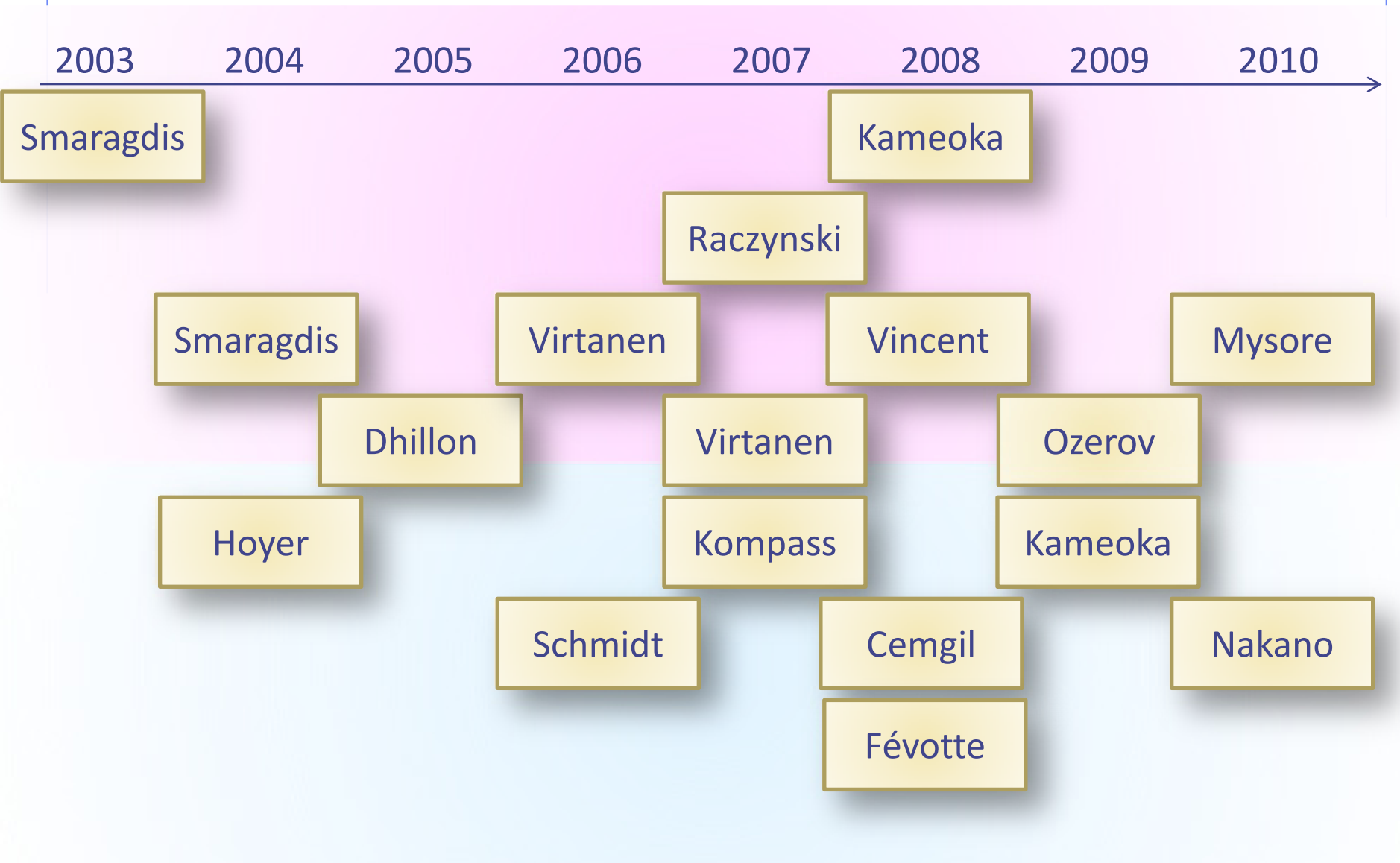
→ Factorial HMM [Ghahramani1997]と同形のモデルになる

A. Ozerov, C. Févotte, M. Charbit, "Factorial scaled hidden Markov model for polyphonic audio representation and source separation," in *Proc. WASPAA'09*, 2009.

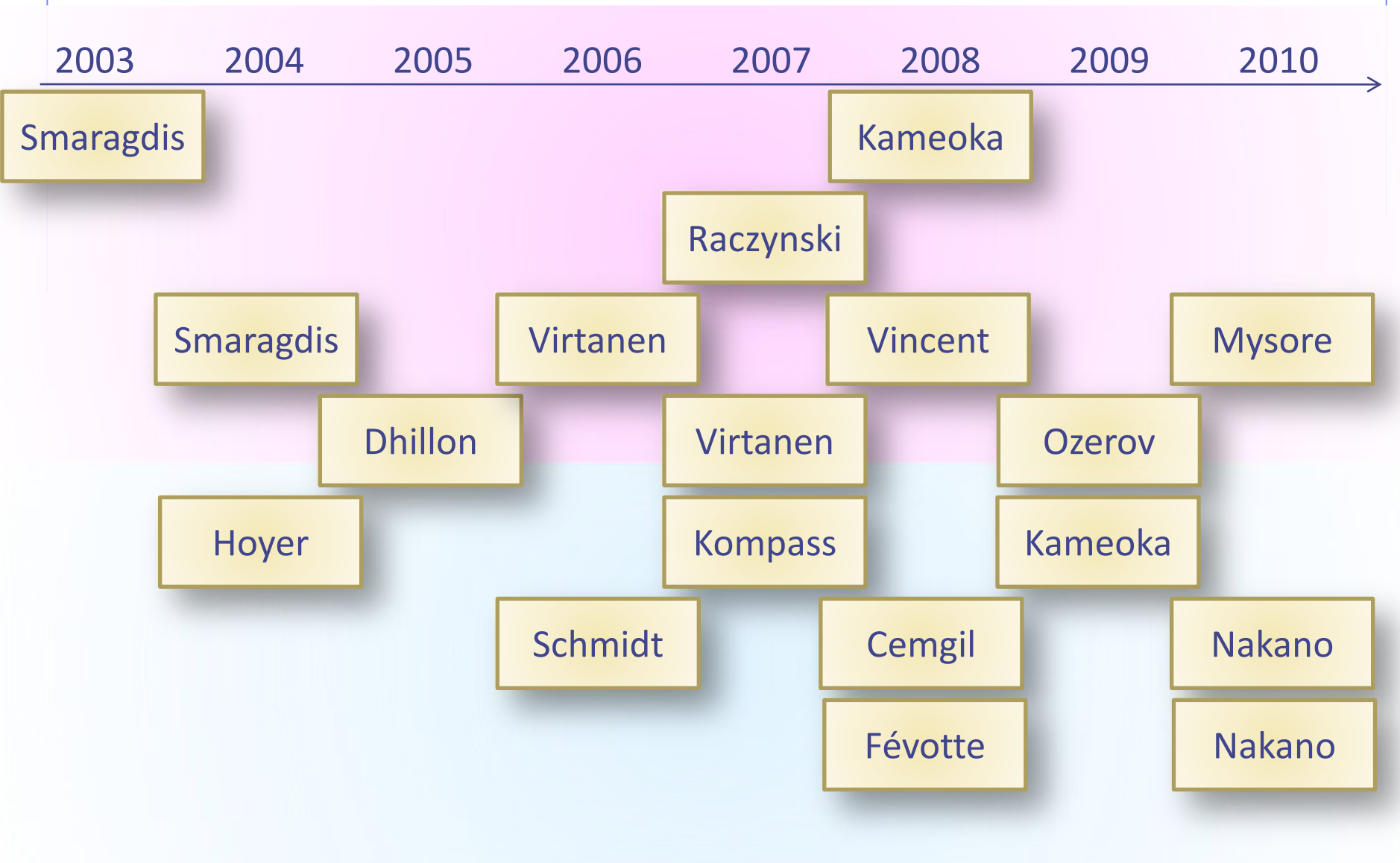
中野, 北野, ルルー, 亀岡, 小野, 嵯峨山, "可変基底NMFに基づく音楽音響信号の解析," 情報処理学会研究報告, 2010-MUS-84-10, Feb. 2010.

G. J. Mysore, P. Smaragdis, B. Raj, "Non-negative hidden Markov modeling of audio with application to source separation," in *Proc. LVA/ICA'2010*. pp.140-148, 2010.

NMFの改良・拡張モデル



NMFの改良・拡張モデル



βダイバージェンス規準NMF [Nakano2010]

◆すべてのβで収束性が保証された乗法更新アルゴリズム

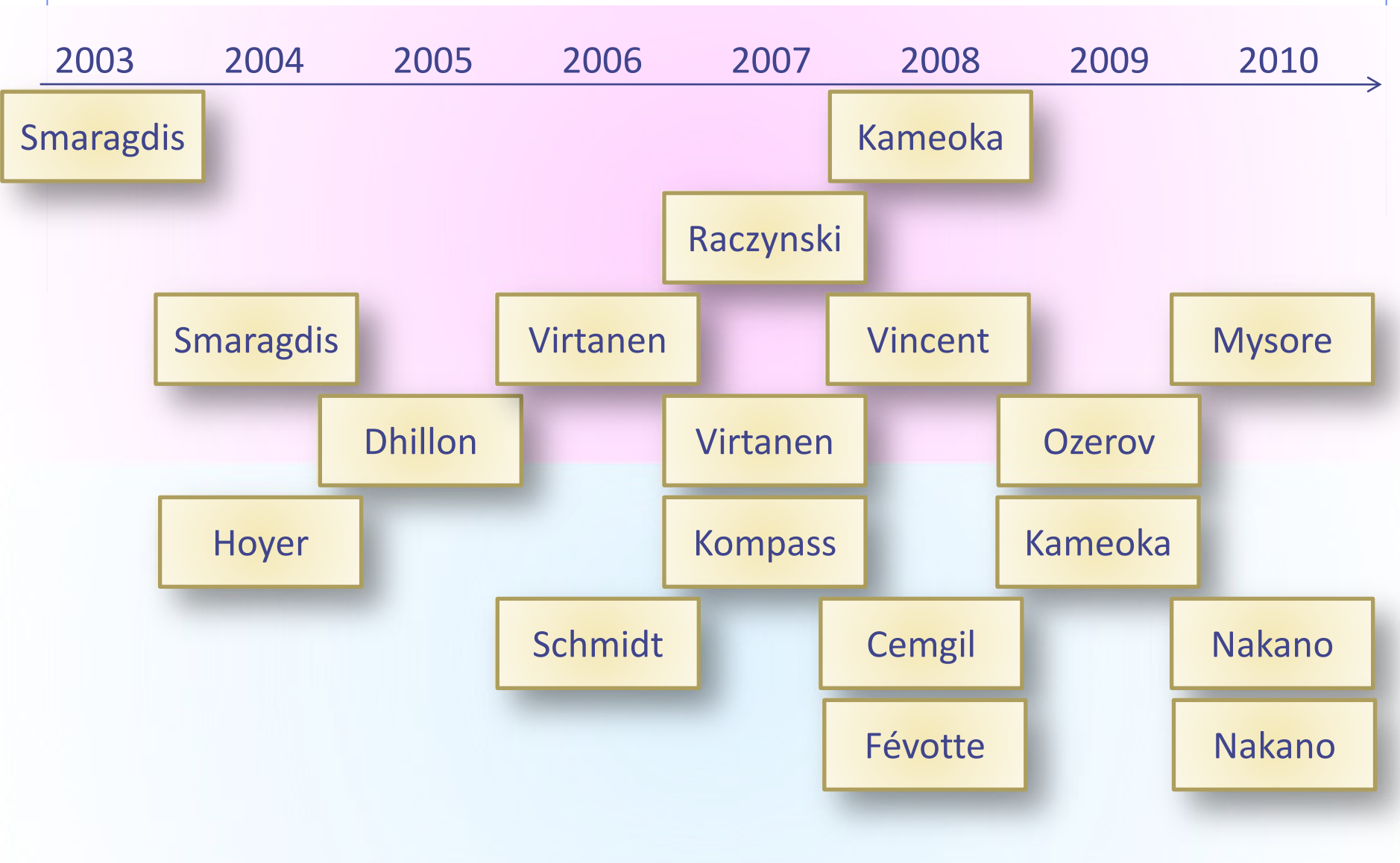
$$\begin{aligned} \mathcal{J} &= \sum_{\omega,t} \mathcal{D}_\beta \left(Y_{\omega,t} \middle| \sum_k H_{\omega,k} U_{k,t} \right) \\ &= \sum_{\omega,t} \left[\frac{Y_{\omega,t}^\beta}{\beta(\beta-1)} + \frac{1}{\beta} \underbrace{\left(\sum_k H_{\omega,k} U_{k,t} \right)^\beta}_{\text{凸/凹}} - \frac{Y_{\omega,t}}{\beta-1} \underbrace{\left(\sum_k H_{\omega,k} U_{k,t} \right)^{\beta-1}}_{\text{凸/凹}} \right] \end{aligned}$$

■ 補助関数の設計方針

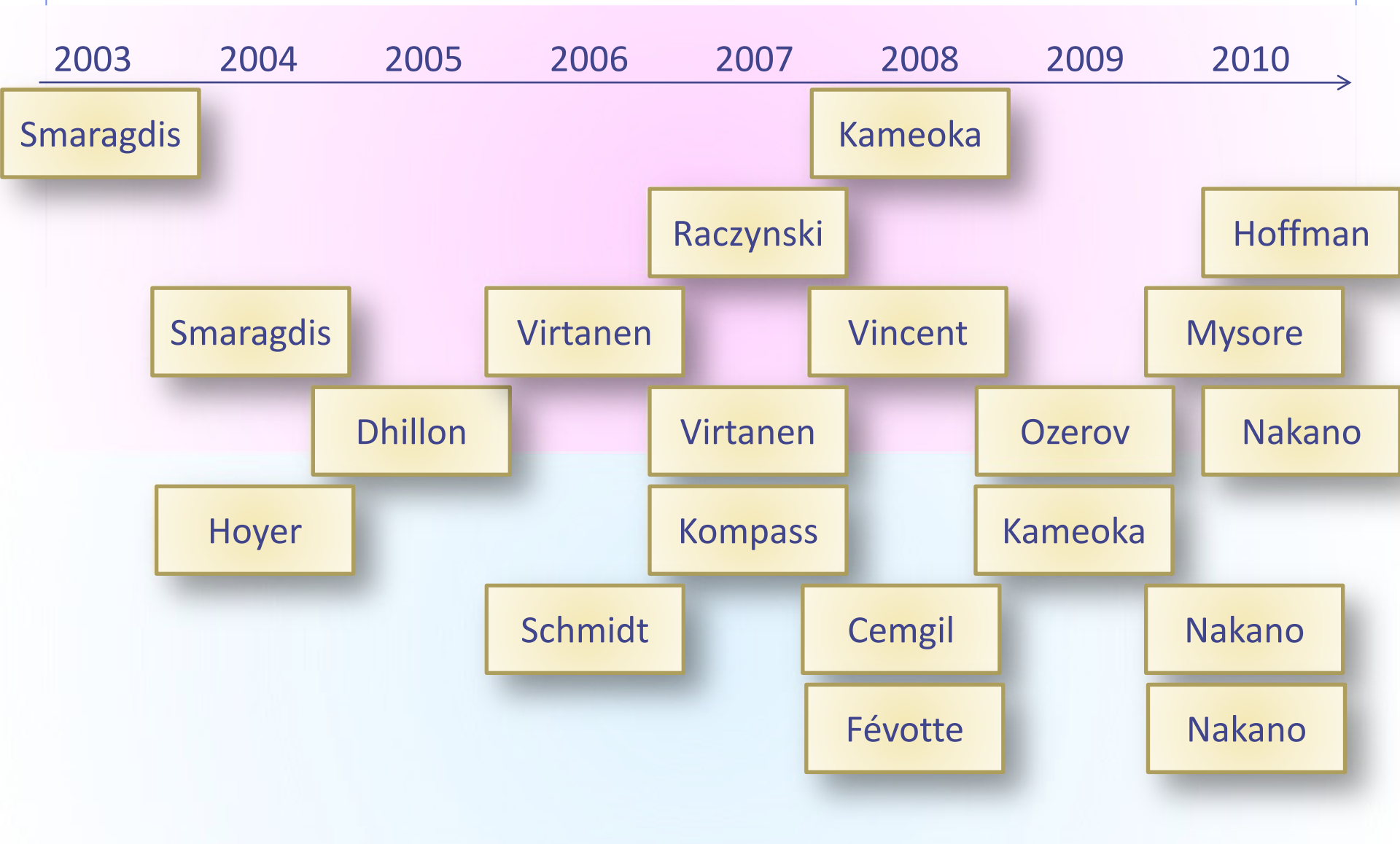
- ◆ Jensen不等式 $\phi \left(\sum_i \lambda_i z_i \right) \leq \sum_i \lambda_i \phi(z_i)$ $\begin{cases} \phi : \text{凸関数} \\ \sum_i \lambda_i = 1 \end{cases}$
 - ◆ 接線不等式 $\phi(z) \leq \phi(c) + \phi'(c)(z-c)$ $\phi : \text{凹関数}$
-

➡ 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, 補助関数を設計 [Kameoka2006]

NMFの改良・拡張モデル



NMFの改良・拡張モデル



ベイズノンパラメトリックNMF [Hoffman2011, Nakano2011]

◆ NMFにおける基底数をデータから推論できるようにしたい

■ Infinite NMF [Schmidt2010]

■ Gamma Process (GaP)-NMF [Hoffman2011]

◆ スペクトログラムの生成モデル

$$H_{\omega,k} \sim \text{Gamma}(a, a)$$

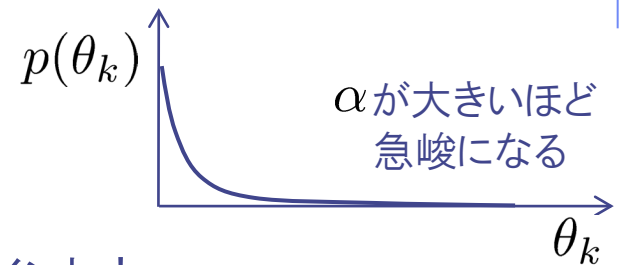
$$U_{k,t} \sim \text{Gamma}(b, b)$$

$$\theta_k \sim \text{Gamma}(\alpha/K, \alpha c) \longrightarrow \text{ここがポイント！}$$

$$Y_{\omega,t} \sim \text{Exponential}\left(\sum_k \theta_k H_{\omega,k} U_{k,t}\right) \rightarrow \text{FévotteのIS-NMF方式}$$

◆ $\theta_l \sim \text{Gamma}(\alpha/K, \alpha c)$ の意味

$\theta_1, \theta_2, \dots$ をスパースにする = 不要なコンポーネントはできるだけ使わない



◆ 状態遷移NMFにおけるコンポーネント数と状態数をデータから推論できるようにしたい

■ Infinite Factorial infinite Hidden Markov Model (iFiHMM) [Nakano2011]

まとめ

1. 非負値行列因子分解(NMF)

- 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

2. 複素NMF [Kameoka2008]

- NMFによるスペクトログラム分解法が根源的に抱える問題点
- これを解決する新モデル

3. NMFの改良・拡張モデル

- スパースNMF, NMFD, NMF2D, ソースフィルタNMF, アクティベーションの連続性規準入りNMF, ハーモニックNMF, 板倉齋藤距離規準NMF, 複素NMF, 状態遷移NMF, ノンパラメトリックベイズNMF, etc...

参考文献

- [1] D. D. Lee and H. S. Seung, "Learning the parts of objects with nonnegative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.
- [2] P. Paatero and 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," in Adv. *NIPS 2000*, pp. 556–562, 2000.
- [5] P. Smaragdis, B. Raj and M.V. Shashanka, "Supervised and semi-supervised separation of sounds from single-channel mixtures," in Proc. *ICA 2007*, pp. 414-421, 2007.
- [6] P. Smaragdis and B. Raj, "Example-driven bandwidth expansion," in Proc. *WASPAA 2007*, pp. 135-138, 2007.
- [7] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *J. Mach. Learning Res.*, vol. 5, pp. 1457–1469, 2004.
- [8] P. Smaragdis, "Non-negative matrix factor deconvolution; extraction of multiple sound sources from monophonic inputs," in Proc. *ICA 2004*, pp. 494–499, 2004.
- [9] 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.

参考文献

- [10] M. N. Schmidt and M. Mørup, “Nonnegative matrix factor 2-D deconvolution for blind single channel source separation,” in *Proc. ICA 2006*, pp. 700-707, 2006.
- [11] T. Virtanen and A. Klapuri, “Analysis of polyphonic audio using source-filter model and non-negative matrix factorization,” in *Adv. NIPS 2006*, 2006.
- [12] S. Eguchi and Y. Kano, “Robustifying maximum likelihood estimation,” Technical report, Institute of Statistical Mathematics, Research Memo. 802, 2001.
- [13] R. Kompass, “A generalized divergence measure for nonnegative matrix factorization,” *Neural Computation*, vol. 19, no. 3, pp. 780–791, 2007.
- [14] T. Virtanen, “Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria,” *IEEE Trans. on Audio, Speech and Language Processing*, vol. 15, no. 3, pp. 1066–1074, 2007.
- [15] S. A. Raczynski, N. Ono and S. Sagayama, “Multipitch analysis with harmonic nonnegative matrix approximation,” in *Proc. of ISMIR 2007*, pp.381-386, 2007.
- [16] E. Vincent, N. Bertin and R. Badeau, “Harmonic and inharmonic nonnegative matrix factorization for polyphonic pitch transcription,” in *Proc. ICASSP 2008*, pp. 109-112, 2008.
- [17] A. T. Cemgil, “Bayesian inference for nonnegative matrix factorization models,” Technical Report CUED/F-INFENG/TR.609, University of Cambridge, 2008.
- [18] C. Févotte, N. Bertin and J.-L. Durrieu, “Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis,” *Neural Computation*, vol. 21, no. 3, pp. 793-830, 2009.
- [19] 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.

参考文献

- [20] 亀岡, ルルー, 大石, 柏野, "Music Factorizer: 音楽音響信号をノート単位で編集できるインタフェース," 情報処理学会研究報告, 2009-MUS-81-9, 2009.
- [21] Z. Ghahramani and M. I. Jordan, "Factorial hidden Markov models," *Machine Learning*, vol. 29, pp. 245-273, 1997.
- [22] A. Ozerov, C. Févotte and M. Charbit, "Factorial scaled hidden Markov model for polyphonic audio representation and source separation," in *Proc. WASPAA 2009*, pp. 121-124, 2009.
- [23] 中野, 北野, ルルー, 亀岡, 小野, 嵯峨山, "可変基底NMFに基づく音楽音響信号の解析," 情報処理学会研究報告, 2010-MUS-84-10, 2010.
- [24] G. J. Mysore, P. Smaragdis and B. Raj, "Non-negative hidden Markov modeling of audio with application to source separation," in *Proc. LVA/ICA 2010*. pp.140-148, 2010.
- [25] 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," 情報処理学会研究報告, 2006-MUS-66-13, pp. 77-84, 2006.
- [26] M. Nakano, H. Kameoka, J. Le Roux, Y. Kitano, N. Ono and S. Sagayama, "Convergence-guaranteed multiplicative algorithms for non-negative matrix factorization with beta-divergence," In *Proc. MLSP 2010*, pp. 283-288, 2010.
- [27] M. N. Schmidt and M. Mørup, "Infinite non-negative matrix factorization," in *Proc. EUSIPCO*, 2010.
- [28] M. Hoffman, D. Blei and P. Cook, "Bayesian nonparametric matrix factorization for recorded music," in *Proc. ICML*, pp.439-446, 2010.
- [29] 中野, ルルー, 亀岡, 中村, 小野, 嵯峨山, "スペクトログラムのベイジアンノンパラメトリックモデリングに基づく音楽信号の解析," 情報処理学会研究報告, 2011-MUS-91-6, 2011.

レポート課題の対象論文

- ◆ A. T. Cemgil, “Bayesian inference for nonnegative matrix factorization models,” Technical Report CUED/F-INFENG/TR.609, University of Cambridge, 2008.
- ◆ C. Févotte, N. Bertin and J.-L. Durrieu, “Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis,” *Neural Computation*, vol. 21, no. 3, pp. 793-830, 2009.