

東京大学工学部 4年生 夏学期 [03501130]

応用音響学 第7回 (5/27)

亀岡弘和

東京大学大学院情報理工学系研究科
システム情報学専攻
kameoka@hil.t.u-tokyo.ac.jp

講義スケジュール

前半(亀岡担当)

- 4/08: 第1回
- 4/15: 第2回
- 4/22: 第3回
- 4/29: 休日
- 5/01(木): 第4回
- 5/06: 休日
- 5/13: 第5回
- 5/20: 第6回
- 5/27: 第7回

後半(牧野担当)

- 6/03: 第8回
- 6/10: 第9回
- 6/17: 第10回
- 6/24: 第11回
- 7/01: 第12回
- 7/08: 第13回
- 7/15: 第14回
- 7/22: 学期末試験

講義資料と成績評価

■講義資料

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

■成績評価

- 出席点
- 学期末試験

本日の話題

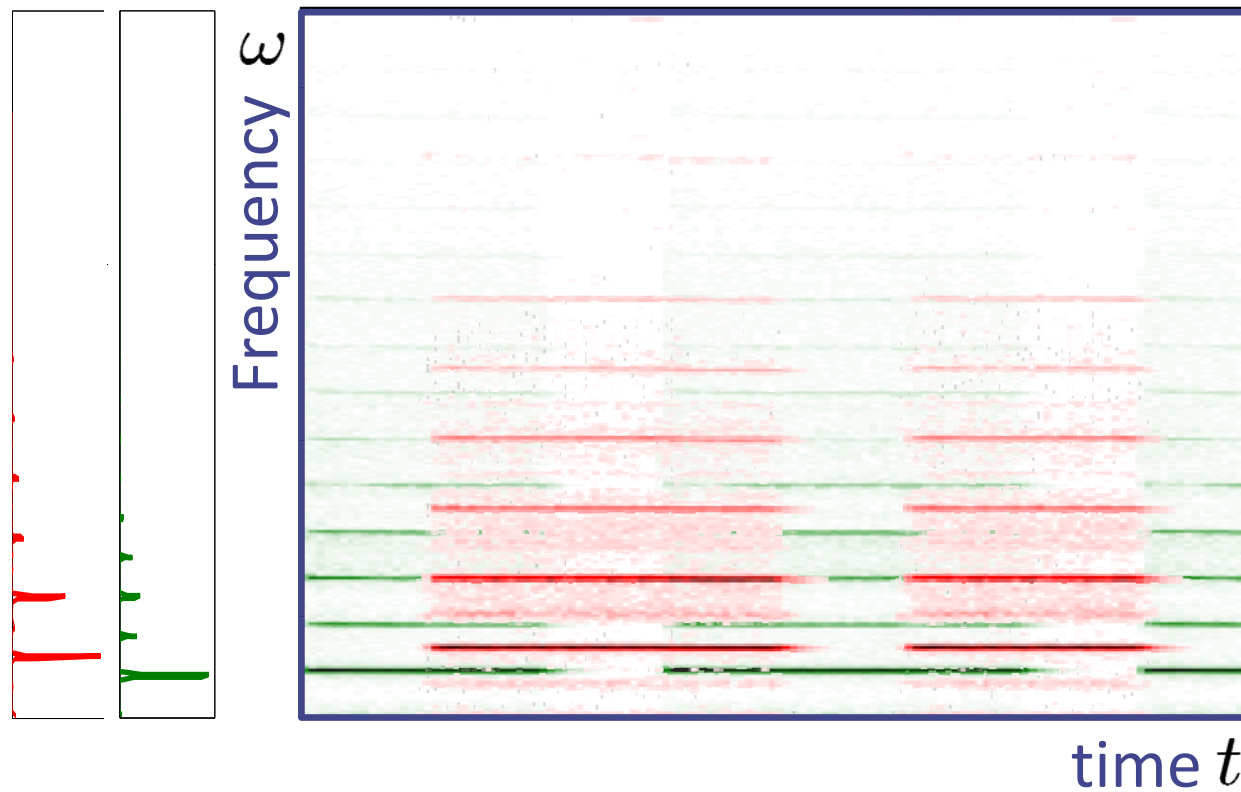
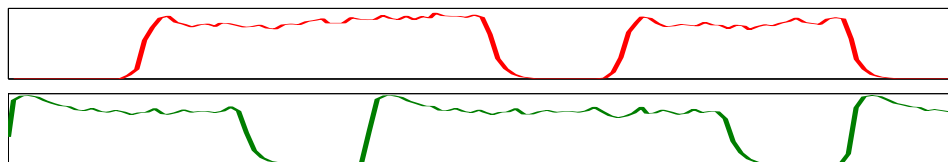
■ 非負値行列因子分解

- 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
- 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

■ 参考資料

- A. Cichocki, R. Zdunek, A. Huy Phan, S. Amari, *Non-negative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, John Wiley & Sons, 2009.
- 亀岡弘和, "非負値行列因子分解," 計測と制御, Vol. 51, No. 9, pp. 835-844, Sep. 2012.

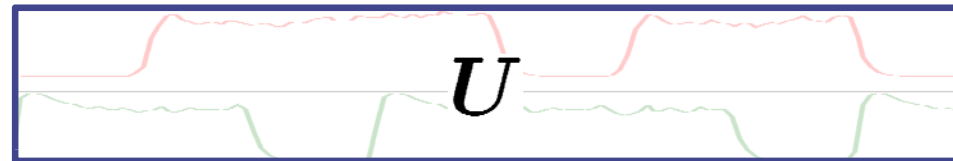
「行列の積」としてのスペクトログラム



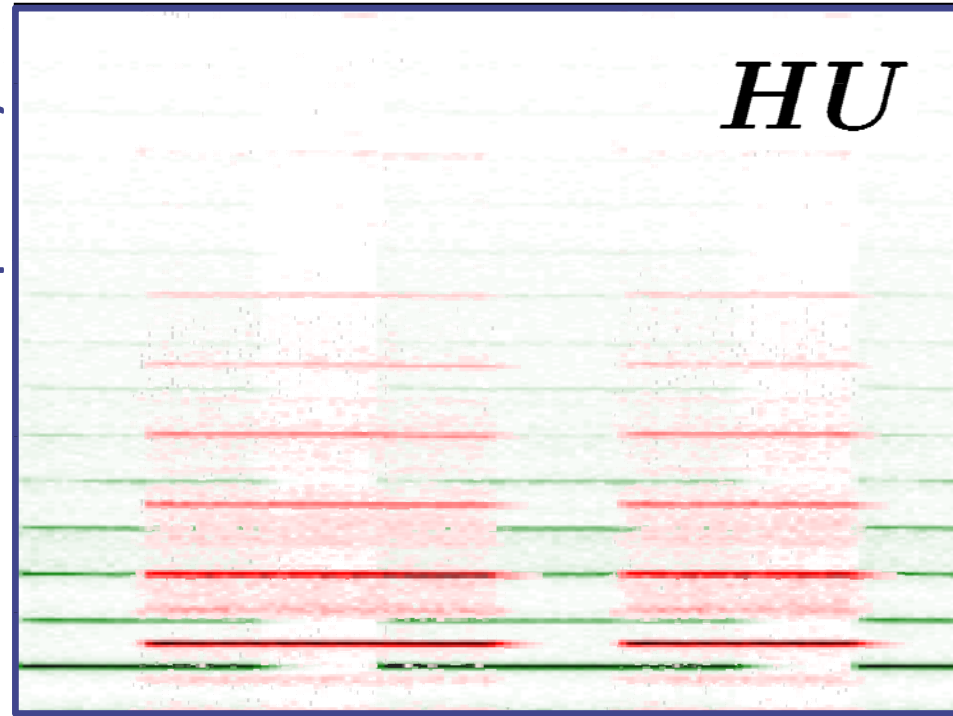
「行列の積」としてのスペクトログラム

各基底の「アクティベーション」

「基底スペクトル」

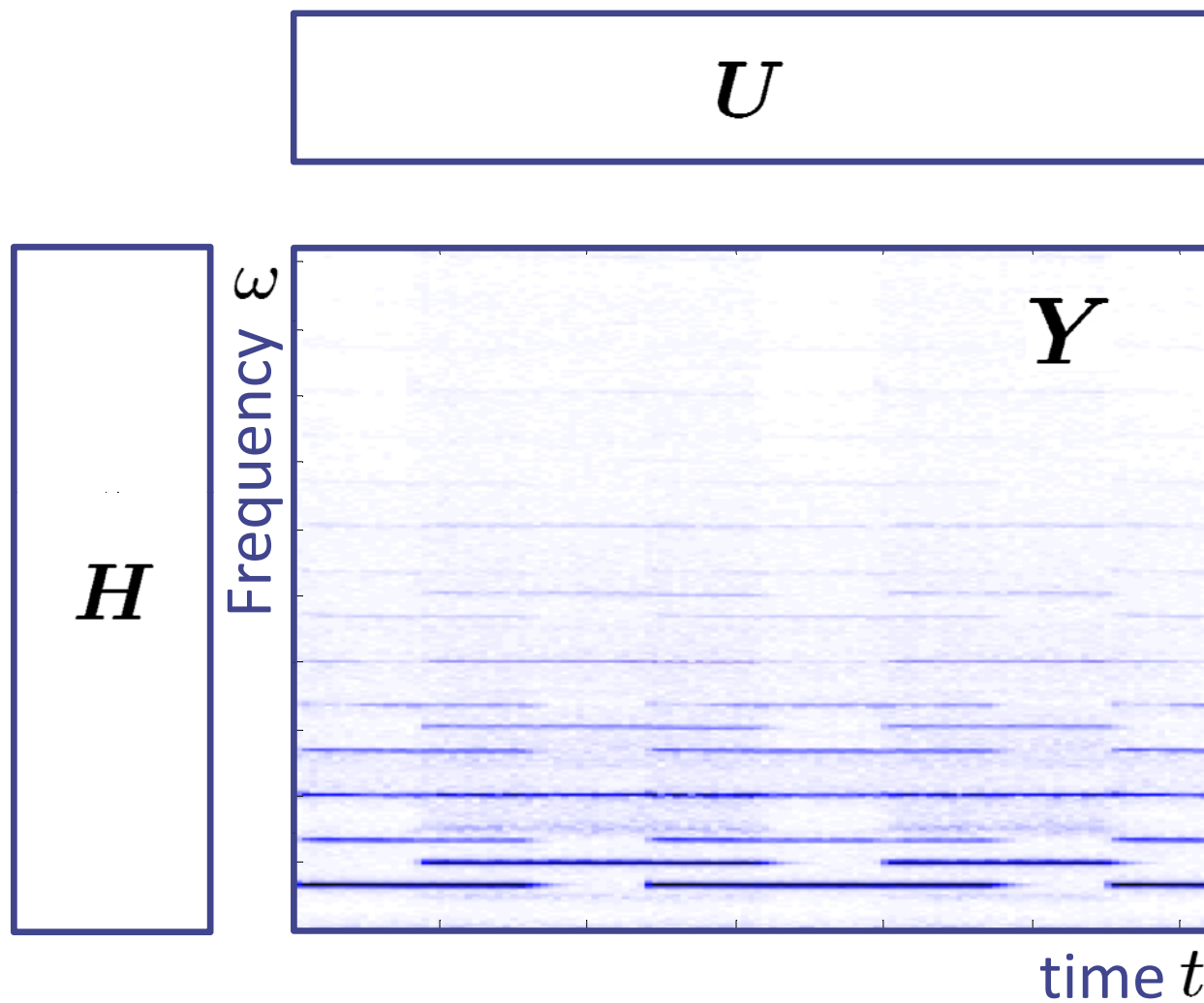


Frequency ω



time t

「行列の積」としてのスペクトログラム



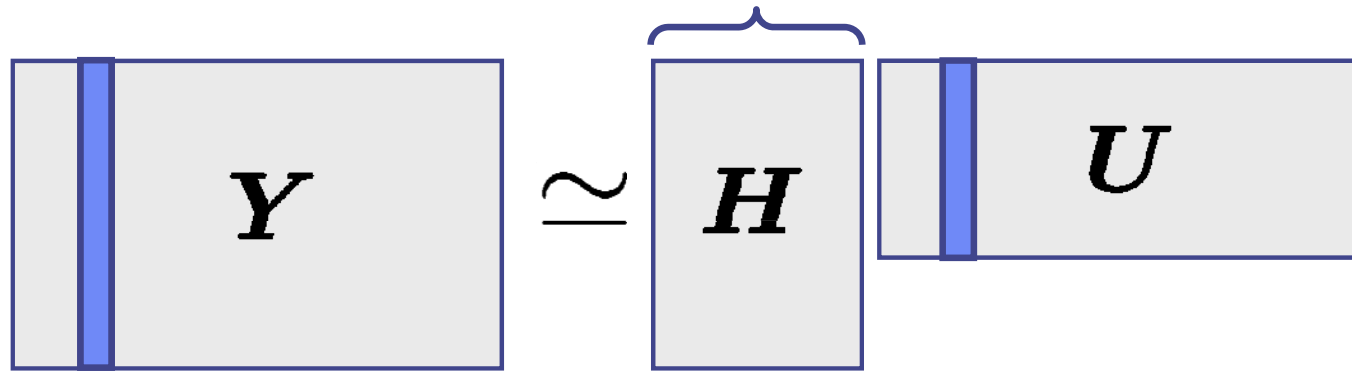
本日の話題

■ 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

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

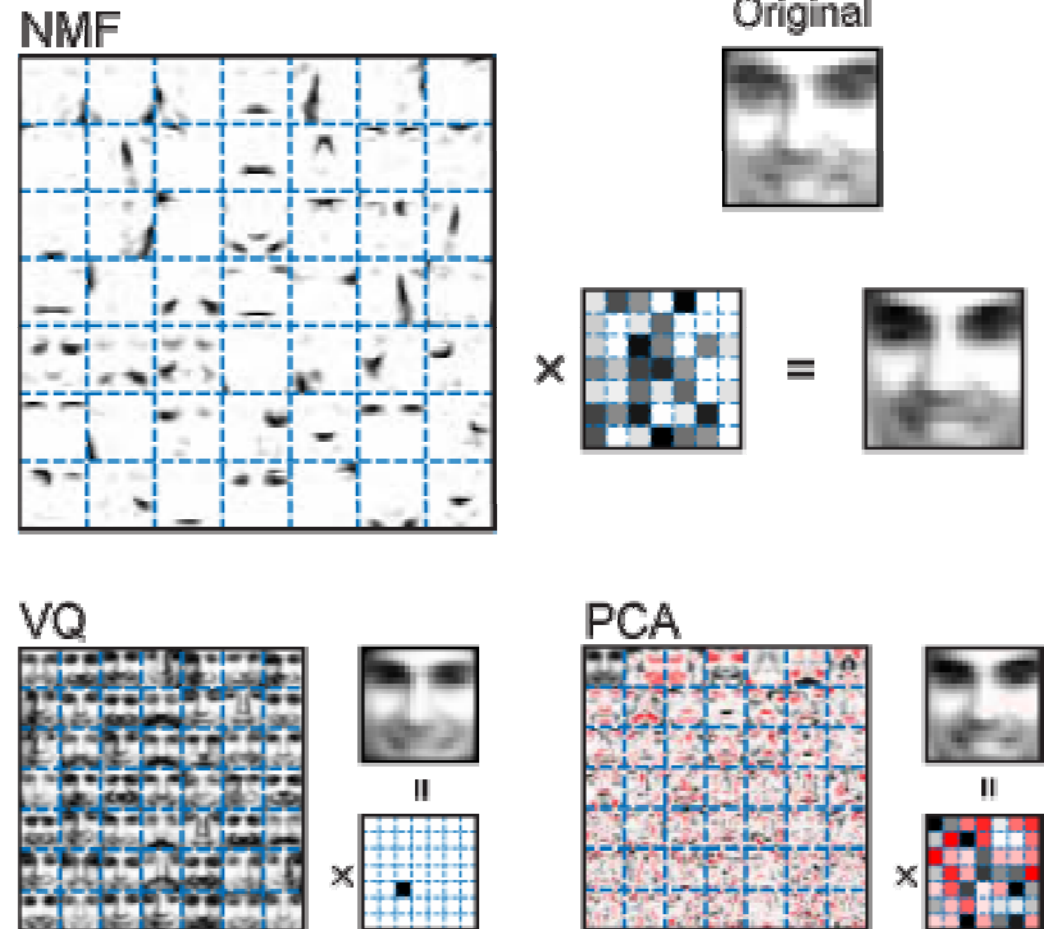
NMFが生まれた背景

- ◆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]

$$H_{\omega,k} \leftarrow H_{\omega,k} \frac{[YU^T]_{\omega,k}}{[HUU^T]_{\omega,k}}$$
$$U_{k,t} \leftarrow U_{k,t} \frac{[H^T Y]_{k,t}}{[H^T H U]_{k,t}}$$



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

Observed
vector

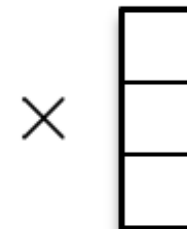


Basis
vectors

(a)



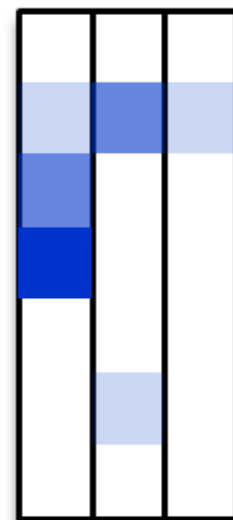
Coeffi-
cients



×

■ 係数が負の値を
取ってはならない場合

(b)



×



スパース

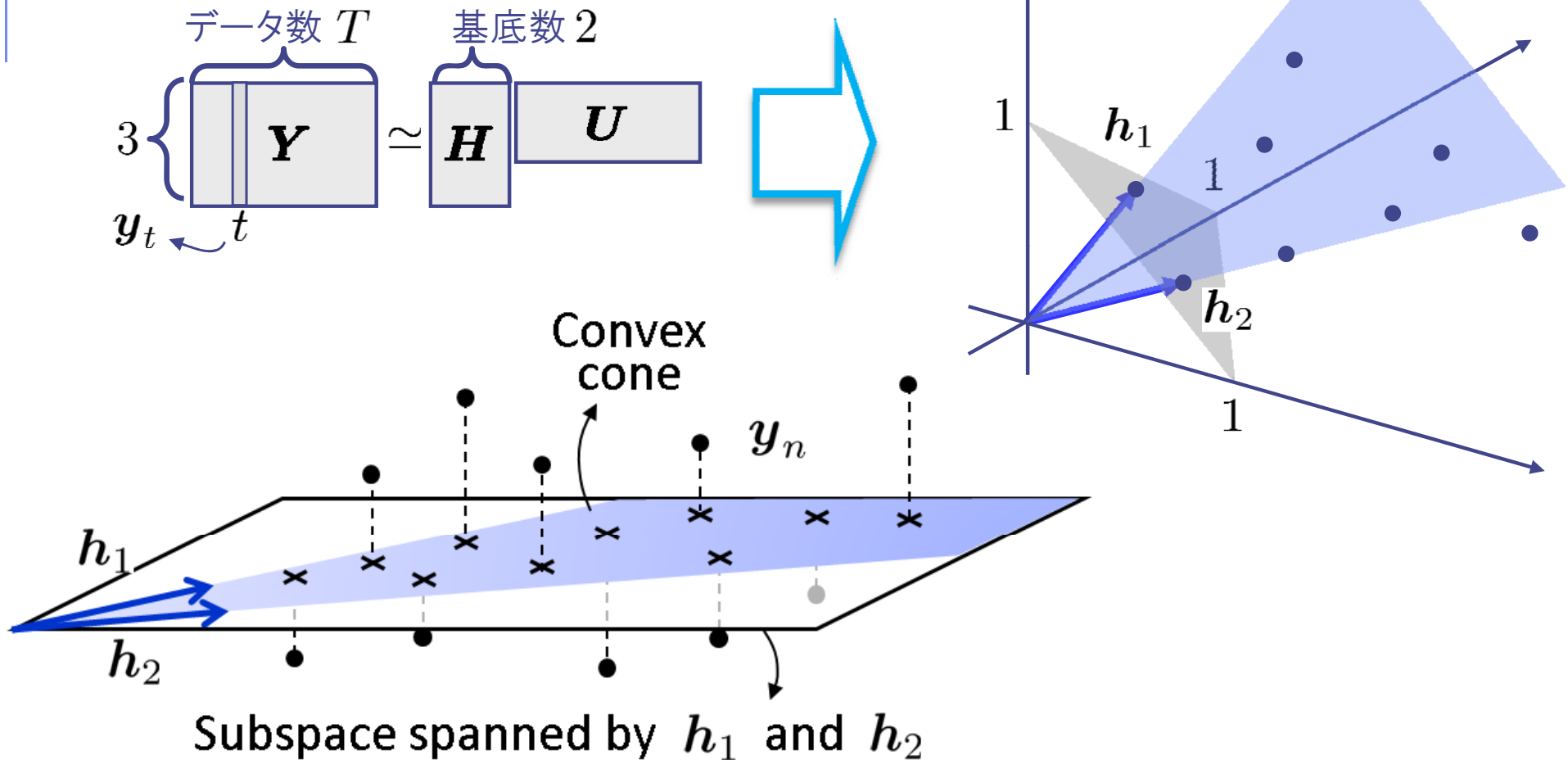
スパース性 $\approx$ 統計的独立性

NMFの幾何学的イメージ

■ 二乗誤差規準のNMF

$$\text{minimize } \|Y - HU\|_F^2 = \sum_t \left\| y_t - \sum_k h_k U_{k,t} \right\|_2^2$$

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



共起成分のグルーピング効果

- 低ランク近似問題：
少ない基底でデータを
良く表現したい

$$Y \simeq H U$$

- H をどう
した方が
「節約的」？

The diagram illustrates the low-rank approximation $Y \approx HU$. On the left, a large matrix Y is shown, composed of several small colored squares. This is followed by an equals sign, then a matrix H with three columns labeled a , b , and c . Matrix H has a few colored squares, indicating a sparse basis. This is followed by a multiplication sign $\times$, and then a matrix U labeled "Coefficients". Matrix U has three rows and several columns, with colored squares representing the coefficients for each basis vector in H .

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)

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

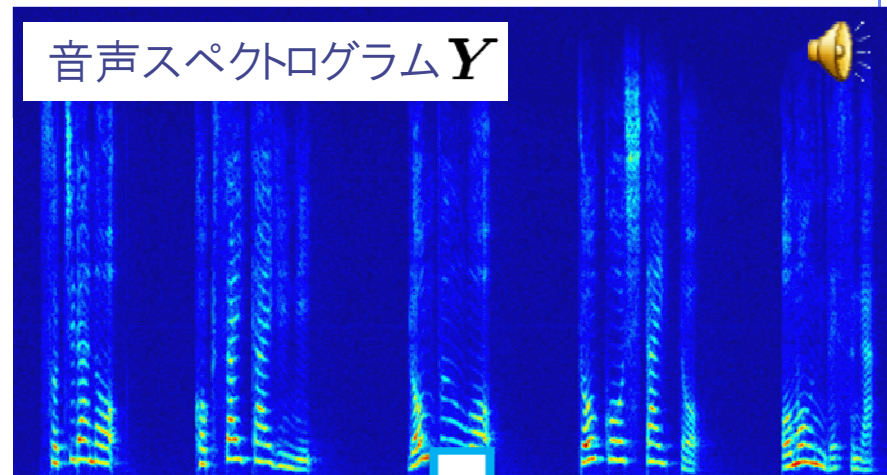
絶対値をとる

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

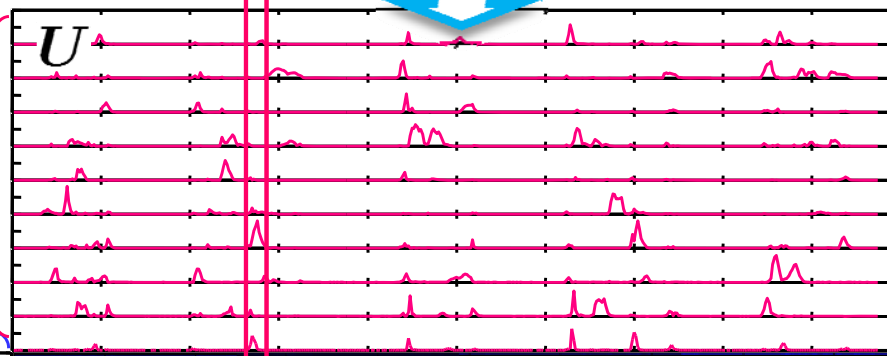
$$Y \simeq HU \quad (H \succeq 0, U \succeq 0)$$

↑
周波数

音声スペクトログラム Y

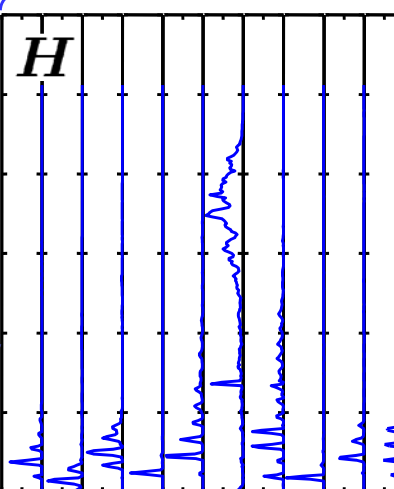


各基底の
アクティビティ



基底スペクトル

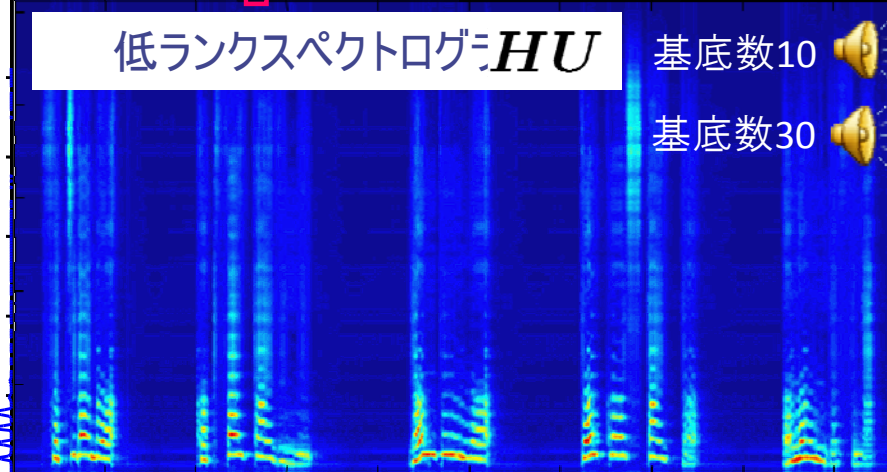
↑
周波数



低ランクスペクトログラム HU

基底数10

基底数30



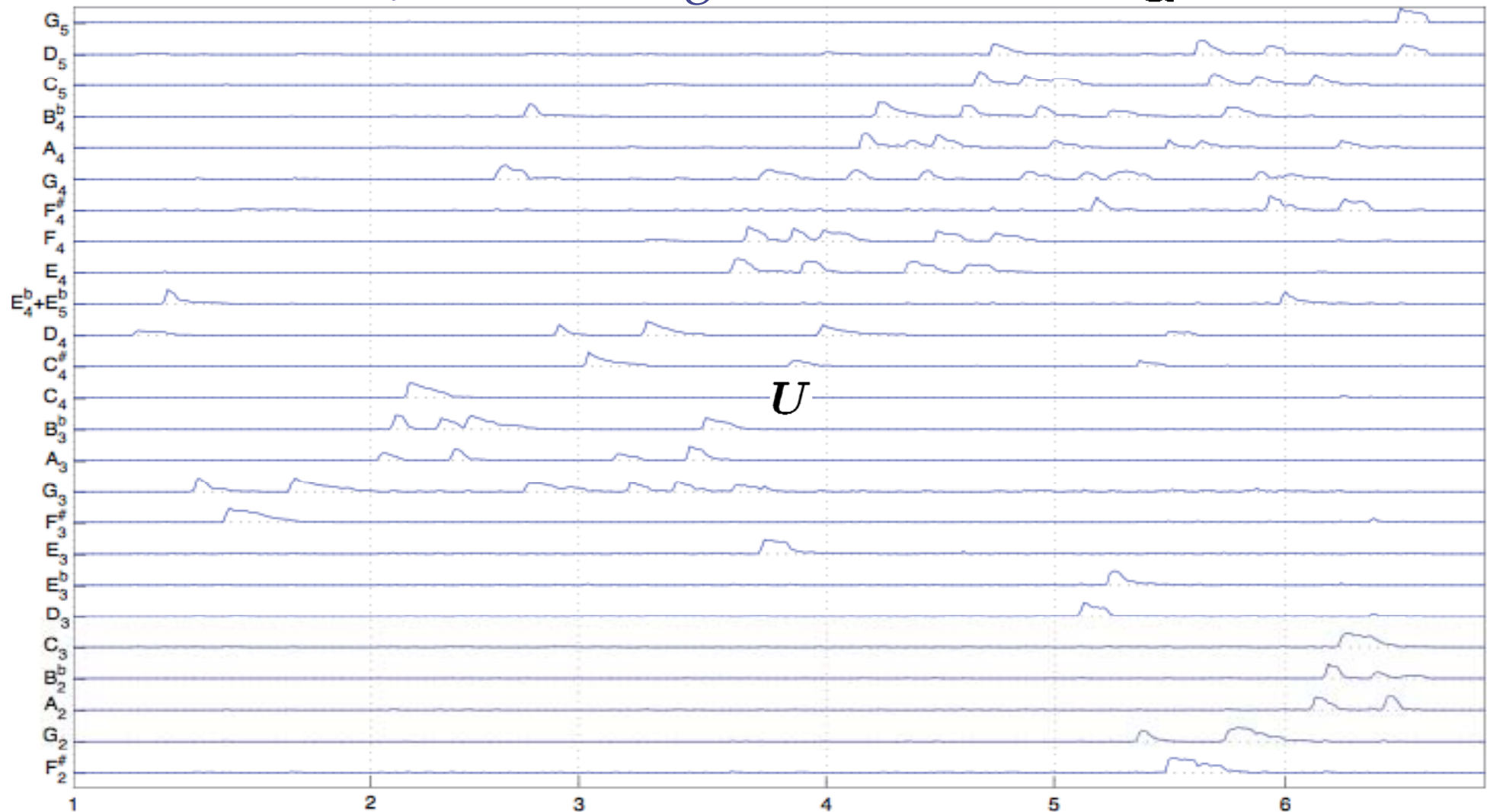
頻出するスペクトル
パーツが獲得される

時刻→

何に使えるのか？(1/5)

■ 自動採譜 [P. Smaragdis et al., 2003]

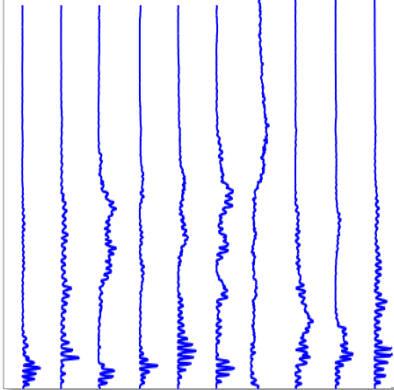
J.S. Bach: Fuge #16 in G minor 🗣️



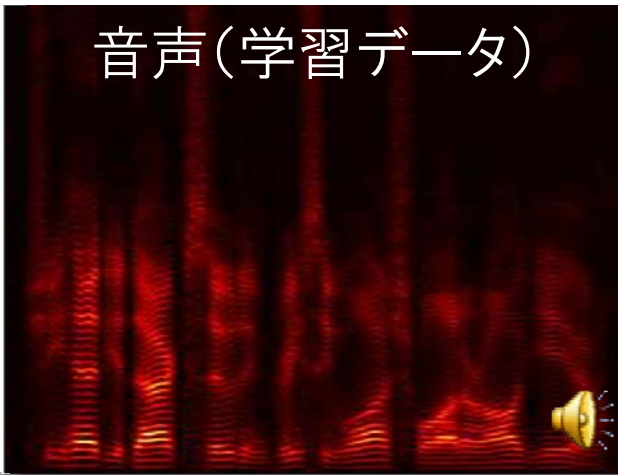
何に使えるのか？(2/5)

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

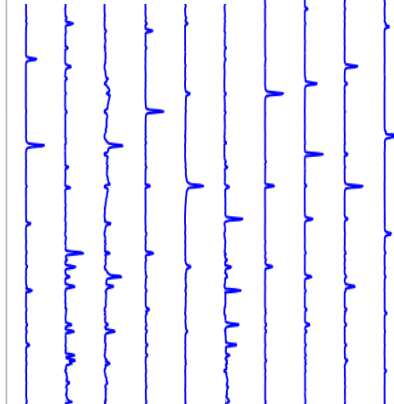
H_{speech}



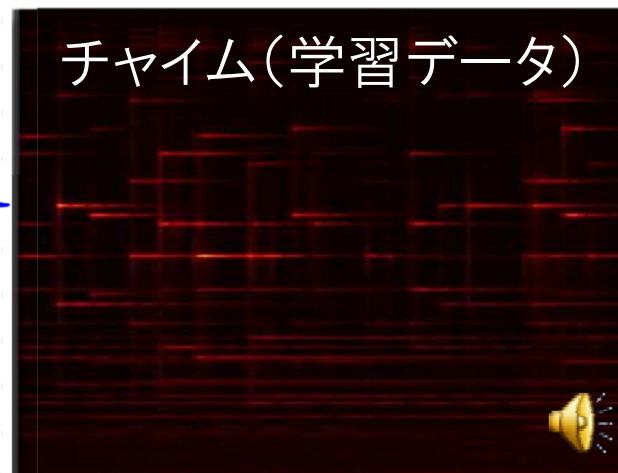
音声(学習データ)



H_{chime}

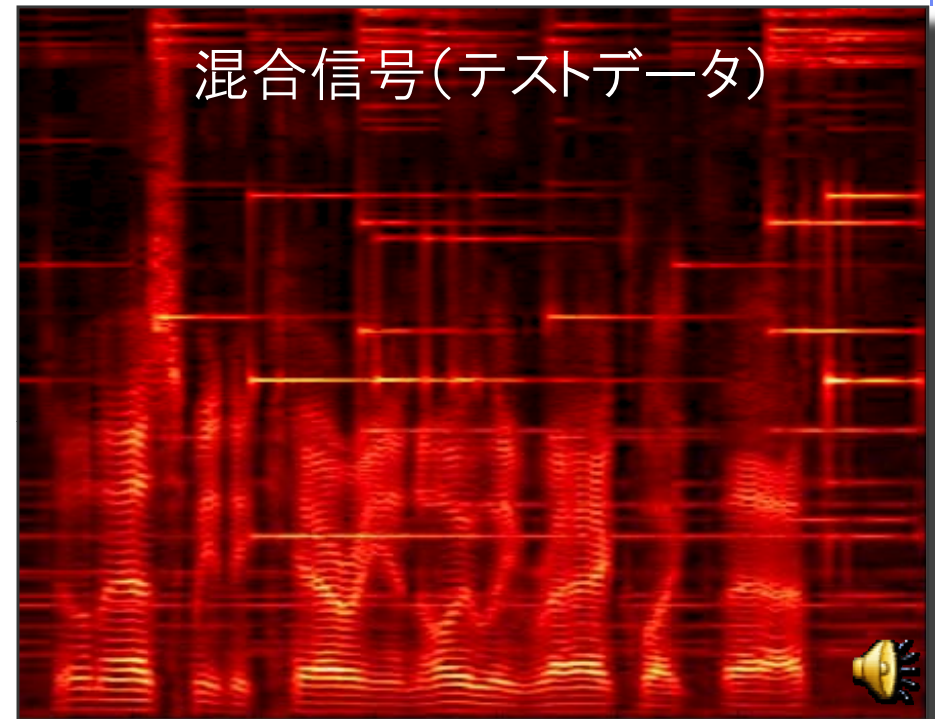


チャイム(学習データ)



Y

混合信号(テストデータ)



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

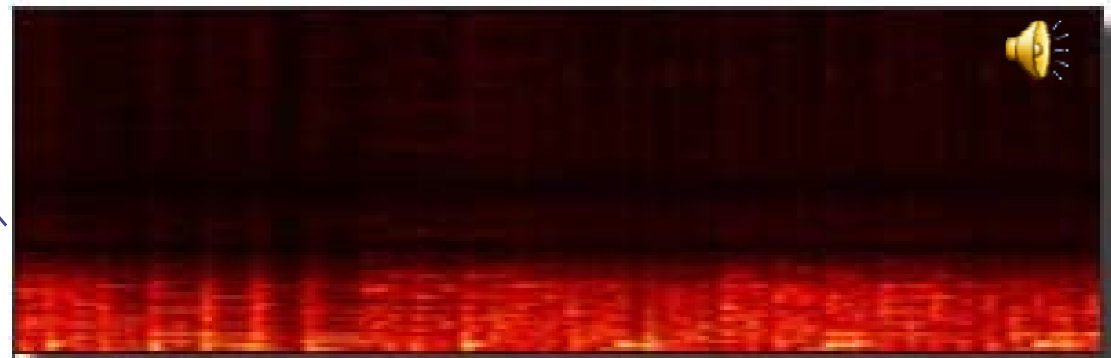
固定

何に使えるのか？(3/5)

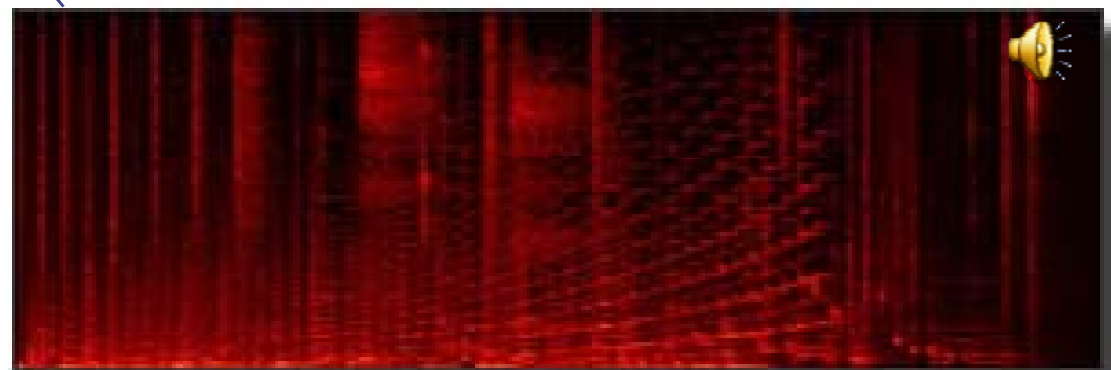
■音の「超解像」

[P. Smaragdis & B. Raj, 2007]

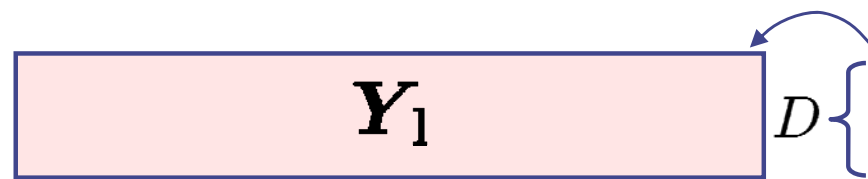
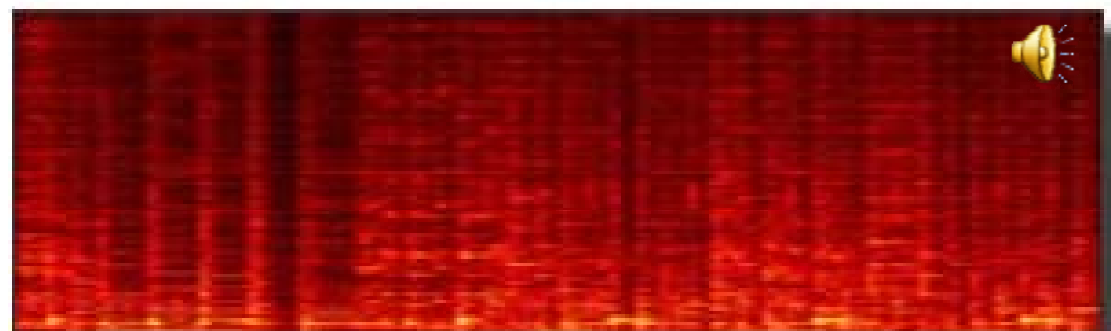
低サンプリングレートの信号



高サンプリングレート信号の学習データ

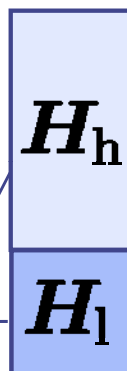


高周波帯域が復元された信号

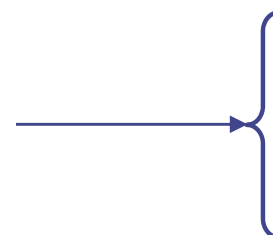


$$Y_1 \simeq H_1 U$$

(固定)



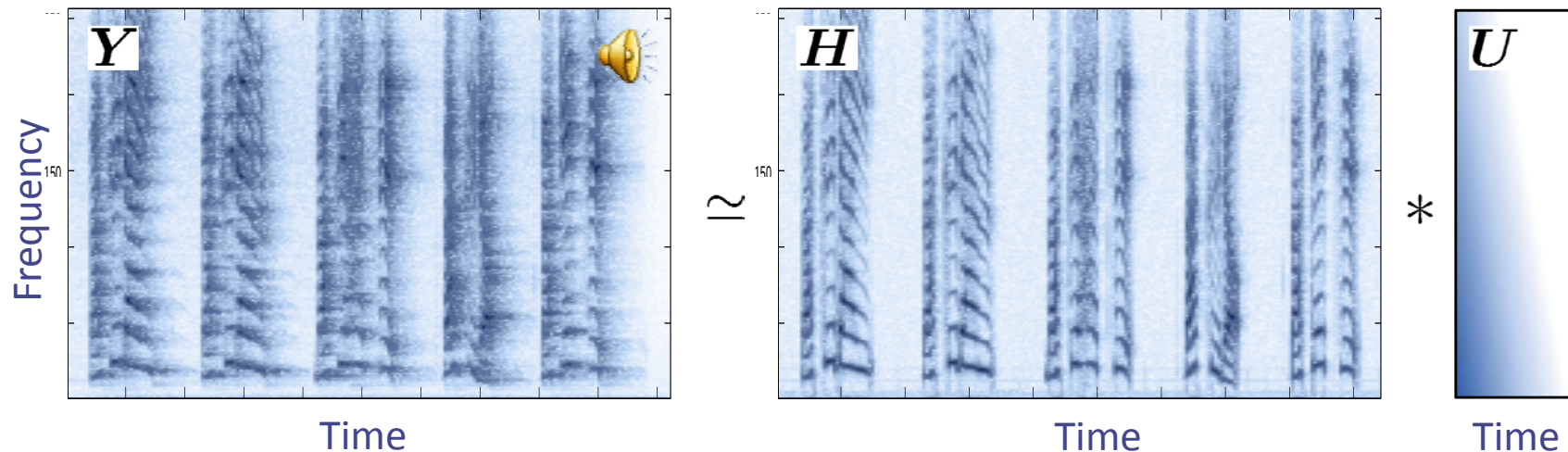
$$H_h U = Y_h$$



何に使えるのか？(4/5)

■ブラインド残響除去 [H. Kameoka et al., 2008]

非負値行列「逆畳み込み」



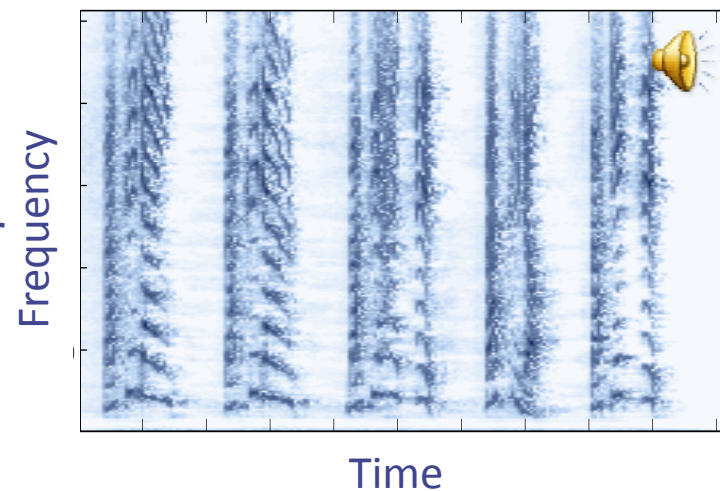
$Y_{\omega,t}$

$\approx$

$$\sum_{\tau} H_{\omega,\tau} U_{\omega,t-\tau}$$

■残響除去音声

非負値行列逆畳み込み
によって求めた実際の H



何に使えるのか？(5/5)

- エコーキャンセラ

[戸上&川口 2009]

- 音響符号化

[Nikunen & Virtanen 2010]

- ブラインド音源分離 (NMFの多チャンネル拡張)

[Ozerov et al. 2010], [Kitano et al. 2010], [Takeda et al. 2011],
[Sawada et al. 2011], [Sawada et al. 2012]

- フォルマントラッキング

[Durrieu et al. 2011]

- 音素特徴量抽出

[Hurmalainen et al. 2011]

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

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

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

$\mathbf{Y}$ 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

■ 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)

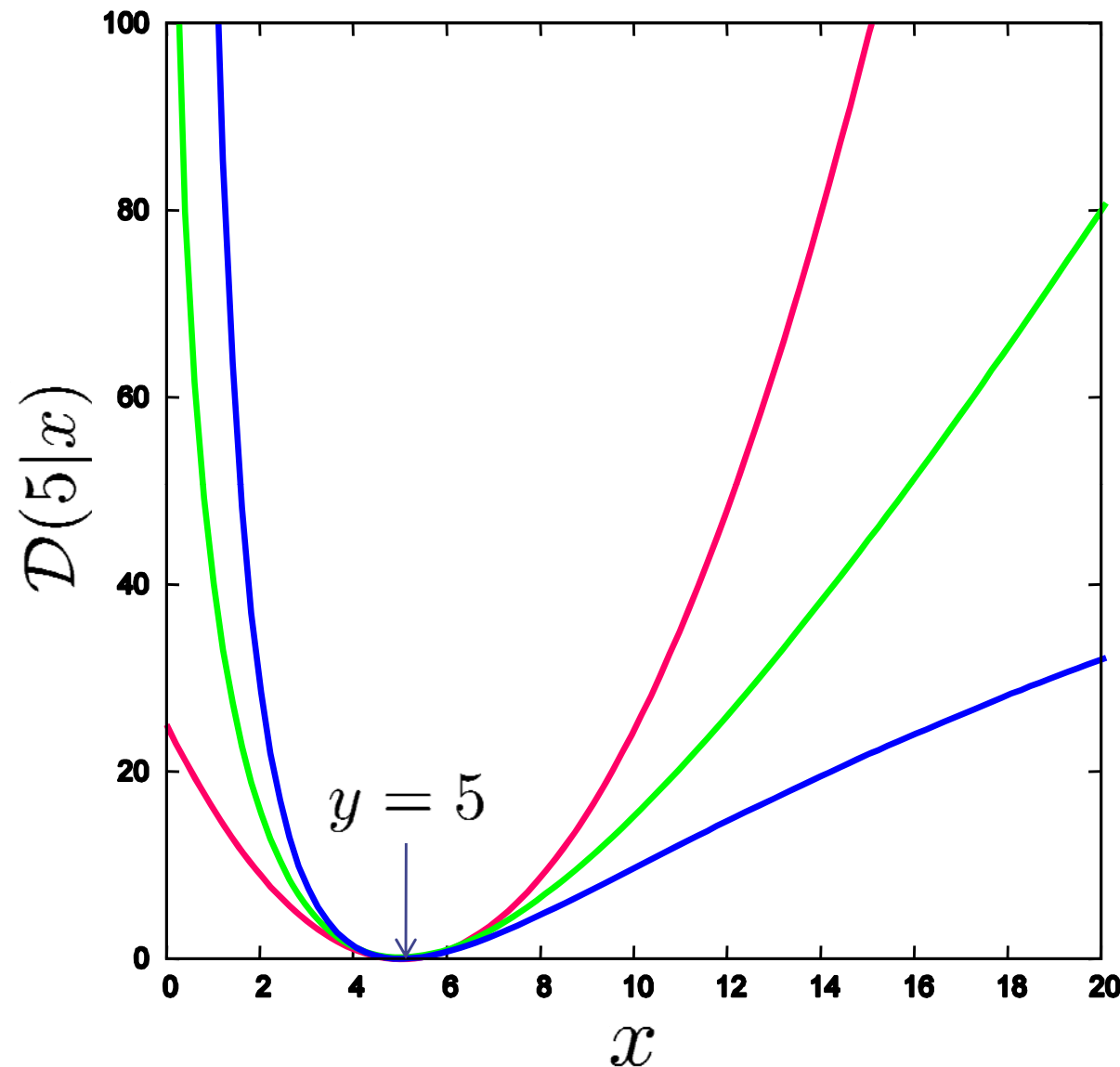
■ 押さえておくべき基本原理

■ 補助関数法

■ 凸不等式(Jensenの不等式)

二乗誤差と | ダイバージェンス

- $\mathcal{D}(y|x)$: x の y からの近さの度合い



— 二乗誤差

$$\mathcal{D}(x|y) = (y - x)^2$$

— | ダイバージェンス

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

— 板倉齋藤距離

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

NMFにおける代表的な最適化規準

■ 二乗誤差

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

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 $\mathbf{Y} = \mathbf{H}\mathbf{U}$ 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

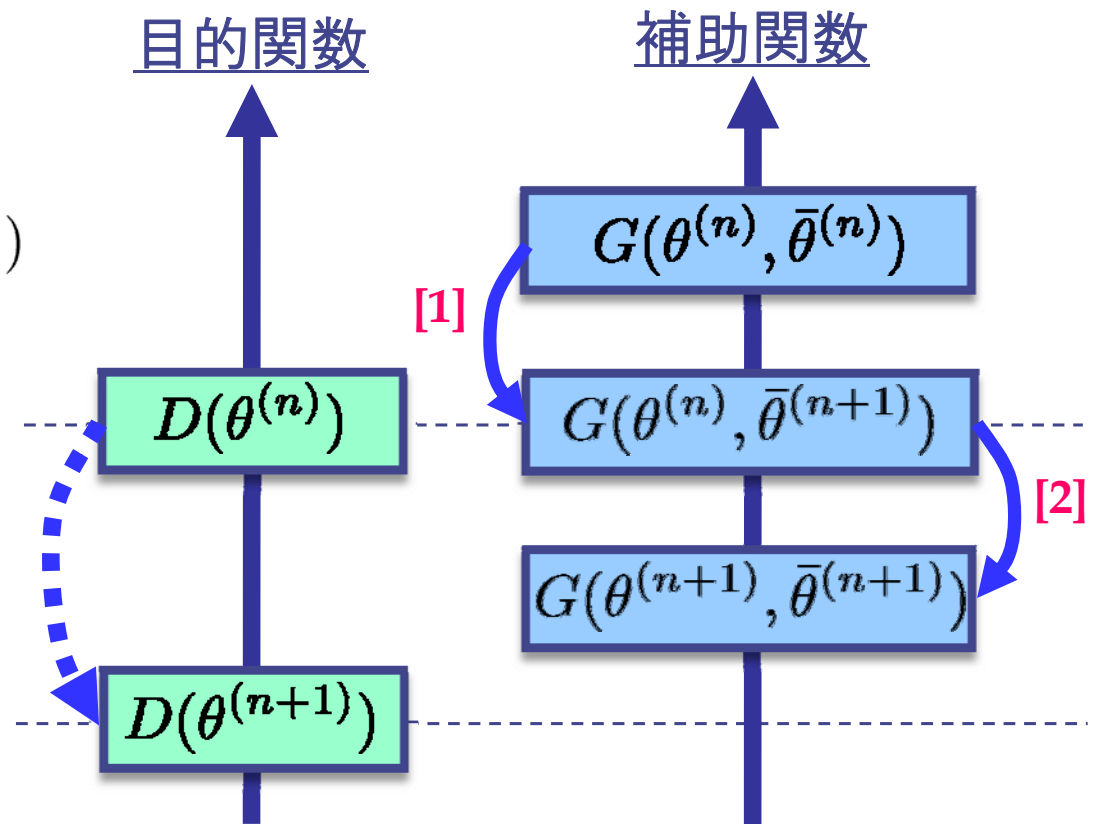
- $D(\theta) = \min_{\bar{\theta}} G(\theta, \bar{\theta})$ を満たす $G(\theta, \bar{\theta})$ を補助関数と定義
- 反復アルゴリズム

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

[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

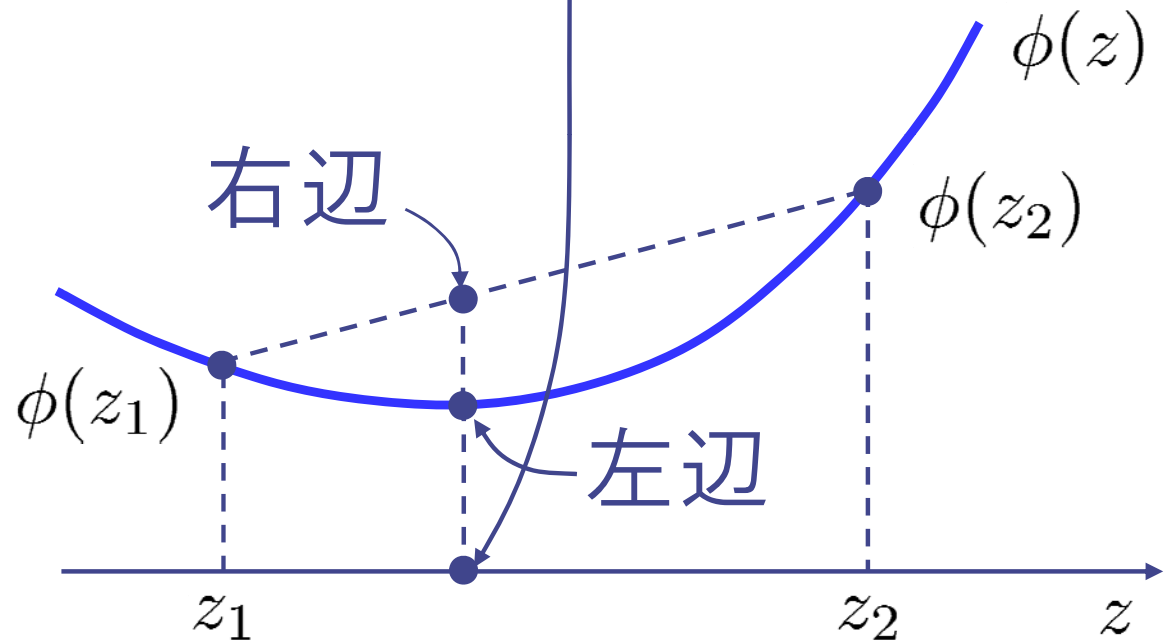
■ $\lambda_i \geq 0, \sum_i \lambda_i = 1$

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

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

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

- 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

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

- 下線部に対してJensenの不等式を立ててみる

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

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

- 補助関数が完成

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

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!

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

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

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

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

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

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

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

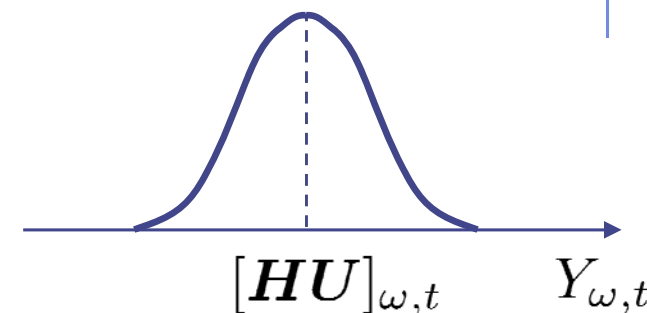
■ Frobeniusノルム規準

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

正規分布

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

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



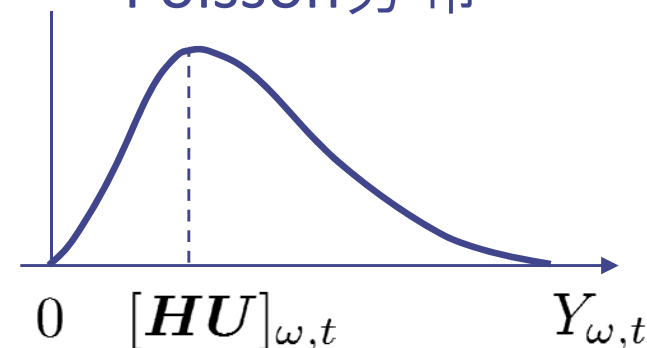
■ ダイバージェンス規準

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

Poisson分布

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

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



板倉齋藤距離規準NMF [Févotte 2008]

- 観測信号をガウス性信号の重ね合わせとしてモデル化
- 各要素信号のパワースペクトログラムがRank-1構造をもつと仮定

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

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

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



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

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



$$\begin{aligned} \mathcal{J} &= \sum_{\omega,t} \mathcal{D}\left(|Y_{\omega,t}|^2 \middle| \sum_k H_{\omega,k} U_{k,t}\right) \\ &= \sum_{\omega,t} \left(\frac{|Y_{\omega,t}|^2}{\sum_k H_{\omega,k} U_{k,t}} - \log \frac{|Y_{\omega,t}|^2}{\sum_k H_{\omega,k} U_{k,t}} - 1 \right) \end{aligned}$$

板倉斎藤距離規準のNMFアルゴリズム [Kameoka 2006]

- 目的関数

$$\mathcal{D}_{\text{IS}}(\mathbf{H}, \mathbf{U}) = \sum_{\omega, t} \left(\frac{|Y_{\omega, t}|^2}{\sum_k H_{\omega, k} U_{k, t}} + \log \sum_k H_{\omega, k} U_{k, t} - \dots \right)$$

- 逆数関数は凸関数 $\Rightarrow$ Jensenの不等式

$$\frac{1}{\sum_i z_i} = \frac{1}{\sum_i \lambda_i \frac{z_i}{\lambda_i}} \leq \sum_i \lambda_i \frac{1}{\frac{z_i}{\lambda_i}} = \sum_i \frac{\lambda_i^2}{z_i}$$

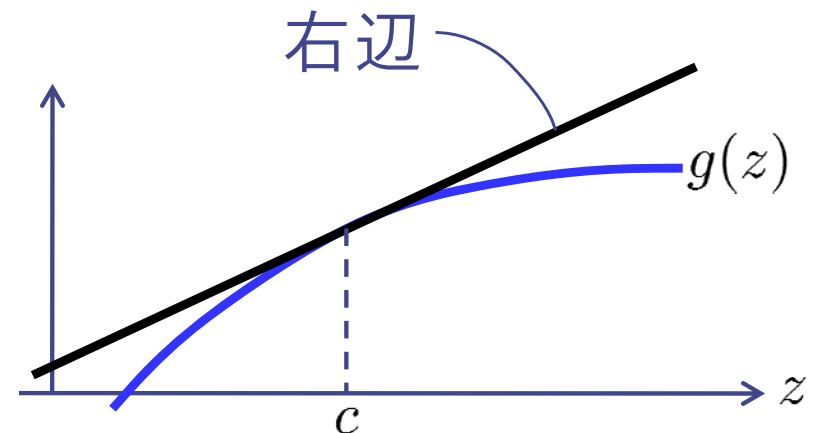
適用 $\Rightarrow \frac{1}{\sum_k H_{\omega, k} U_{k, t}} \leq \sum_k \frac{\lambda_{k, \omega, t}^2}{H_{\omega, k} U_{k, t}}$

- 対数関数は凹関数 $\Rightarrow$ この場合はどうしたら??

$$g(z) \leq g'(c)(z - c) + g(c)$$

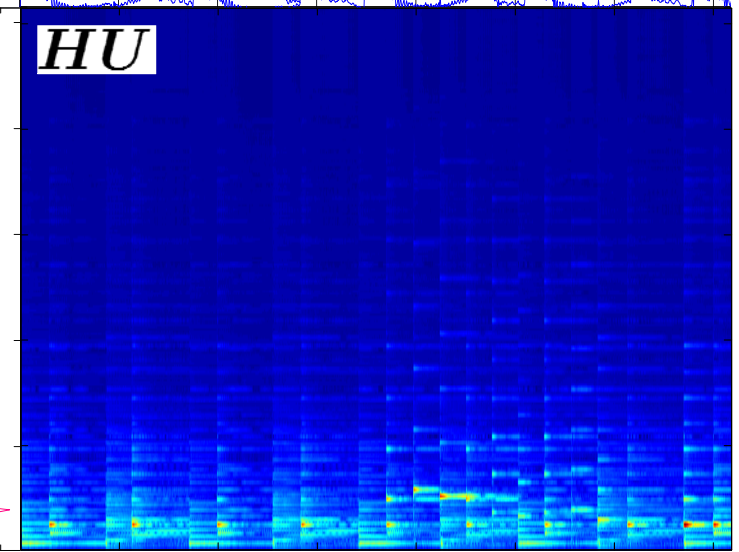
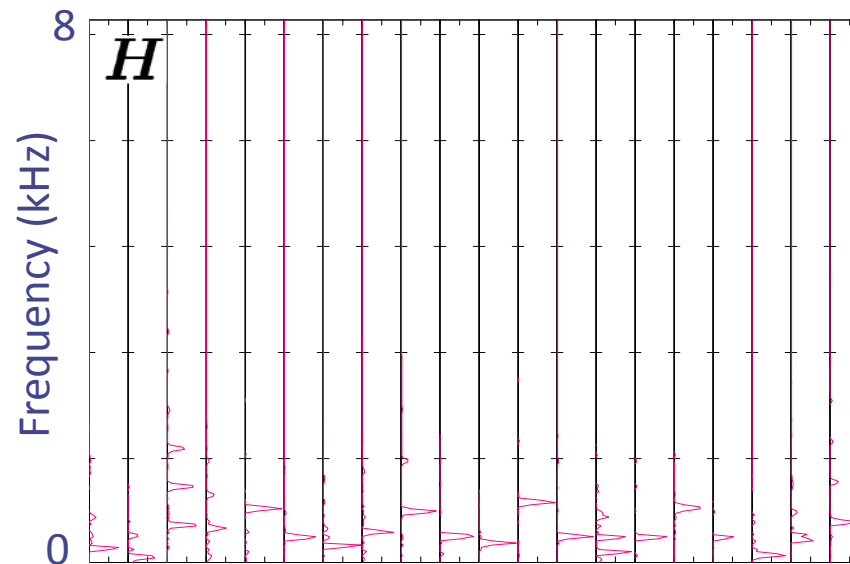
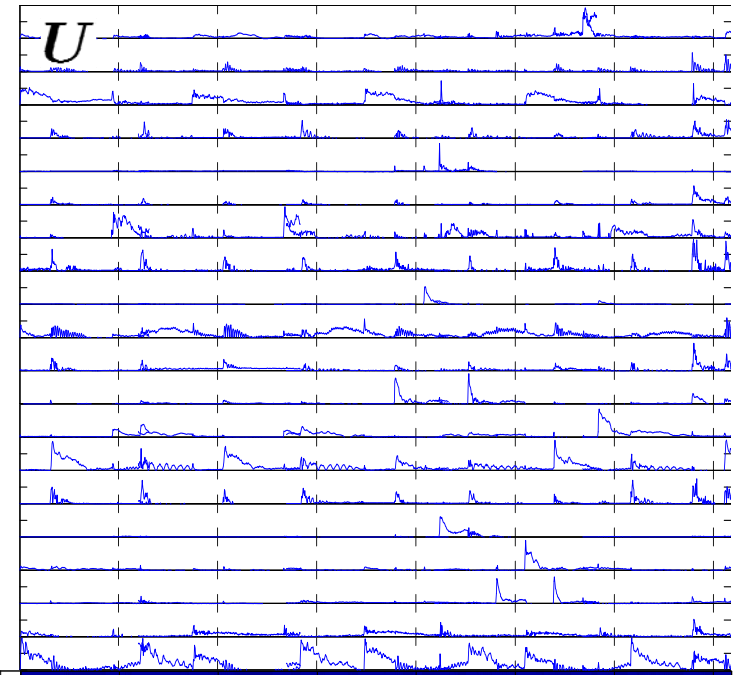
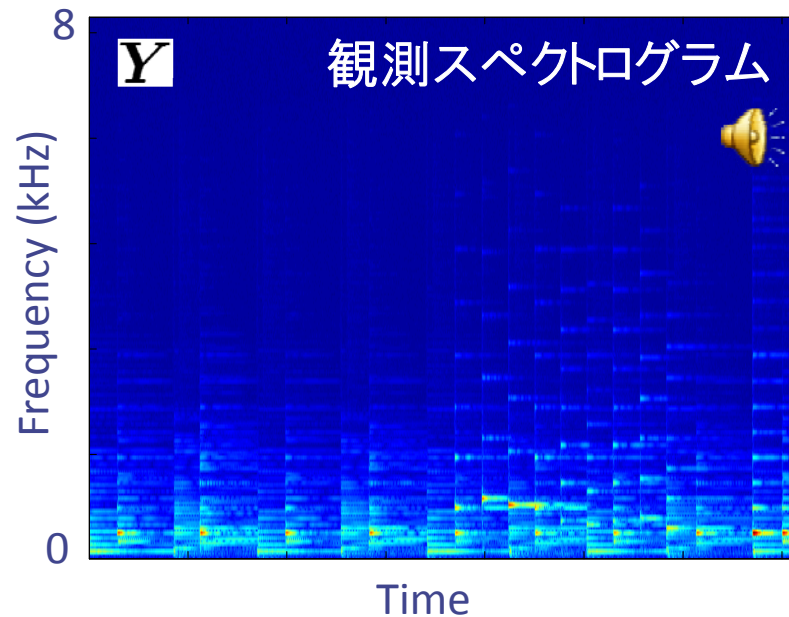
適用

$$\log \sum_k H_{\omega, k} U_{k, t} \leq \frac{1}{c_{\omega, t}} \left(\sum_k H_{\omega, k} U_{k, t} - c_{\omega, t} \right) + \log c_{\omega, t}$$



NMFの規準別の比較

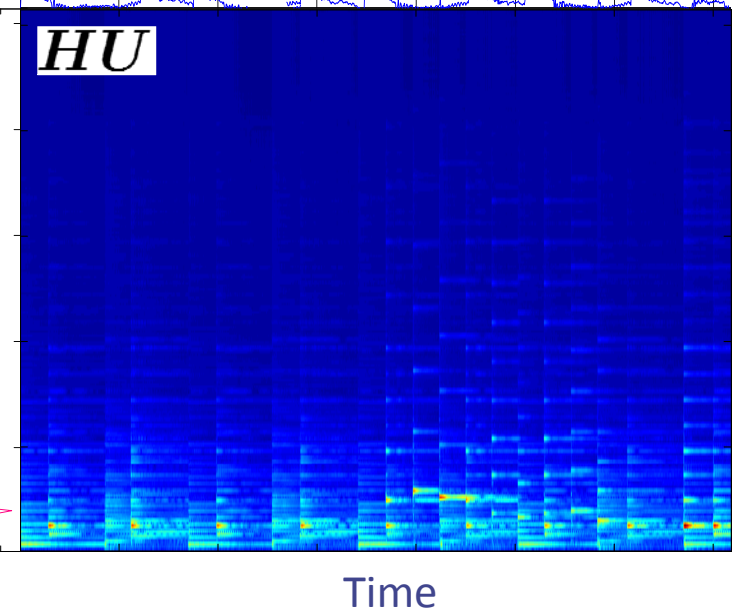
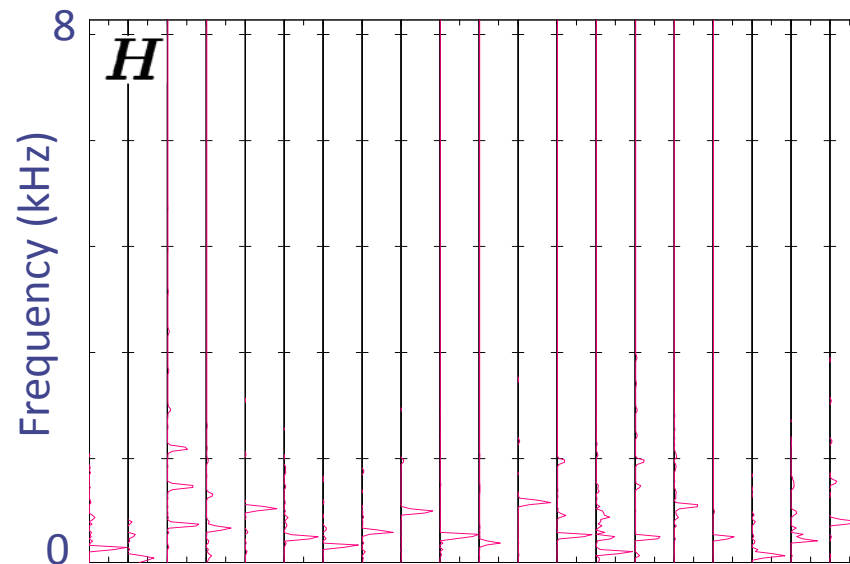
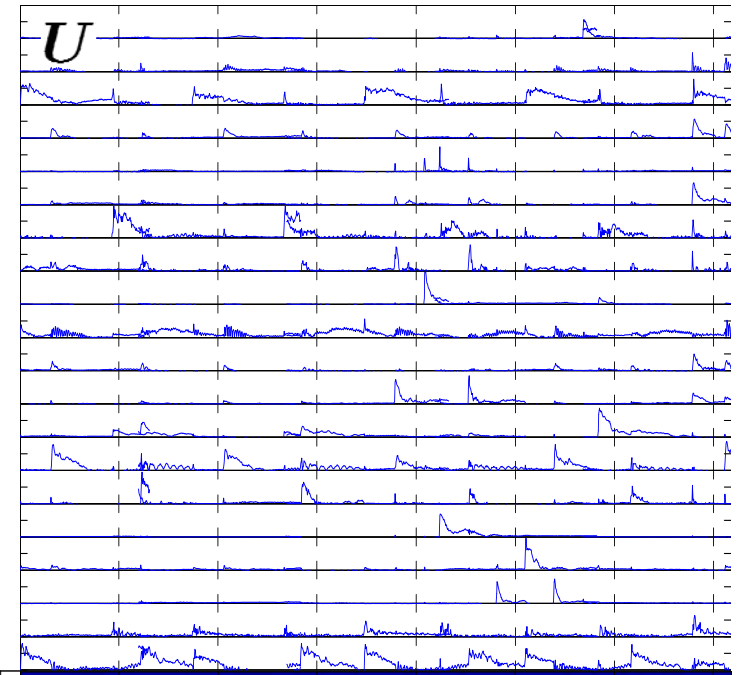
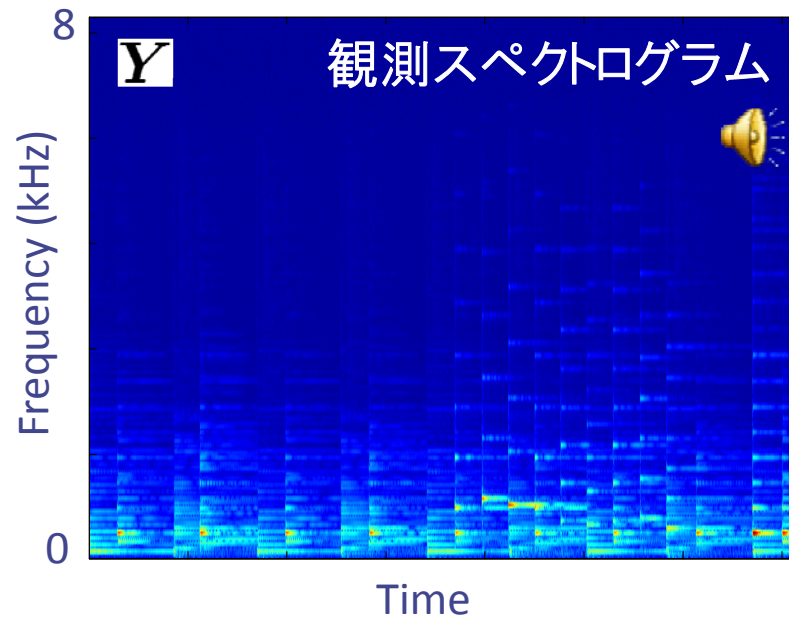
■ 二乗誤差規準



Time

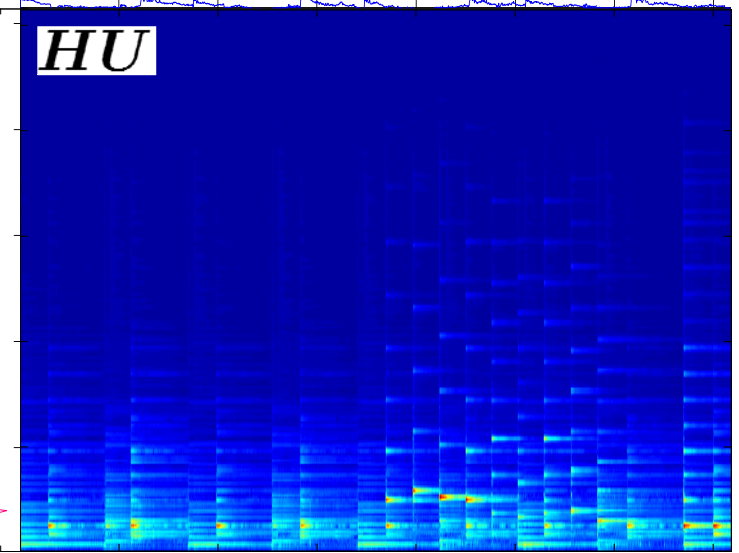
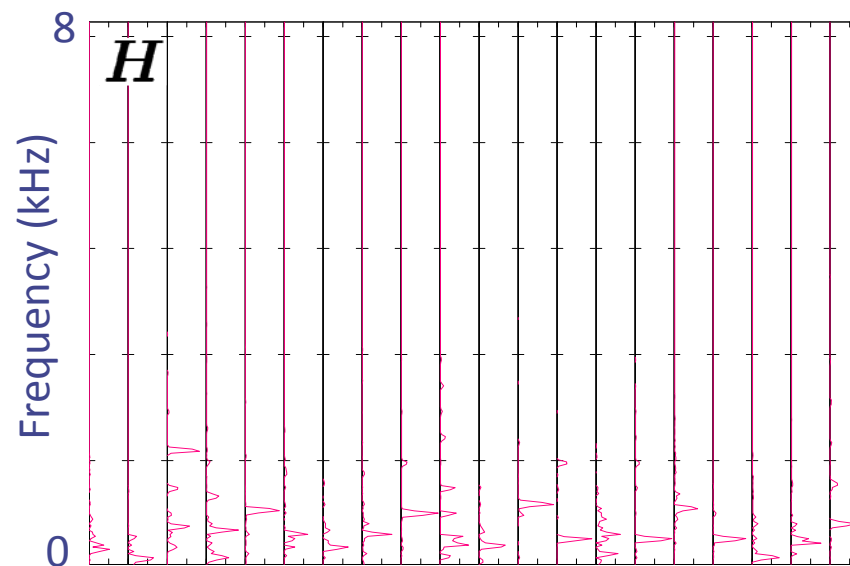
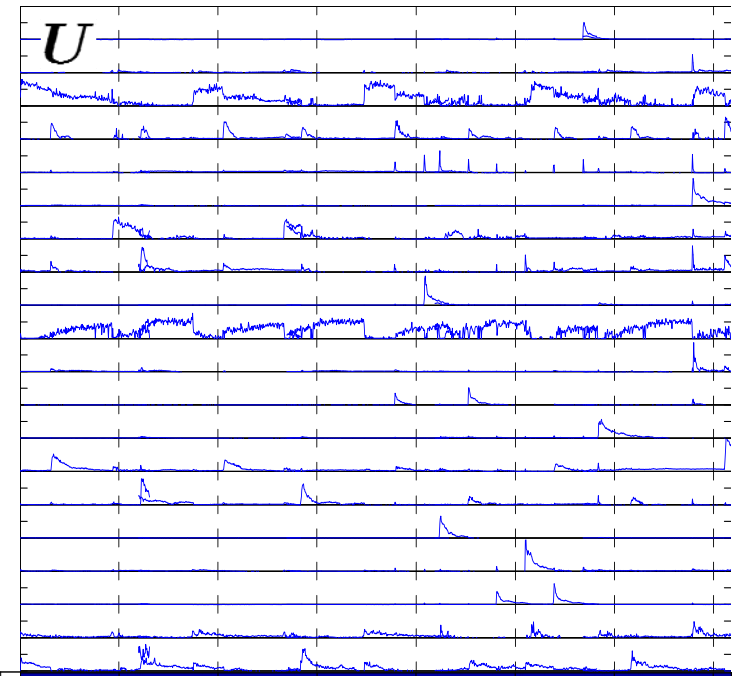
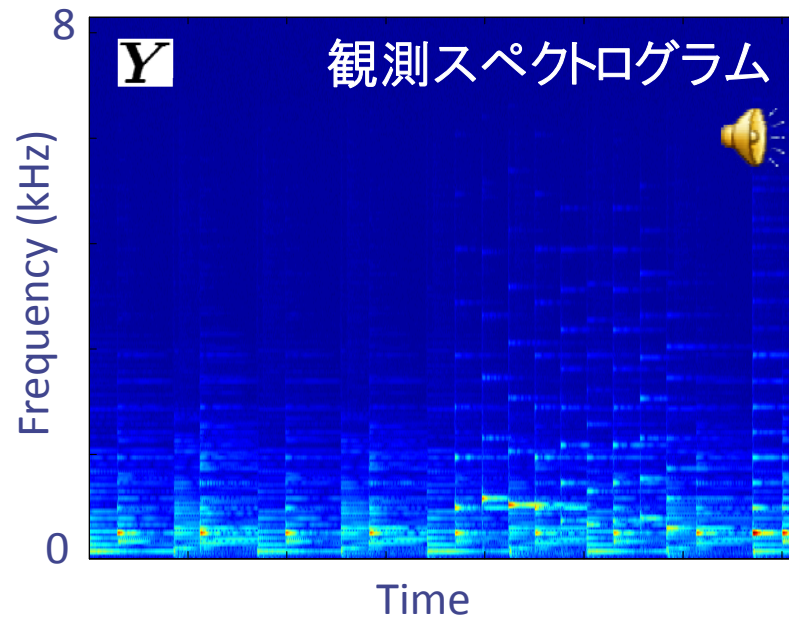
NMFの規準別の比較

■ I ダイバージェンス規準



NMFの規準別の比較

■ 板倉斎藤距離規準



Time

本日の講義のまとめ

■ 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