20 スムース正則化付きNMFと打楽器スペクトル表現への適用

☆小川 遼，北村 大地，綾野 翔馬（香川高専）

1. 研究背景

- 非負値行列因子分解（nonnegative matrix factorization: NMF）
  - 非負行列を非負低ランク行列で近似する手法

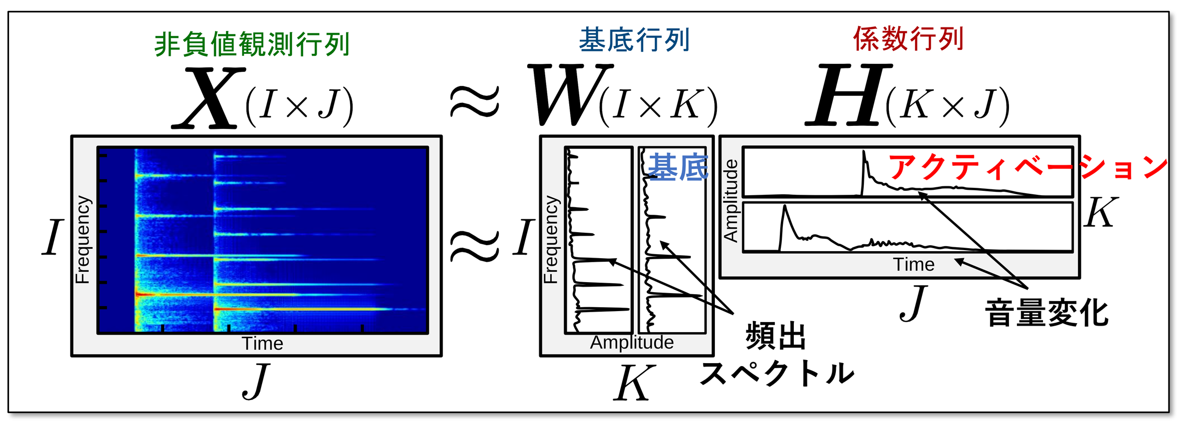
非負値観測行列　基底行列　係数行列　近似行列

$$\boldsymbol{X}^{(I\times J)} \approx \boldsymbol{W}^{(I\times K)}\boldsymbol{H}^{(K\times J)} = \widehat{\boldsymbol{X}}^{(I\times J)}$$

$$I\begin{bmatrix}1&2&0&0&1\\1&3&1&2&2\\0&0&3&5&3\end{bmatrix}_J \approx I\begin{bmatrix}1&0\\1&1\\0&3\end{bmatrix}_K \begin{bmatrix}1&2&0&0&1\\0&0&1&2&1\end{bmatrix}^K_J = \begin{bmatrix}1&2&0&0&1\\1&2&1&2&2\\0&0&3&6&3\end{bmatrix}$$

基底　アクティベーション

  - 音響信号のスペクトログラムを近似することにも使える



- NMFの音響分野での応用
  - 音源分離，自動採譜，音色変換，楽器同定など

2. 既存の正則化付きNMF

- NMFにおける正則化の導入
  - 目的関数に正則化項を追加することでなんらかの事前情報を反映
- 正則化の従来手法の例
  - スパース正則化
    - スパース性とは要素の多くがゼロである性質

$$\underset{\boldsymbol{W},\boldsymbol{H}}{\text{Minimize}}\ \mathcal{D}(\boldsymbol{X}|\boldsymbol{W}\boldsymbol{H}) + \underbrace{\mu\sum_{i,k} w_{i,k}^2}_{\text{正則化項}}\ \text{s.t.}\ w_{i,k}, h_{k,j} \geq 0\ \forall i,j,k$$

正則化なし $\begin{bmatrix}3.6&0.7&0.7\\0.5&0.7&2.4\\0.7&0.1&0.9\\0.1&1.4&0.9\end{bmatrix}$

スパース正則化あり $\begin{bmatrix}5&0&0\\0&0&4\\0&0&0\\0&3&0\end{bmatrix}$

  - スムース正則化
    - スムース性とは要素が滑らかに変化する性質

$$\underset{\boldsymbol{W},\boldsymbol{H}}{\text{Minimize}}\ \mathcal{D}(\boldsymbol{X}|\boldsymbol{W}\boldsymbol{H}) + \underbrace{\mu\sum_{i,k} |w_{i,k} - w_{i+1,k}|^2}_{\text{正則化項}}\ \text{s.t.}\ w_{i,k}, h_{k,j} \geq 0\ \forall i,j,k$$

スムース正則化あり $\begin{bmatrix}1.4&0.5&1.1\\1.5&0.4&1.2\\1.7&0.3&1.0\\1.8&0.3&1.0\end{bmatrix}$

3. 提案手法

- ディリクレ分布に基づくNMF（Dirichlet NMF）
  - 基底行列の各列（基底ベクトル）にスムース性を促す正則化付きNMFを提案
  - 正則化項として，基底ベクトルがディリクレ分布から生成されることを仮定
- ディリクレ分布
  - 0～1の要素を持ち総和が1のベクトルを生成する多変量分布

対称ディリクレ分布の確率密度関数

$$\boldsymbol{x} \sim p(\boldsymbol{x}; a) = \frac{1}{B(a)}\prod_i x_i^{a-1}\ \ \text{s.t.}\ \sum_i x_i = 1, x_i \geq 0\ \forall i$$

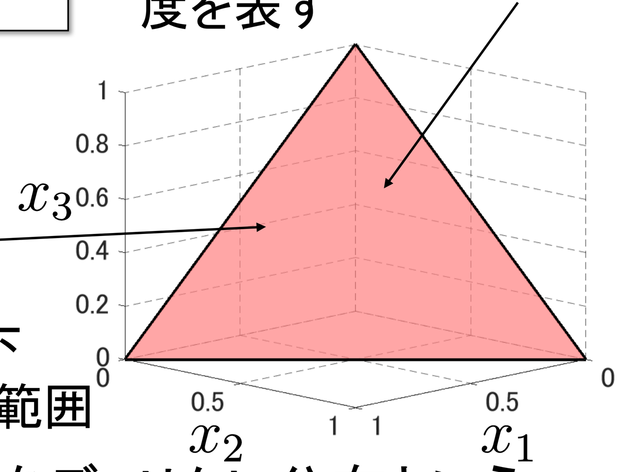
$B(a)$: 多変量ベータ関数

ディリクレ分布は，この単体上の確率変数ベクトル$\boldsymbol{x}$の出現確率密度を表す

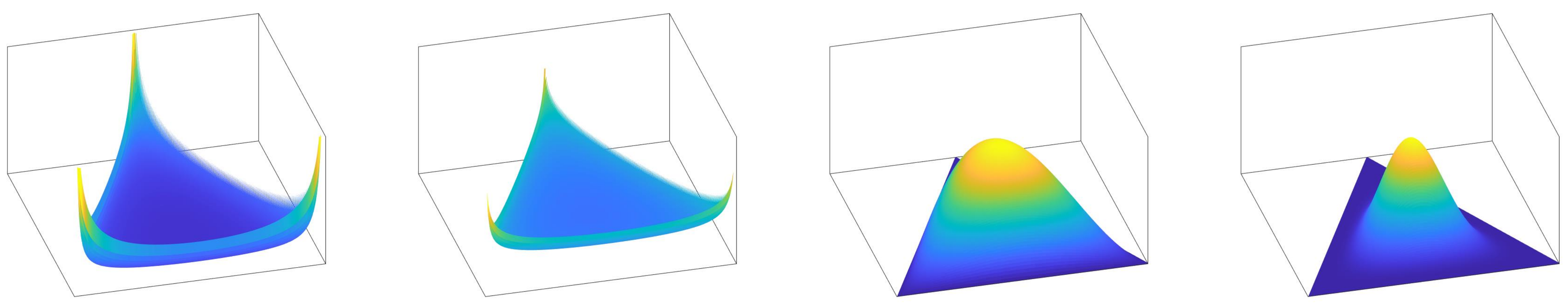
例：3次元の場合

$$\boldsymbol{x} = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\quad x_1+x_2+x_3=1,\quad x_1,x_2,x_3 \geq 0$$

この制約条件下で$\boldsymbol{x}$が取りうる範囲



  - このようなベクトルを確率変数にもつ分布をディリクレ分布という



$a = 0.5$　$a = 0.8$　$a = 2.0$　$a = 5.0$

  - aを大きくするとスムースなベクトルを生成しやすくなる　$\boldsymbol{x} = \begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}$
- 目的関数

$$\underset{\boldsymbol{W},\boldsymbol{H}}{\text{Minimize}}\ \mathcal{D}(\boldsymbol{X}|\boldsymbol{W}\boldsymbol{H}) + \sum_{i,k}(a_k - 1)\log w_{ik}^{-1}$$
$$\text{s.t.}\ w_{ik}, h_{kj} \geq 0,\ \sum_i w_{ik} = 1\ \forall i,j,k$$

  - a_kの値でk番目の基底に対しての正則化のパラメータを調整
    - k番目の基底に対してスムース度を決定できる
  - 対称ディリクレ分布による制約で行列$\boldsymbol{W}$の列方向の総和が1という制約

- 補助関数法
  - 目的関数の上限となるように定義された補助関数を反復的に降下させることによって，近似的に解を探索する手法 [Lee+, 2000]
  - 反復更新による最適化アルゴリズムが導出され，毎回の反復で目的関数の値が単調に減少することが保証されている

$$\mathcal{J}^+ = \sum_{i,j}\left[x_{ij}\log x_{ij} - x_{ij}\sum_k \delta_{ijk}\log\frac{w_{ik}h_{kj}}{\delta_{ijk}} - x_{ij} + \sum_k w_{ik}h_{kj}\right] + \sum_{i,k}(a_k - 1)\log w_{ik}^{-1}$$

  - Dirichlet NMFの目的関数に補助関数法を適用したものが上式
- Dirichlet NMFの反復更新式
  - 補助関数法を適用したDirichlet NMFの目的関数を導出
  - $\sum_i w_{ik} = 1$というディリクレ分布の制約が存在する
  - 補助関数に対してラグランジュの未定乗数法を適用し，制約条件を満たしたまま最適化する更新則を新たに提案

$$w_{ik}^{(t+1)} = \frac{w_{ik}^{(t)}\sum_j \frac{x_{ij}}{\sum_k' w_{ik'}^{(t)}h_{k'j}^{(t)}}h_{kj}^{(t)} + a_k - 1}{\sum_i\left(w_{ik}^{(t)}\sum_j \frac{x_{ij}}{\sum_k' w_{ik'}^{(t)}h_{k'j}^{(t)}}h_{kj}^{(t)} + a_k - 1\right)}$$

  - ただし，上式が $a_k \geq 1$での条件でのみ有効
    - $0 < a_k < 1$でも有効な更新式はASJ2025春研発で発表予定
    - $0 < a_k < 1$ではスパース正則化（one-hot化）も可能になる

4. 実験

- ①Dirichlet NMFによる正則化の効果の確認
  - 乱数とスムースな基底から構成される観測行列$\boldsymbol{X}$を作成
  - 観測行列にDirichlet NMFを適用し正解の基底行列と比較する

正解の基底行列　擬似乱数行列　非負値観測行列　基底行列　係数行列

$$\overline{\boldsymbol{W}}\ \overline{\boldsymbol{H}} = \boldsymbol{X} \approx \boldsymbol{W}\boldsymbol{H}$$

$$\text{乱数}\begin{bmatrix}0.1&0&0.2\\0.4&0.1&0.2\\0&0.4&0.2\\0.3&0&0.2\\0.2&0.5&0.2\end{bmatrix}\text{スムース}\begin{bmatrix}1&2&3&4&5\\5&4&3&2&1\\1&2&3&4&5\end{bmatrix} = \begin{bmatrix}0.3&0.6&0.9&1.2&1.5\\1.1&1.6&2.1&2.6&3.1\\2.2&2.0&1.8&1.6&1.4\\0.5&1.0&1.5&2.0&2.5\\2.9&2.8&2.7&2.6&2.5\end{bmatrix} \approx \boldsymbol{W}\begin{bmatrix}\ ?\ \end{bmatrix}$$

  - 各基底に対してパラメータを$a_1=1, a_2=1, a_3=50$に設定
    - パラメータが1だと正則化なしNMFと等価

0.1	0	0.2
0.4	0.1	0.2
0	0.4	0.2
0.3	0	0.2
0.2	0.5	0.2

スムース

正解の基底行列の$\boldsymbol{W}$

0.09561	0.1215	0.08452
0.1837	0.2932	0.2257
0.01097	0.1025	0.468
0.101	0.2072	0.1919
0.6087	0.2756	0.02987

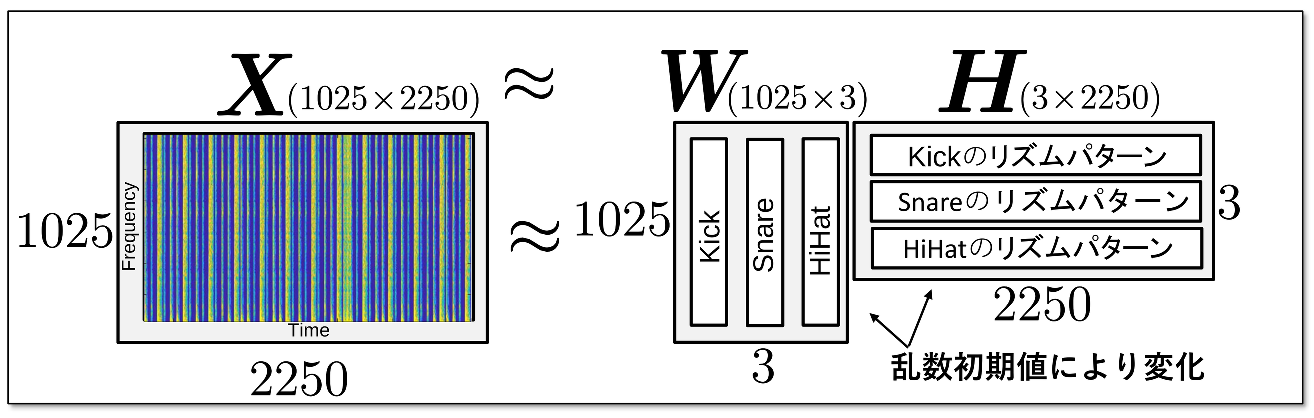
正則化なしNMFの$\boldsymbol{W}$

0.04073	0.01981	0.1997
0.2525	0.2356	0.2015
0.04947	0.3394	0.1983
0.1403	0.1112	0.2018
0.517	0.294	0.1987

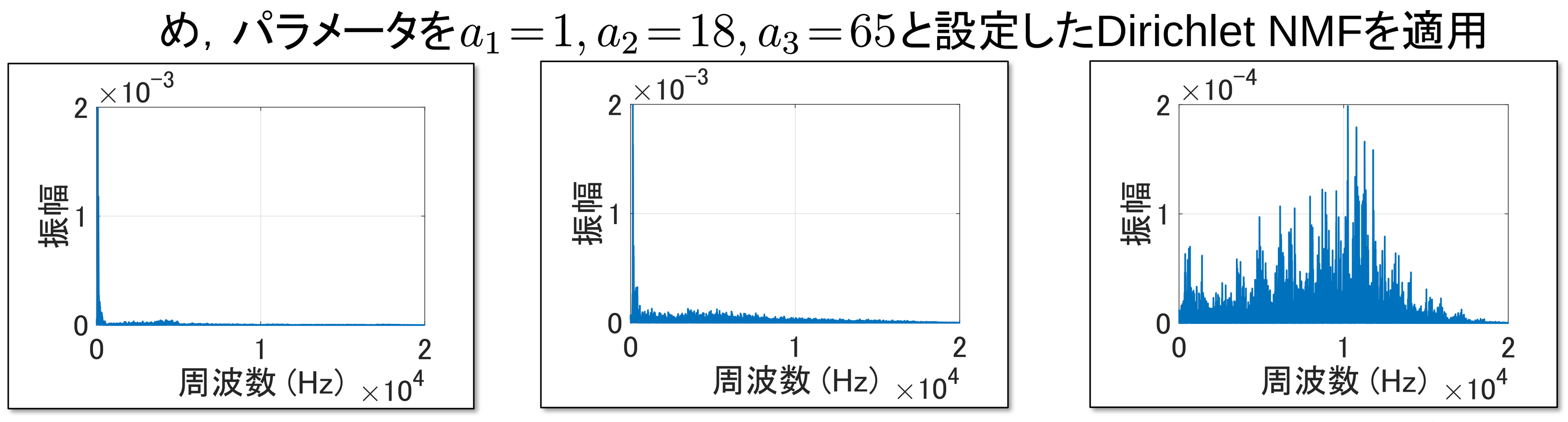
スムース

Dirichlet NMFの$\boldsymbol{W}$

- ②Dirichlet NMFの打楽器スペクトル表現への適用
  - ドラムのKick・Snare・HiHatを含む音源のスペクトログラムを，基底3本のNMFで分解し，Kick・Snare・HiHatのスペクトルを得る

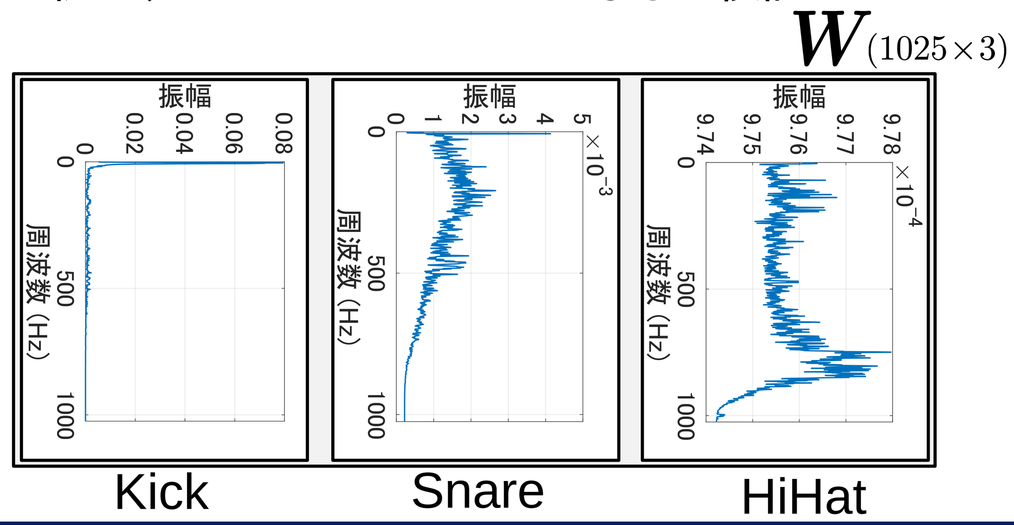


    - 正則化なしのNMFでは，得られる基底の順番が$\boldsymbol{W}$や$\boldsymbol{H}$の乱数初期値に依存して変化（必ずしもKick・Snare・HiHatの順番にならない）
    - Snareのスペクトルはややスムース，HiHatのスペクトルはかなりスムースなため，パラメータを$a_1=1, a_2=18, a_3=65$と設定したDirichlet NMFを適用



Kickのスペクトル　Snareのスペクトル　HiHatのスペクトル

    - 基底の順番が乱数初期値に依らずKick・Snare・HiHatになるか検証
  - 1000種類中934種類の乱数初期値で基底がKick・Snare・HiHatの順番となった（正則化で誘導された）



Dirichlet NMFで得られる基底行列の一例

Kick　Snare　HiHat