

ブラインド音源分離

～時空間スモールデータの非ガウス・
低ランクモデリングとその最適化の数理～

東京大学・大学院情報理工学系研究科
猿渡 洋

hiroshi_saruwatari@ipc.i.u-tokyo.ac.jp

目次

- ブラインド音源分離について
 - 定式化, 歴史, 応用
- ブラインド音源分離手法の概観: 何をモデリング?
 - 周波数領域独立成分分析(FDICA)
 - 独立ベクトル分析(IVA)
 - 独立低ランク行列分析(ILRMA)
- アルゴリズム概観: どうやって最適化?
 - 「生成モデル型」vs. 「分離モデル型」?
 - 多チャネルNMFとの関係
- 実験的比較・実用化例
- スパースな生成モデルの導入
- 事前分布・正則化の導入
- まとめ

音の分離:なぜ必要なのか？

■カクテルパーティー効果

- 人の聞き分け能力の模擬
- 補聴器への応用



■音声認識

- マイクロホンと口の間の距離が大きくなるにつれて増大してくる妨害音を抑圧・除去
- 音声対話インターフェイスやロボットへの応用
- 環境音認識・音監視システム

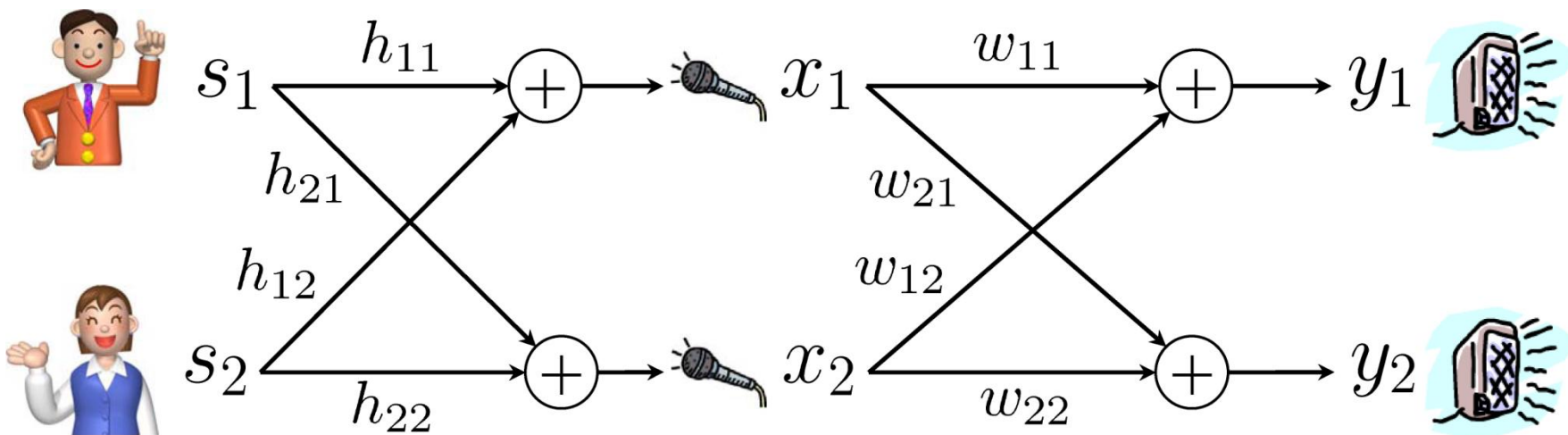


■音楽／楽器音分析

- ミックスダウンされてしまった録音データの解析・エディット
- 自動採譜・ユーザオリエンテッドな音楽情報処理
- VR・ARにおける音響信号の加工

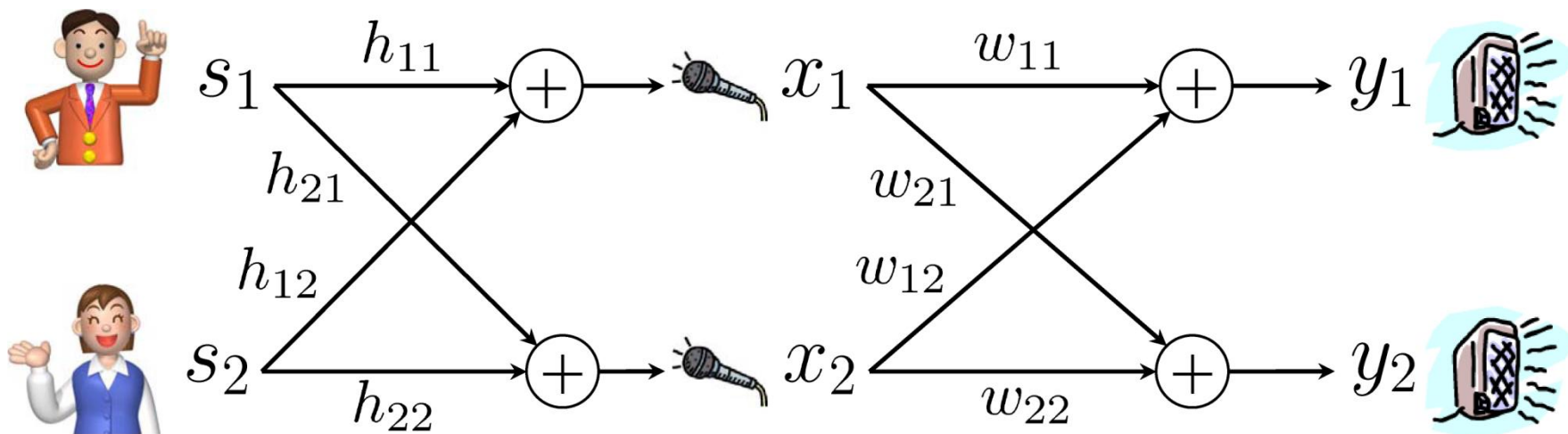
ブラインド音源分離 (Blind Source Separation)

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



ブラインド音源分離 (Blind Source Separation)

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



実は上記は**2つのことを同時に推定**している

- [空間] 統計的に独立な音源の分類問題 (分離行列 W の推定)
- [信号] 各音源が属する確率分布 $p(y)$ の推定問題

上記を閉形式で解く方法は存在せず凸問題でもない⇒**大変困難!**

何をどうモデル化して最適化するのか

Multichannel NMF: Full-rank spatial model + NMF-based source model

Spatial rank-1 approximation

FDICA

- Source Model:
Super-Gauss time series
- Spatial Constraint:
Up to permutation solver
 - Frequency correlation
(no spatial constraint)
 - Source direction
(need spatial separability)
- Optimization algorithm:
 - Natural gradient
(conv. is not guaranteed)

IVA

- Source Model:
Rank-1 in spectrogram
 - Time: sparse
 - Frequency: **non-sparse**
- Spatial Constraint:
None
- Optimization algorithm:
 - Natural gradient
 - Auxiliary func. method
(conv. guaranteed)

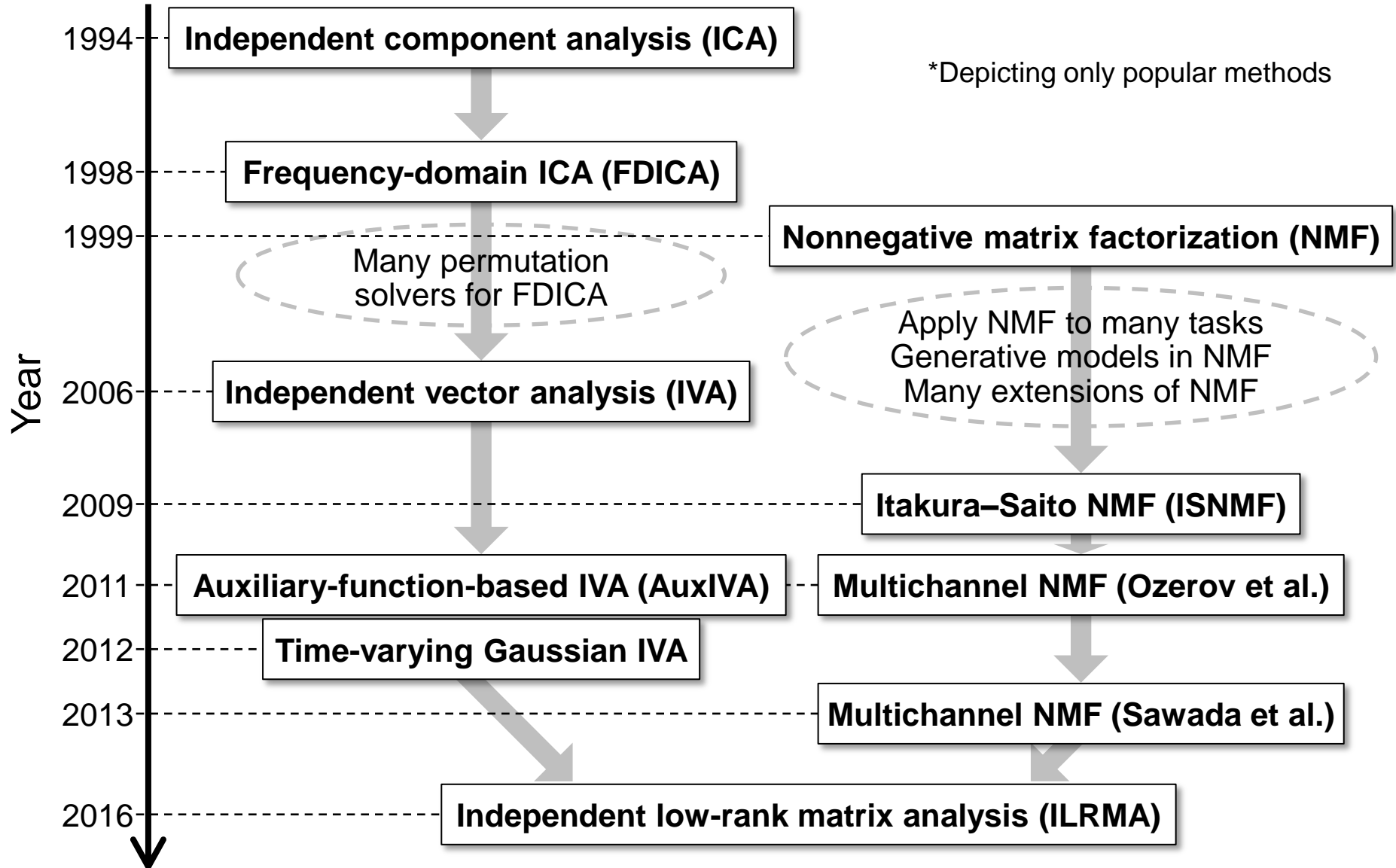
ILRMA

- Source Model:
Low-rank in spectrogram
 - Time: sparse
 - Frequency: sparse
- Spatial Constraint:
None
- Optimization algorithm:
 - Auxiliary func. method
(conv. guaranteed)

いろいろなやり方があり、いろいろ関連しあっているようだ…

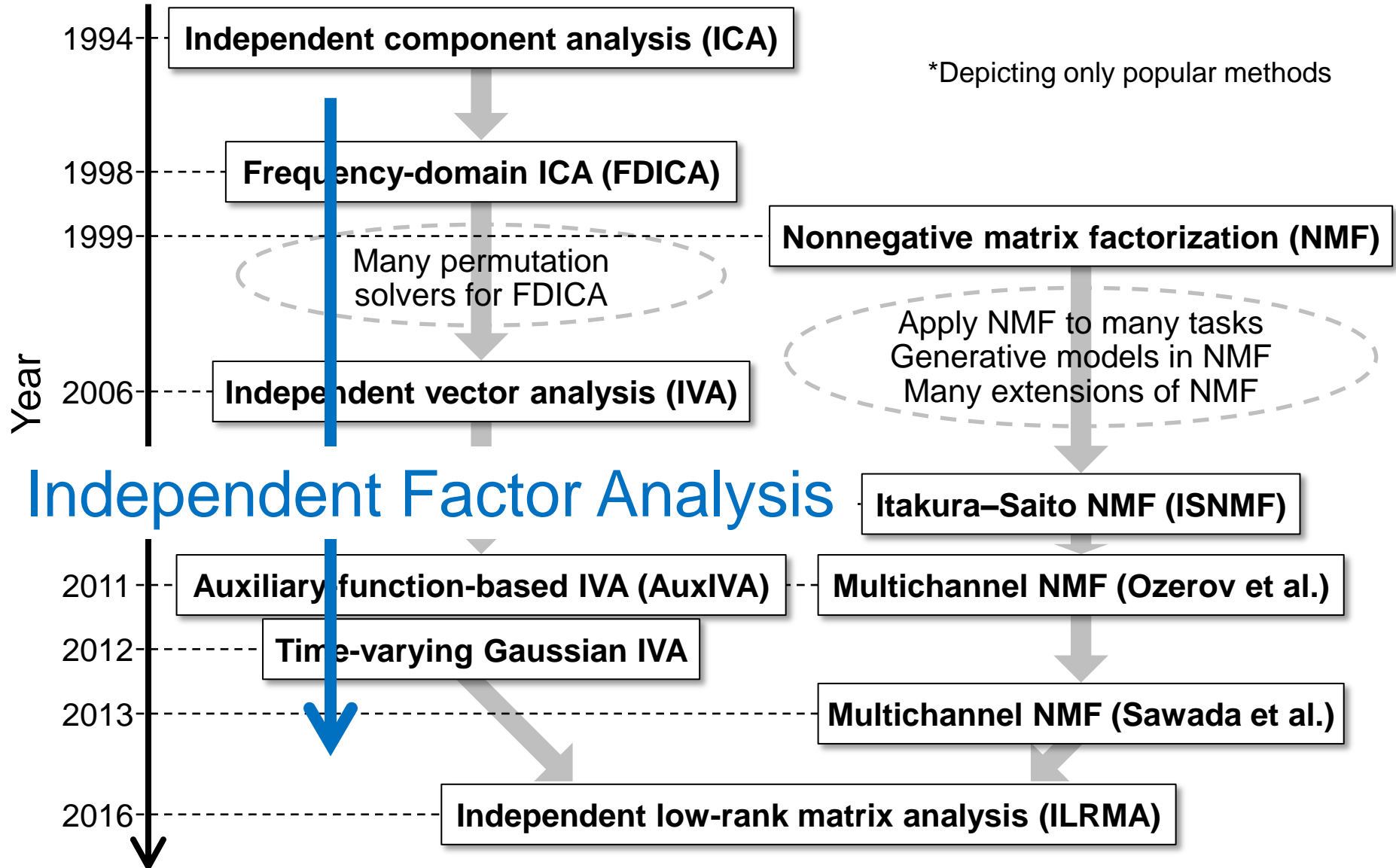
歴史をひもといてみると...

- Basic theories and their evolution



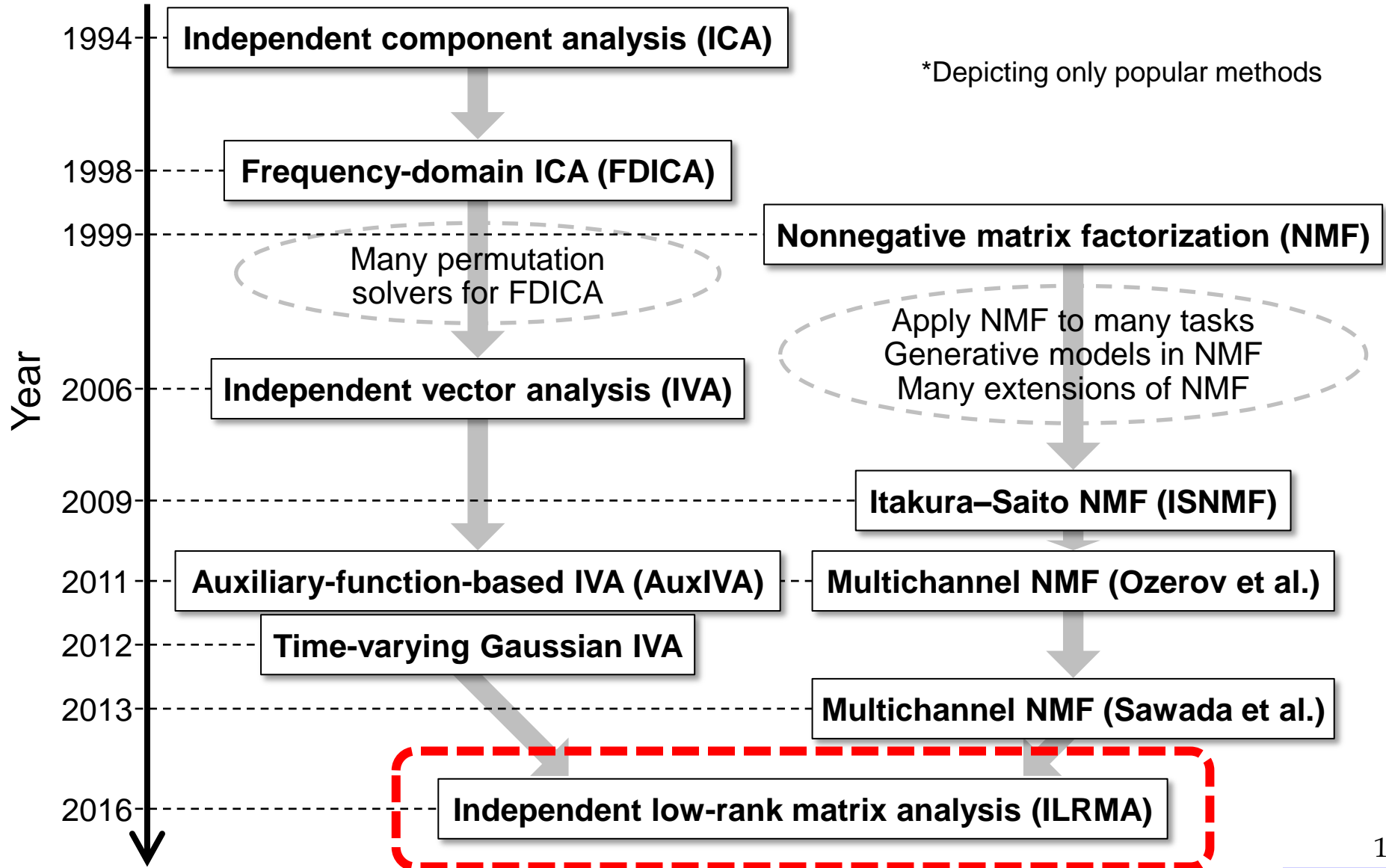
歴史をひもといてみると...

- Basic theories and their evolution



歴史をひもといてみると...

- Basic theories and their evolution



周波数領域 独立成分分析 (FDICA)

BSSを有名にした立役者

最尤推定法 (Infomax) [Bell et al., 1995]

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

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

- 線形変換と確率密度関数

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

- 音源信号の独立性と非Gauss性を仮定

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

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

最尤推定法

- 目的関数:

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

- 勾配法により \mathbf{W} を反復的に更新

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

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

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

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

$\phi(y)$ の具体形

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

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

最尤推定法

- 目的関数: 周辺エントロピー和最小化 = 非ガウス化の促進

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

- 勾配法により \mathbf{W} を反復的に更新

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

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

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

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

$\phi(y)$ の具体形

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

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

周波数領域ICAによる音響信号BSS(1998~)

• 周波数領域ICA (frequency-domain ICA: FDICA)

[Smaragdis, 1998], [Ikeda-Murata, 1998], [Saruwatari+, 2000], etc.

- 音響信号は空間伝達関数(時間遅れ・畳み込み)を経て混合
- 時間領域では逆畳み込みの問題になってしまう
- 上記を解決するため、短時間フーリエ変換によって時間-周波数領域へ変換し、そのドメインで**周波数別**に分離行列 W をかける

周波数帯域別にICAを行うと...

ICAでは因子の順番は不問

⇒ 周波数帯域毎に分離信号が入れ替わってしまう

パーミュテーション問題

- 何らかの事前仮説が必要

「**プロジェクションバック**」法という

ICAでは因子の大きさは不問

⇒ 周波数帯域毎に分離信号の利得がバラバラに...

スケーリング問題

- 基準センサへの射影で解決
- $W^{-1}\text{diag}\{0, \dots, 1, \dots, 0\}$ の乗算

目的音出力番号

IVAの登場

パーミュテーションフリー なICAを目指して

周波数成分間の高次相関を用いたBSS(2006~)

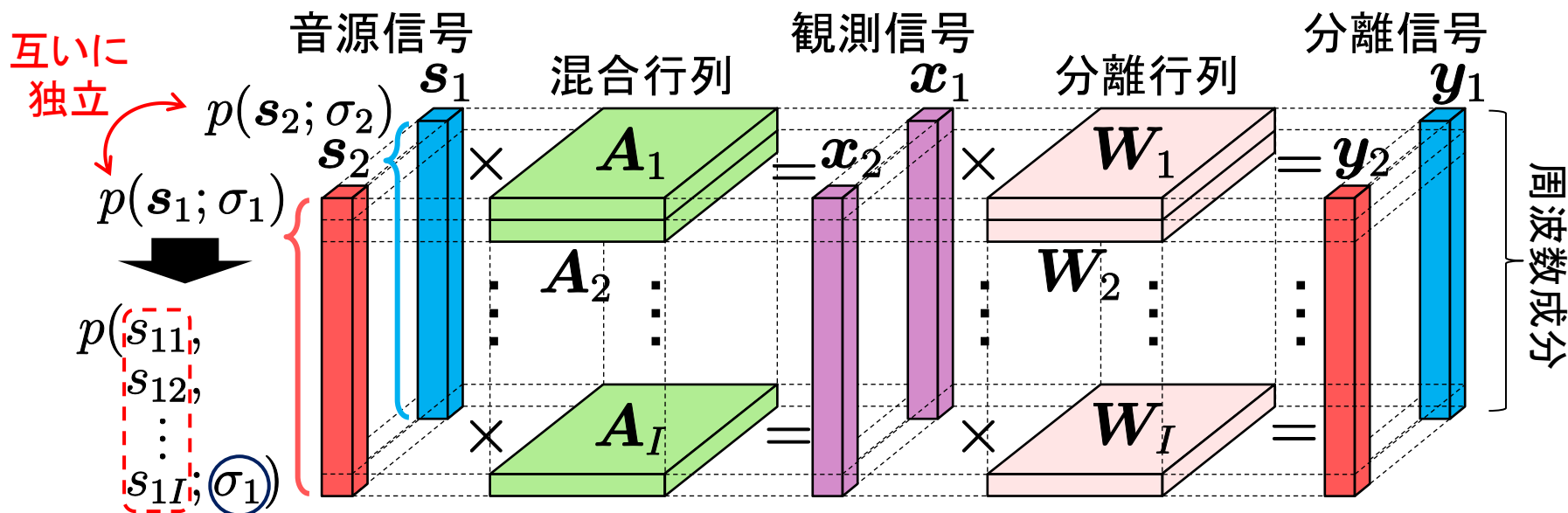
独立ベクトル分析 (independent vector analysis: IVA)

[Hiroe, 2006], [Kim, 2006], [Kim, 2007]

– FDICAの多変量拡張手法, パーミュテーション問題を回避

– 球対称な多変量(多次元)分布モデルの最尤推定

- ICAにおける音源の事前分布 $p(s_n)$ を多次元分布 $p(s_{n1}, \dots, s_{nI}; \sigma_n)$ に拡張
- 球対称性: 無相関な周波数成分間で同じ分散を持つ多変量分布



周波数成分間 共通の分散
は無相関だが を持つ
依存性をもつ

IVAの仮定する周波数成分間の依存性

- FDICAとIVAの違いはコントラスト関数のみ
 - コントラスト関数: パラメタの対数尤度のgradient

$$\varphi_i(s_{n1}, \dots, s_{nI}) = -\frac{\partial p(s_{n1}, \dots, s_{nI}; \sigma_n)}{\partial s_{ni}}$$

- 「音源の事前分布が一次元か多次元か」のみ

- IVAの仮定する音源の事前分布 $p(\mathbf{s}_n)$

- ラプラス分布 (音声信号のモデルとして有効な **スパース分布**)

周波数毎に独立な事前分布

$$p(\mathbf{s}_n) = \prod_{i=1}^I p(s_{ni}) = \alpha \prod_{i=1}^I \exp\left(-\frac{|s_{in} - \mu_{in}|}{\sigma_{ni}}\right)$$

周波数間で高次相関をもつ事前分布

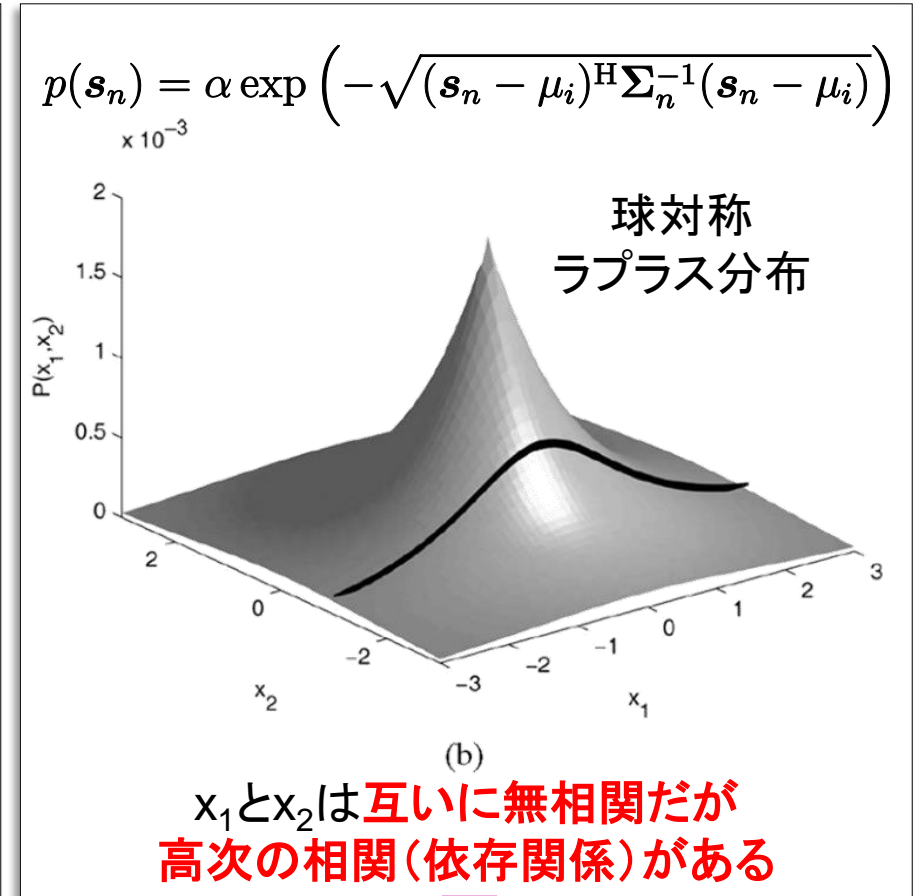
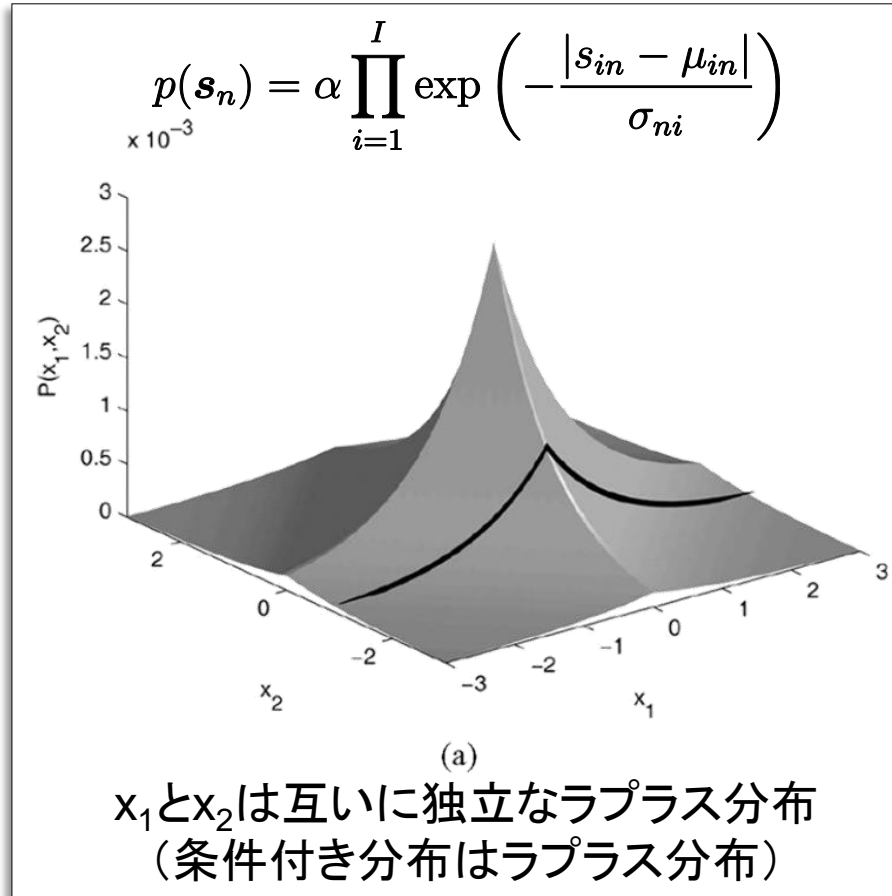
$$p(\mathbf{s}_n) = \alpha \exp\left(-\sqrt{(\mathbf{s}_n - \boldsymbol{\mu}_i)^H \boldsymbol{\Sigma}_n^{-1} (\mathbf{s}_n - \boldsymbol{\mu}_i)}\right)$$

分散共分散行列

- 後者は $\boldsymbol{\Sigma}_n = \mathbf{I}$ (互いに無相関) の場合でも, s_{ni} が互いに依存

IVAの仮定する周波数成分間の依存性

- 図は [Kim, 2007] より引用



周波数間を一つにまとめる効果が生じる
(2次元スペクトログラム分布の先駆けであった)

独立低ランク行列分析

NMFと融合した多チャネルBSS

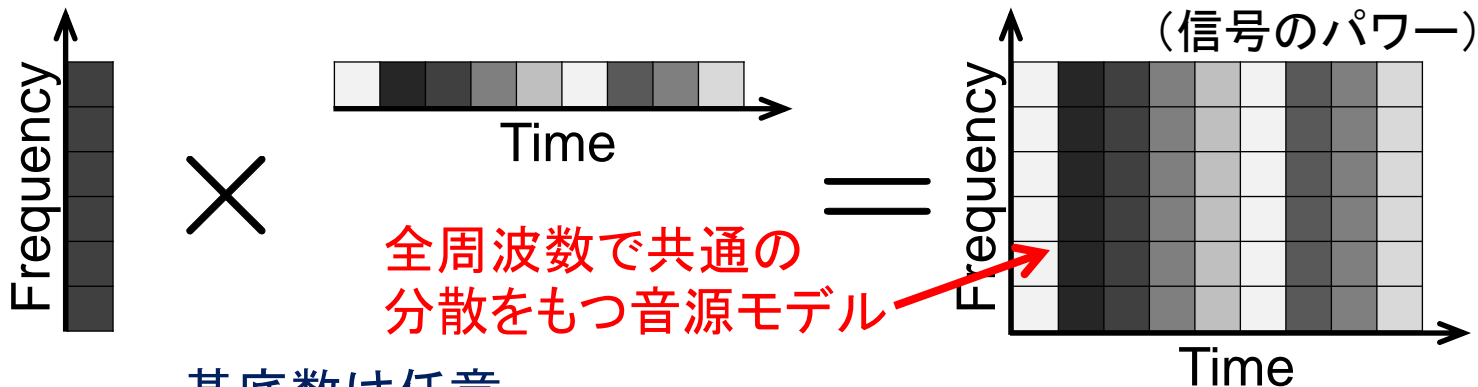
低ランク性を導入したBSS(2016~)

- 独立低ランク行列分析 (independent low-rank matrix analysis: ILRMA)

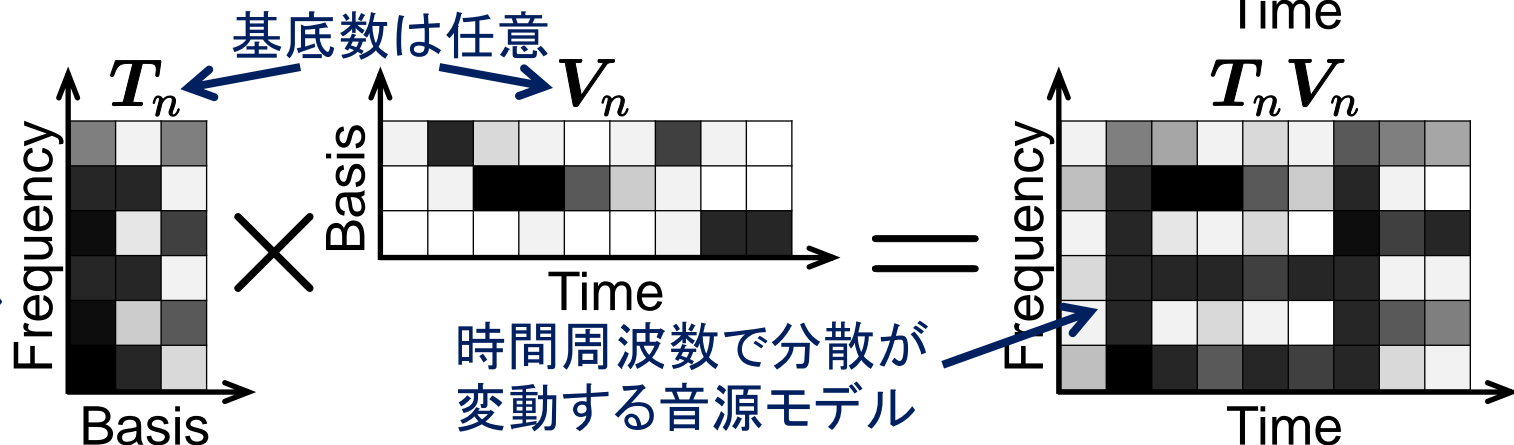
[Kitamura, Saruwatari+, IEEE/ACM Trans. ASLP 2016]

- 独立音源の詳細な時間周波数構造(のパワー)を**低ランク非負値行列**として捉えながら線形分離フィルタを学習

IVAの
音源モデル



ILRMAの
音源モデル



- 板倉斎藤擬距離NMF (ISNMF)に対応する音源モデルを想定

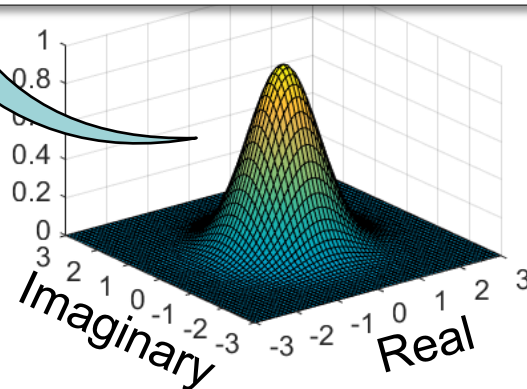
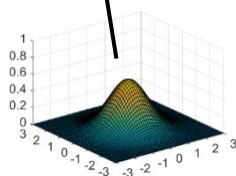
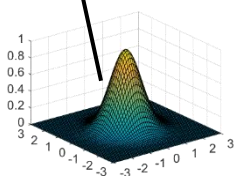
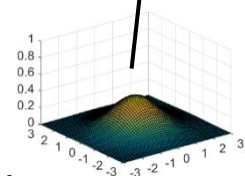
板倉斎藤擬距離基準NMFにおける生成モデル

- X を複素スペクトログラムとする(STFTしたそのもの)
 - 各時間周波数要素は複素要素 c_{ijk} を K 個足し合わせたもの

とある時間周波数要素 $x_{ij} = \sum_{k=1}^K c_{ijk}$

平均0, 分散 $t_{ik}v_{kj}$ の複素球対称ガウス分布
 $c_{ijk} \sim \mathcal{N}_c(0, t_{ik}v_{kj})$

$$x_{ij} = c_{ij1} + c_{ij2} + \dots + c_{ijK}$$



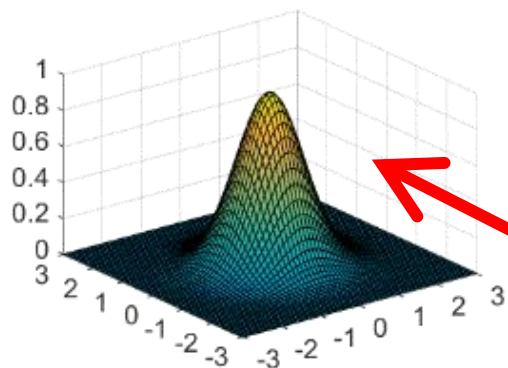
これらの複素ガウス分布は互いに独立 (i.i.d.)

- 複素ガウス分布の線形結合なので x_{ij} も複素ガウス分布
 - ガウス分布の再生性 (α -安定分布)
 - x_{ij} の複素ガウス分布の分散は $\sum_{k=1}^K t_{ik}v_{kj}$
- 分散が時間周波数で変動する複素ガウス分布が生成モデル

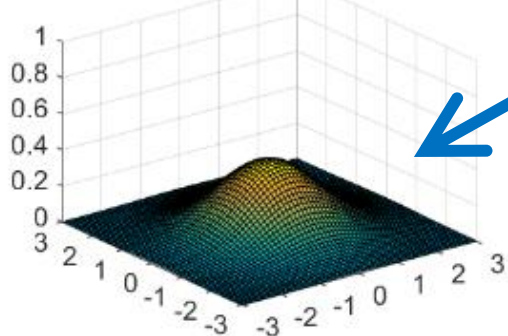
板倉斎藤擬距離基準NMFにおける生成モデル

- パワースペクトログラムは複素ガウスの分散に対応

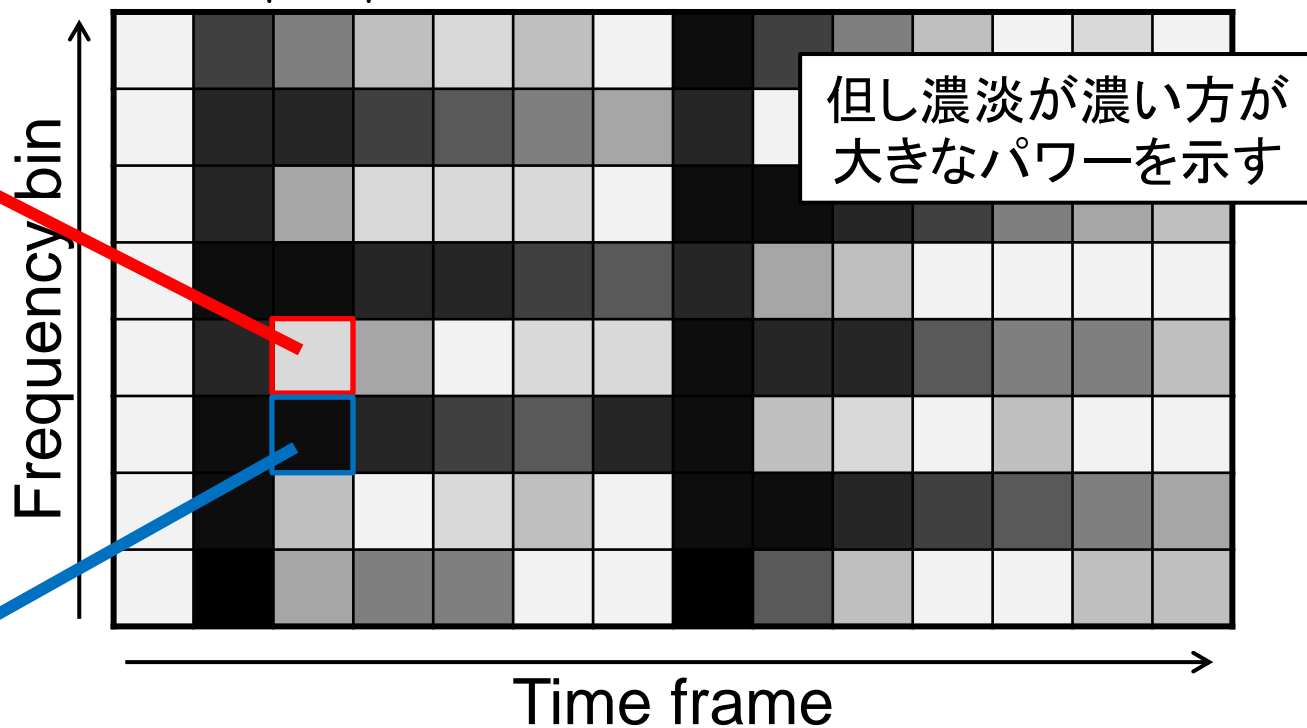
$|X|^2$: パワースペクトログラム



パワーが小＝分散が小
殆ど0付近の複素数しか
生成しない



パワーが大＝分散が大
大きな振幅の複素数も
生成しうる



各時間周波数で分散が変動する複素ガウス分布

巨視的(マクロ)に考えると分散が逐一変動する為、スペクトログラム全体の密度分布はスーパーガウシアン(カートシスがガウス分布より大)な分布になっている

ILRMAのコスト関数と更新則

- ILRMAのコスト(対数尤度)関数

$$\mathcal{J} = \sum_{i,j} \left[\sum_m \log \sum_k z_{mk} t_{ik} v_{kj} + \sum_m \frac{|y_{ij,m}|^2}{\sum_k z_{mk} t_{ik} v_{kj}} - 2 \log |\det \mathbf{W}_i| \right]$$

主なパラメータ

空間モデル

\mathbf{W}_i : 分離フィルタ行列

$y_{ij,m} = \mathbf{w}_{i,m}^H \mathbf{x}_{ij}$: 分離信号

音源モデル

t_{ik} : 基底行列要素

(TV : K ランクNMFモデル)

v_{kj} : アクティベーション行列要素

z_{mk} : 分割関数(基底の共有を許す重み)

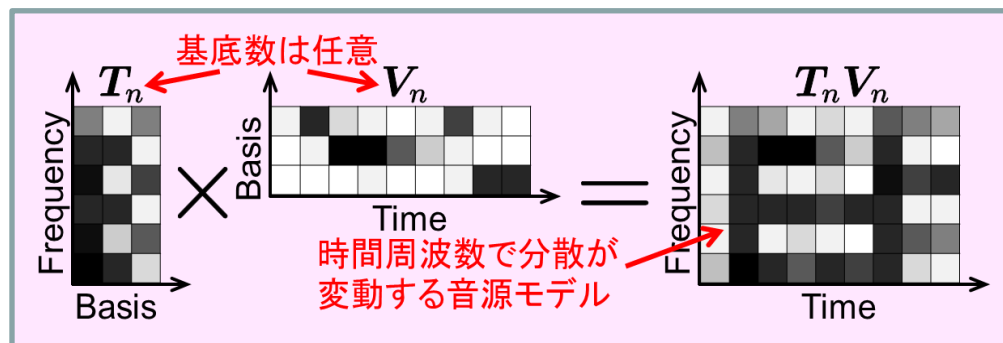
ILRMAのコスト関数と更新則

- ILRMAのコスト(対数尤度)関数→これを最小化

$$\mathcal{J} = \sum_{i,j} \left[\sum_m \log \sum_k z_{mk} t_{ik} v_{kj} + \sum_m \frac{|y_{ij,m}|^2}{\sum_k z_{mk} t_{ik} v_{kj}} - 2 \log |\det \mathbf{W}_i| \right]$$

音源の低ランク性コスト関数
(ISNMFモデルの推定に寄与)

音源の独立性コスト関数
(空間モデル W の推定に寄与)



$$p(\mathbf{y}) = p(y_1) p(y_2) \dots p(y_m)$$

となる W を推定

両者を交互にMajorization-Minimization(補助関数)アルゴリズムで反復最小化

- ✓ コスト値の単調減少性を保証(勾配法には無い特徴)
- ✓ 高速かつ安定な求解法を実現(従来の多入力NMFと比較して2ケタ速い)

アルゴリズム概観

「生成モデル」 vs. 「分離モデル」

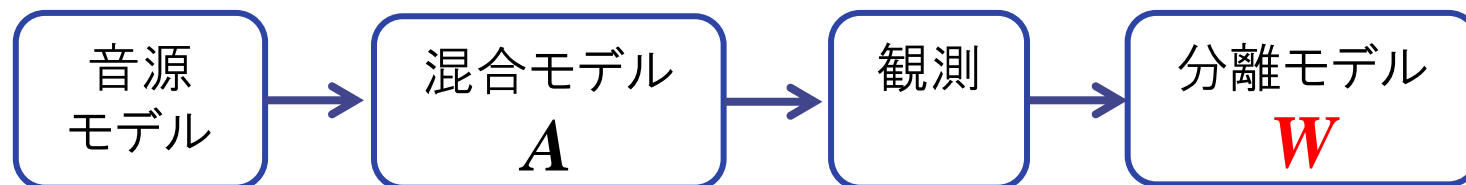
音源分離問題における双対性

生成モデル型



- 音源と混合モデル A のパラメータを同時に推定
- 混合系が劣決定の場合でも適用可能
- 多チャネルNMFなどに利用されるモデル

分離モデル型



- 音源と分離モデル $W=A^{-1}$ のパラメータを同時に推定
- 混合系がランク1空間混合の場合に適用可能
- ICA, IVA, ILRMA等利用されるモデル

何が違うのか？どちらが良いのか？

FDICAにおける最適化: 自然勾配法

- 勾配法では \mathbf{W} の逆行列計算が厄介

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

- 自然勾配 (Natural Gradient) [Amari, 1997]

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

分離フィルタ
のみの更新

- 逆行列計算が不要 (音響系では「逆フィルタ」計算が不要！)
- 混合系の推定よりも安定
- ビームフォーマとの類似性
- アルゴリズムは以下のとおり

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

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

IVAにおける最適化：補助関数法でWを求める

※首都大・小野先生より参考資料を頂きました。感謝いたします。

■ 解くべき最適化問題

$$J(\Theta) \rightarrow \min$$

■ 補助関数

$$Q(\Theta, \tilde{\Theta}) \geq J(\Theta)$$

■ 反復更新則

補助変数の更新 (E-step)

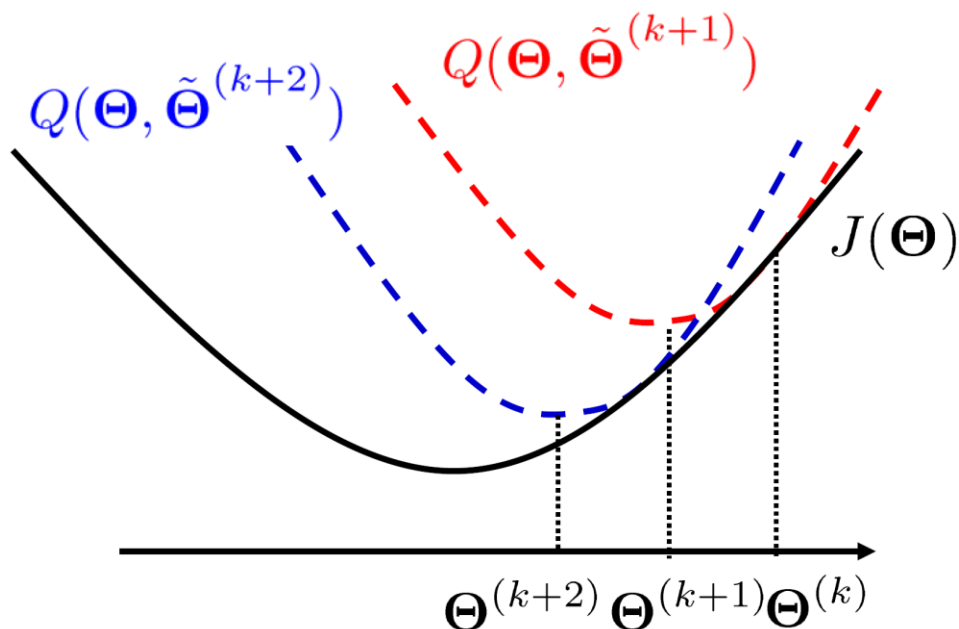
$$\tilde{\Theta}^{(k+1)} = \operatorname{argmin}_{\tilde{\Theta}} Q(\Theta^{(k)}, \tilde{\Theta})$$

パラメータの更新 (M-step)

$$\Theta^{(k+1)} = \operatorname{argmin}_{\Theta} Q(\Theta, \tilde{\Theta}^{(k+1)})$$

■ 補助関数の例(絶対値関数の最小化)

$$|x| \leq \frac{1}{2x_0} |x|^2 + \frac{x_0}{2} \quad \Rightarrow \quad \begin{array}{l} \text{二次関数の和も二次関数なので、} \\ \text{絶対値関数の和も容易に最小化できる} \end{array}$$



利点：目的関数の収束を保証しつつ比較的容易に反復最小化可能

補助関数型IVA [ono, 2011]: 補助関数の定義

- IVAのコントラスト関数として以下の関数を考えてみる

$$G(\mathbf{z}) = G_R(\|\mathbf{z}\|_2) \quad \begin{array}{l} \mathbf{z} \text{ は多次元複素変数} \\ \text{球対称(ノルムだけの)分布} \end{array}$$

IVAのコスト関数 J 及びその補助関数 Q を以下のように定める

$$J(\mathbf{W}) = \frac{1}{N_\tau} \sum_{\tau=1}^{N_\tau} \sum_{k=1}^K \underline{G(\mathbf{y}_k(\tau))} - \sum_{\omega=1}^{N_\omega} \log |\det W(\omega)|$$



$$Q(\mathbf{W}, \mathbf{V}) = \sum_{\omega=1}^{N_\omega} \left[\frac{1}{N_\tau} \sum_{\tau=1}^{N_\tau} \sum_{k=1}^K \underline{\mathbf{w}_k^h(\omega) V_k(\omega) \mathbf{w}_k(\omega)} - \log |\det W(\omega)| \right]$$

補助変数

$$V_k(\omega) = \frac{1}{N_\tau} \sum_{\tau=1}^{N_\tau} \left[\frac{G'(r_k(\tau))}{r_k(\tau)} \mathbf{x}(\omega, \tau) \mathbf{x}^h(\omega, \tau) \right] \quad r_k(\tau) = \sqrt{\sum_{\omega=1}^{N_\omega} |\mathbf{w}_k^h(\omega) \mathbf{x}(\omega, \tau)|^2}$$

補助関数型IVA: パラメータ更新

■ 重み付き共分散行列(補助変数)の更新

$$V_k(\omega) = \frac{1}{N_\tau} \sum_{\tau=1}^{N_\tau} \left[\frac{G'(r_k(\tau))}{r_k(\tau)} \mathbf{x}(\omega, \tau) \mathbf{x}^h(\omega, \tau) \right] \quad r_k(\tau) = \sqrt{\sum_{\omega=1}^{N_\omega} |\mathbf{w}_k^h(\omega) \mathbf{x}(\omega, \tau)|^2}$$

■ 補助関数 Q が小さくなるように W の更新

$$\frac{\partial Q}{\partial w_1} = 0$$

$$\frac{\partial Q}{\partial w_2} = 0$$

$$\frac{\partial Q}{\partial w_3} = 0$$

$$w_1^h V_1 w_1 = 1$$

$$w_2^h V_1 w_1 = 0$$

$$w_3^h V_1 w_1 = 0$$

$$w_1^h V_2 w_2 = 0$$

$$w_2^h V_2 w_2 = 1$$

$$w_3^h V_2 w_2 = 0$$

$$w_1^h V_3 w_3 = 0$$

$$w_2^h V_3 w_3 = 0$$

$$w_3^h V_3 w_3 = 1$$

どう
解く
のか？

$K=2$ ならば一般化固有値問題(閉解法あり)、 $K=3$ 以上は閉形式解が無い。
分離行列を行ベクトル毎に更新することは閉形式で可能(補助関数は減少する)。

補助関数型IVA:パラメータ更新

- 重み付き共分散行列(補助変数)の更新

$$V_k(\omega) = \frac{1}{N_\tau} \sum_{\tau=1}^{N_\tau} \left[\frac{G'(r_k(\tau))}{r_k(\tau)} \mathbf{x}(\omega, \tau) \mathbf{x}^h(\omega, \tau) \right] \quad r_k(\tau) = \sqrt{\sum_{\omega=1}^{N_\omega} |\mathbf{w}_k^h(\omega) \mathbf{x}(\omega, \tau)|^2}$$

- 補助関数 Q が小さくなるように W の更新

【IP(反復射影)アルゴリズム】

for $k = 1 : K$

$$\mathbf{w}_k(\omega) \leftarrow (W(\omega) V_k(\omega))^{-1} \mathbf{i}_k$$

$$\mathbf{w}_k(\omega) \leftarrow \mathbf{w}_k(\omega) / \sqrt{\mathbf{w}_k^h(\omega) V_k(\omega) \mathbf{w}_k(\omega)}$$

end

(\mathbf{i}_k は第 k 要素のみ
が1の単位ベクトル)

$K=2$ ならば一般化固有値問題(閉解法あり)、 $K=3$ 以上は閉形式解が無い。
分離行列を行ベクトル毎に更新することは閉形式で可能(補助関数は減少する)。

ILRMAにおける最適化

- ILRMAコスト関数はIVAコストとNMFコストの複合
 - 全変数の最適化は**全て補助関数法に基づく更新式**
 - 分離フィルタは**IP更新**、音源モデルは**乗法更新**

分離フィルタと分離信号の更新

$$V_{i,m} = \frac{1}{J} \sum_j \frac{1}{r_{ij,m}} \mathbf{x}_{ij} \mathbf{x}_{ij}^H$$

$$\mathbf{w}_{i,m} \leftarrow (\mathbf{W}_i V_{i,m})^{-1} \mathbf{e}_m$$

$$\mathbf{w}_{i,m} \leftarrow \mathbf{w}_{i,m} (\mathbf{w}_{i,m}^H V_{i,m} \mathbf{w}_{i,m})^{-\frac{1}{2}}$$

$$y_{ij,m} \leftarrow \mathbf{w}_{i,m}^H \mathbf{x}_{ij}$$

但し, $\mathbf{W}_i = (\mathbf{w}_{i,1} \cdots \mathbf{w}_{i,M})^H$,
 \mathbf{e}_m は m 番目の要素のみ1で
 他が0の縦ベクトル

音源モデルと推定分散の更新

$$z_{mk} \leftarrow z_{mk} \sqrt{\frac{\sum_{i,j} |y_{ij,m}|^2 t_{ik} v_{kj} (\sum_{k'} z_{mk'} t_{ik'} v_{k'j})^{-2}}{\sum_{i,j} t_{ik} v_{kj} (\sum_{k'} z_{mk'} t_{ik'} v_{k'j})^{-1}}}$$

$$t_{ik} \leftarrow t_{ik} \sqrt{\frac{\sum_{j,m} |y_{ij,m}|^2 z_{mk} v_{kj} (\sum_{k'} z_{mk'} t_{ik'} v_{k'j})^{-2}}{\sum_{j,m} z_{mk} v_{kj} (\sum_{k'} z_{mk'} t_{ik'} v_{k'j})^{-1}}}$$

$$v_{kj} \leftarrow v_{kj} \sqrt{\frac{\sum_{i,m} |y_{ij,m}|^2 z_{mk} t_{ik} (\sum_{k'} z_{mk'} t_{ik'} v_{k'j})^{-2}}{\sum_{i,m} z_{mk} t_{ik} (\sum_{k'} z_{mk'} t_{ik'} v_{k'j})^{-1}}}$$

$$r_{ij,m} = \sum_k z_{mk} t_{ik} v_{kj} : \text{音源毎の推定分散}$$

- 複雑なスペクトrogram音源モデルを内包しつつ、反復更新で尤度が単調増加することを保証するアルゴリズム (cf. 自然勾配)

多チャンネルNMFとは何が違うの？

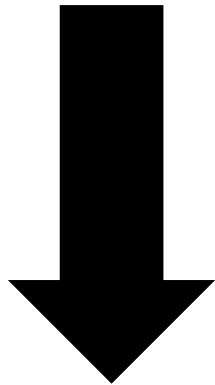
- 全音源の時間周波数構造をNMFでモデル化し音源毎の空間伝達特性(空間相関行列 R)を用いてクラスタリング
 - 「NMFモデル」と「音源毎の空間相関行列」を同時に推定

提案者と年代	空間的な混合系と音源モデル	各音源の空間相関行列	音源のスペクトログラム	各変数の最適化手法
Ozerov and Fevotte, 2010	$R^{(x)} = R^{(s)} \circ TV + R^{(b)}$	$R^{(s)}$ はランク1 $R^{(b)}$ はフルランク	NMF	EMで $R^{(s)}$, $R^{(b)}$, NMF変数を推定
Arberet et al., 2010	$R^{(x)} = R^{(s)} \circ TV + R^{(b)}$	$R^{(s)}$ も $R^{(b)}$ もフルランク	NMF	EMで $R^{(s)}$, $R^{(b)}$, NMF変数を推定
Ozerov et al., 2011	$R^{(x)} = (R^{(s)} Z) \circ TV + R^{(b)}$	$R^{(s)}$ はランク1 $R^{(b)}$ はフルランク	分割関数付きのNMF	EMで $R^{(s)}$, $R^{(b)}$, NMF変数を推定
Sawada et al., 2013	$R^{(x)} = (R^{(s)} Z) \circ TV$	フルランクの $R^{(s)}$	分割関数付きのNMF	乗算更新式で $R^{(s)}$, NMF変数を推定
Kitamura et al., 2015	$R^{(x)} = (R^{(s)} Z) \circ TV$	ランク1の $R^{(s)}$	分割関数付きのNMF	反復射影(IP)で分離行列 $W = A^{-1}$, 乗算更新式で NMF変数を推定

多チャンネルNMFとILRMAのミッシングリンク

- 多チャンネルNMFの目的関数にランク1空間モデルを導入

$$\mathcal{J} = \sum_{i,j} \left[\text{tr} \left(X_{ij} \left(\sum_k \left(\sum_m \mathbf{R}_i^{(s_m)} z_{mk} \right) t_{ik} v_{kj} \right)^{-1} \right) + \log \det \sum_k \left(\sum_m \mathbf{R}_i^{(s_m)} z_{mk} \right) t_{ik} v_{kj} \right]$$



1. ランク1チャンネル間相関を導入 ($\mathbf{R}_i^{(s_m)} = \mathbf{a}_{i,m} \mathbf{a}_{i,m}^H$)
2. 混合行列 \mathbf{A}_i を用いて表現しなおす
3. 分離行列 $\mathbf{W}_i = \mathbf{A}_i^{-1}$ と分離信号 $\mathbf{y}_{ij} = \mathbf{W}_i \mathbf{x}_{ij}$ に変数変換
ここで逆行列ドメインでの最適化にしたのがポイント!

ILRMAのコスト関数

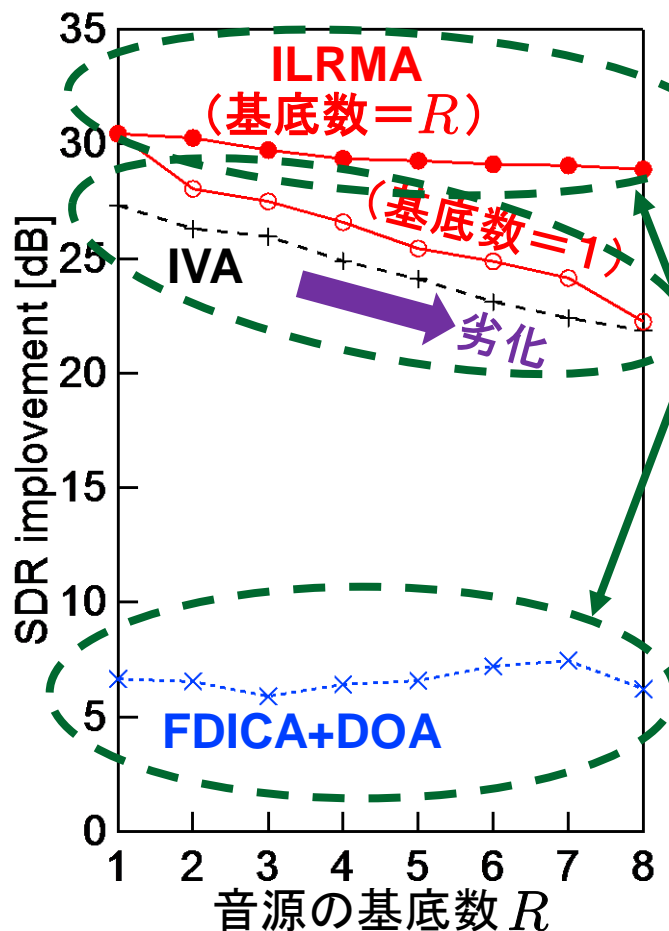
$$\mathcal{J} = \sum_{i,j} \left[\sum_m \log \sum_k z_{mk} t_{ik} v_{kj} + \sum_m \frac{|y_{ij,m}|^2}{\sum_k z_{mk} t_{ik} v_{kj}} - 2 \log |\det \mathbf{W}_i| \right]$$

- ランク1空間モデル制約付き多チャンネルNMFとILRMAは双対であるが、ILRMAは「分離モデル型」であり高速な最適化が可能

実験の比較例・実用化例

音源の基底数の違いによる分離精度比較

- 基底数 R を変化させた人工音源の分離実験



音源1の結果グラフ

FDICA+DOAは R によらない

- 具体的な音源モデルが無いため

IVAとILRMA(基底数=1)は音源の基底数の増加に伴って精度が劣化する

- 音源モデルが表現できなくなるため

ILRMA(基底数= R)は精度劣化しない

- 適切で具体的な音源モデルが仮定されているため

ILRMAの持つ音源モデル(NMF表現)が最も柔軟

各種音源分離手法の比較

- 実際の音響信号 & 空間混合の分離実験
 - SiSECデータ(実録音・2音源混合)による分離実験結果
 - 演算時間はIVAを基準に正規化した

Algorithm	SDR improv.	Comp. time
Soft masking	-0.1	-
IVA	2.6	1.0
Ozerov's MNMF	1.2	-
Sawada's MNMF	5.0	49.1
ILRMA	8.7	1.3

ILRMAが高い分離精度と低演算時間を実現

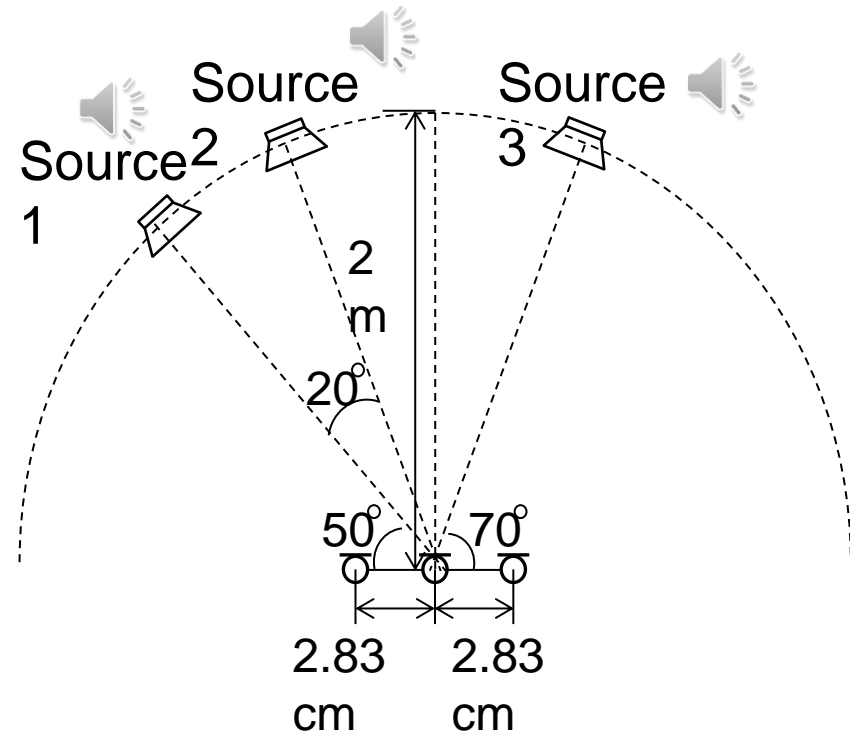
高速ICA及びILRMAによるデモ

- リアルタイム音声聞き分け(警察備品に採用)

[Mori, Saruwatari+, EURASIP JASP 2006]



- ドラム、弦楽器、音声からなる複合音の分離



2013年ドコモモバイルサイエンス賞を受賞

ロボット対話 デモンストレーション

Demonstration for
spoken-oriented hands-free
interactive robot.

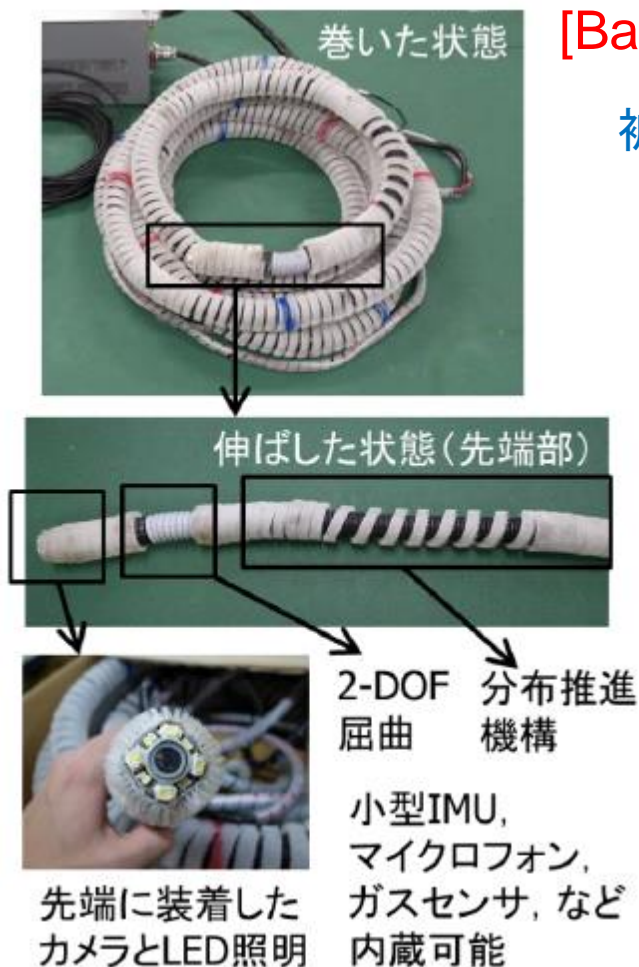
[Takahashi, Saruwatari+, IEEE Trans. ASLP 2009]

人工知能研究会優秀賞・平成23年度RIEC Award を受賞

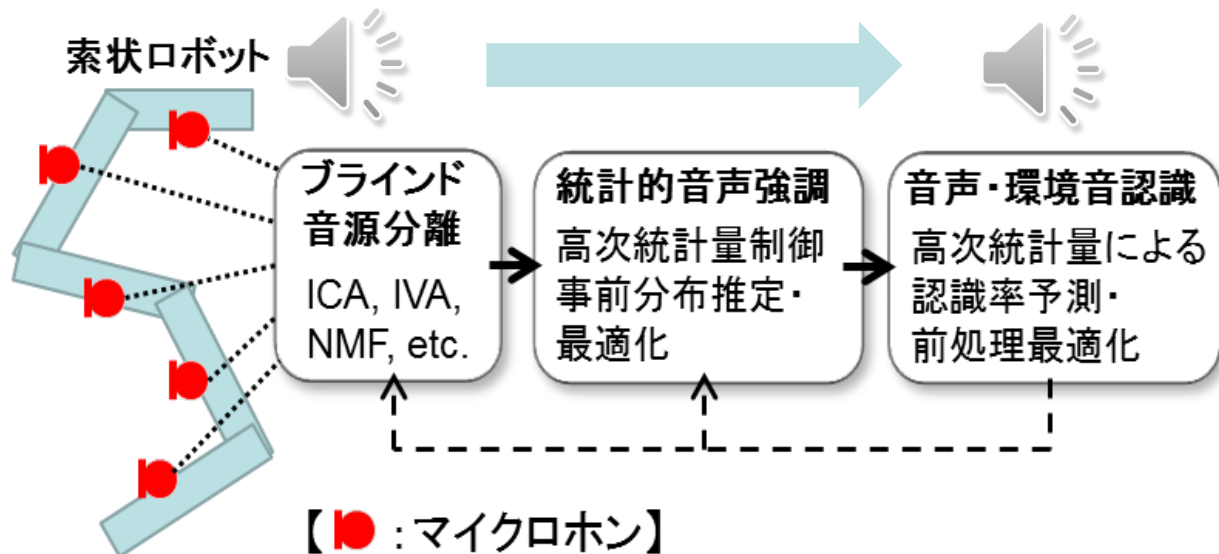
内閣府ImPACT災害対応タフロボット [2016年6月プレスリリース]

- 災害時の倒壊家屋に入り込んで被災者発見
- 環境音認識による状況把握・救助支援

[Bando, Saruwatari+, J. Robotics & Mechatronics 2017]



被災者はいるか？



いかなる曲がりくねった形状においても位置不定マイク同士が協調して騒音の中から被災者の声を見つけ出す

スパースな生成モデルの導入 ～スパース性と低ランク性の関係～

安定分布と Student's t 分布

• Symmetric α -stable ($S\alpha S$) 分布に基づく音源モデル

[A. Liutkus+, 2015], [U. Şimşekli+, 2015], [S. Leglaive+, 2017], [M. Fontaine+, 2017]

- これは複素変数の重ね合わせに関して安定分布

$$\begin{aligned} z_1 &\sim S\alpha S(z; 0, |z_1|^\alpha) \\ z_2 &\sim S\alpha S(z; 0, |z_2|^\alpha) \end{aligned}$$



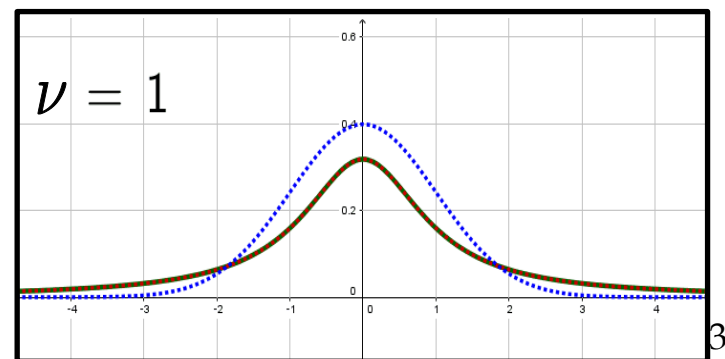
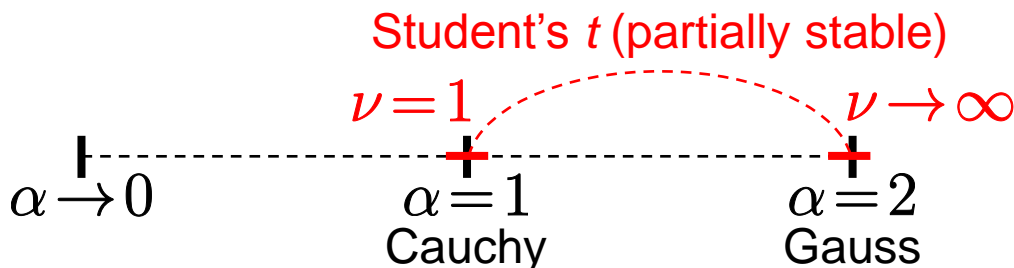
$$z_1 + z_2 \sim S\alpha S(z; 0, |z_1|^\alpha + |z_2|^\alpha)$$

- α を小さくしていくと裾の重いスパースな分布になる

• Student's t 分布

[C. Févotte+, 2006], [K. Yoshii+, 2016], [K. Kitamura+, 2016], [S. Leglaive+, 2017]

- これは 2つの $S\alpha S$ であるCauchy分布($\nu=1$)と Gaussian分布を表すことが出来る ($\nu \rightarrow \infty$)



t -ILRMA: 複素Student's t 分布生成モデル型ILRMA

- 時間周波数領域における複素Student's t 分布

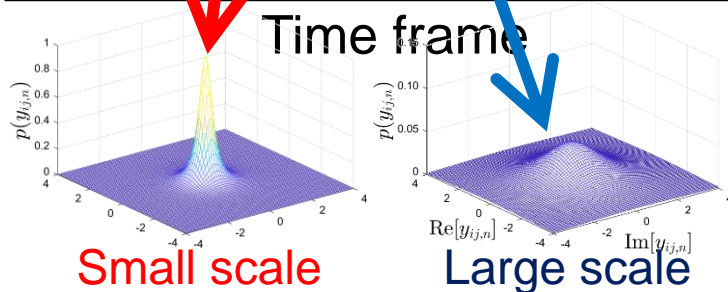
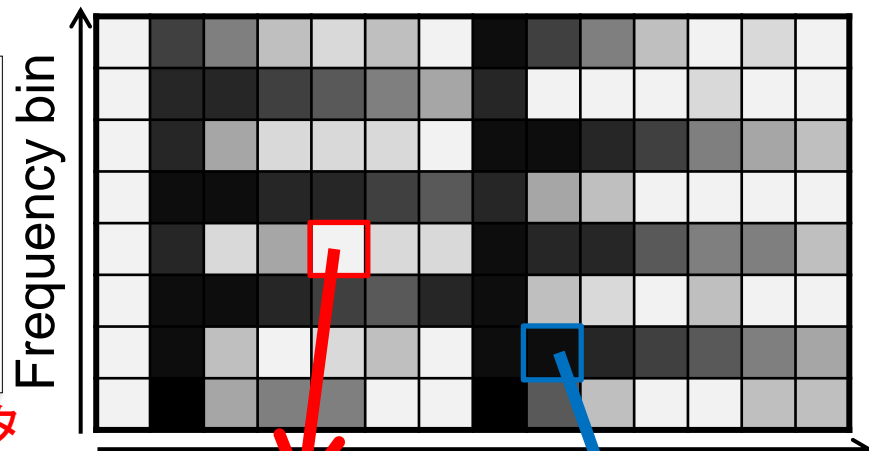
n 番目音源のスペクトログラム

$$p(\mathbf{Y}_n) = \prod_{i,j} p(y_{ij,n})$$

$$= \prod_{i,j} \frac{1}{\pi \sigma_{ij,n}^2} \left(1 + \frac{2 |y_{ij,n}|^2}{\nu \sigma_{ij,n}^2} \right)^{-\frac{2+\nu}{2}}$$

時間周波数グリッドにおけるスケールパラメータ
 ↓ この時間周波数スケールが以下の低ランク構造を持つ

$$\sigma_{ij,n}^p = \sum_k t_{ik,n} v_{kj,n} \quad (\Sigma_n^p = \mathbf{T}_n \mathbf{V}_n)$$



- t -ILRMA [Mogami, Saruwatari+, MLSP2017]におけるコスト関数

$$\mathcal{L} = -2J \sum_i \log |\det \mathbf{W}_i| + \sum_{i,j,n} \left[\left(1 + \frac{\nu}{2} \right) \log \left(1 + \frac{2}{\nu} \frac{|\mathbf{w}_{i,n}^H \mathbf{x}_{ij}|^2}{(\sum_k t_{ik,n} v_{kj,n})^{\frac{2}{p}}} \right) + \log \left(\sum_k t_{ik,n} v_{kj,n} \right)^{\frac{2}{p}} \right]$$

分離行列の最適化: t -ILRMAにおけるIPの適用

- 分離行列 W の更新にはIPを使いたいのだが...
 - IPは “ $-\log |\det \mathbf{W}_i|$ ” と “ $w_{i,n}$ の二次形式” の和の形式にのみ適用が出来る
 - t -ILRMAにおいてはどうなっているのでしょうか？
- コスト関数の比較
 - 従来の(時変ガウス型)ILRMA (IP適用可能)

$$\mathcal{L} = -2J \sum_i \left[\log |\det \mathbf{W}_i| \right] + \sum_{i,j,n} \left(\frac{|\mathbf{w}_{i,n}^H \mathbf{x}_{ij}|^2}{\sum_k t_{ik,n} v_{kj,n}} + \log \sum_k t_{ik,n} v_{kj,n} \right)$$

- t -ILRMA (“log”がついているのでそのままではIP適用不可能)

$$\mathcal{L} = -2J \sum_i \left[\log |\det \mathbf{W}_i| \right] + \sum_{i,j,n} \left[\left(1 + \frac{\nu}{2}\right) \underline{\log} \left(1 + \frac{2}{\nu} \frac{|\mathbf{w}_{i,n}^H \mathbf{x}_{ij}|^2}{\left(\sum_k t_{ik,n} v_{kj,n}\right)^{\frac{2}{p}}} \right) + \log \left(\sum_k t_{ik,n} v_{kj,n} \right)^{\frac{2}{p}} \right]$$

対数関数を外すような工夫が必要!

t-ILRMAにおいてどうやってIPを適用するか？

- 補助関数によって上から抑えてやれば良い

— 例えば接線不等式 $\log(z) \leq \frac{1}{\lambda}(z - \lambda) + \log \lambda$

$$\begin{aligned} \mathcal{L} &= -2J \sum_i \log |\det \mathbf{W}_i| + \sum_{i,j,n} \left[\left(1 + \frac{\nu}{2}\right) \log \left(1 + \frac{2}{\nu} \frac{|\mathbf{w}_{i,n}^H \mathbf{x}_{ij}|^2}{(\sum_k t_{ik,n} v_{kj,n})^{\frac{2}{p}}}\right) + \log \left(\sum_k t_{ik,n} v_{kj,n}\right)^{\frac{2}{p}} \right] \\ &\leq -2J \sum_i \left[\log |\det \mathbf{W}_i| \right] + \sum_{i,j,n} \left[\left(1 + \frac{\nu}{2}\right) \frac{1}{\alpha_{ij,n}} \left(1 + \frac{2}{\nu} \frac{|\mathbf{w}_{i,n}^H \mathbf{x}_{ij}|^2}{(\sum_k t_{ik,n} v_{kj,n})^{\frac{2}{p}}} - \alpha_{ij,n}\right) \right. \\ &\quad \left. + \left(1 + \frac{\nu}{2} \log \alpha_{ij,n}\right) + \log \left(\sum_k t_{ik,n} v_{kj,n}\right)^{\frac{2}{p}} \right] \end{aligned}$$

t-ILRMAにおいてどうやってIPを適用するか？

- 補助関数によって上から抑えてやれば良い

— 例えば接線不等式 $\log(z) \leq \frac{1}{\lambda}(z - \lambda) + \log \lambda$

$$\begin{aligned} \mathcal{L} &= -2J \sum_i \log |\det \mathbf{W}_i| + \sum_{i,j,n} \left[\left(1 + \frac{\nu}{2}\right) \log \left(1 + \frac{2}{\nu} \frac{|\mathbf{w}_{i,n}^H \mathbf{x}_{ij}|^2}{(\sum_k t_{ik,n} v_{kj,n})^{\frac{2}{p}}}\right) + \log \left(\sum_k t_{ik,n} v_{kj,n}\right)^{\frac{2}{p}} \right] \\ &\leq -2J \sum_i \left[\log |\det \mathbf{W}_i| \right] + \sum_{i,j,n} \left[\left(1 + \frac{\nu}{2}\right) \frac{1}{\alpha_{ij,n}} \left(1 + \frac{2}{\nu} \frac{|\mathbf{w}_{i,n}^H \mathbf{x}_{ij}|^2}{(\sum_k t_{ik,n} v_{kj,n})^{\frac{2}{p}}} - \alpha_{ij,n}\right) \right. \\ &\quad \left. + \left(1 + \frac{\nu}{2} \log \alpha_{ij,n}\right) + \log \left(\sum_k t_{ik,n} v_{kj,n}\right)^{\frac{2}{p}} \right] \end{aligned}$$

補助関数の最小点を求めるためにIPを適用可能

- IPに基づく \mathbf{W}_i の最適化

$$\begin{aligned} \mathbf{U}_{i,n} &= \frac{1}{J} \sum_j \frac{\mathbf{x}_{ij} \mathbf{x}_{ij}^H}{\frac{\nu}{\nu+2} (\sum_k t_{ik,n} v_{kj,n})^{\frac{2}{p}} + \frac{\nu}{\nu+2} |y_{ij,n}|^2} \\ \mathbf{w}_{i,n} &\leftarrow (\mathbf{W}_i \mathbf{U}_{i,n})^{-1} \mathbf{e}_n \\ \mathbf{w}_{i,n} &\leftarrow \mathbf{w}_{i,n} (\mathbf{w}_{i,n}^H \mathbf{U}_{i,n} \mathbf{w}_{i,n})^{-\frac{1}{2}} \end{aligned}$$

\mathbf{e}_n : unit vector whose n th element is unity

Cf. MNMF [H. Sawada+, 2013] は一つの空間相関行列を更新するのに J 回の逆行列演算と2回の固有値分解が必要となる。

音源モデル最適化：スパース性と低ランク性の関係

- コスト関数における音源モデルパラメータ項

$$\mathcal{L} = -2J \sum_i \log |\det \mathbf{W}_i| + \sum_{i,j,n} \left[\left(1 + \frac{\nu}{2}\right) \underline{\log} \left(1 + \frac{2}{\nu} \frac{|\mathbf{w}_{i,n}^H \mathbf{x}_{ij}|^2}{\left(\sum_k t_{ik,n} v_{kj,n}\right)^{\frac{2}{p}}}\right) + \log \left(\sum_k t_{ik,n} v_{kj,n}\right)^{\frac{2}{p}} \right]$$

$$\left(\sum_k t_{ik,n} v_{kj,n}\right)^{-\frac{2}{p}} \leq \sum_k \eta_{ij,nk} \left(\frac{t_{ik,n} v_{kj,n}}{\eta_{ij,nk}}\right)^{-\frac{2}{p}} \quad \log(z) \leq \frac{1}{\lambda} (z - \lambda) + \log \lambda$$

更新式

$$t_{ik,n} \leftarrow t_{ik,n} \left[\frac{\sum_j \frac{z_{ij,n}^{\frac{p}{p+2}} v_{kj,n}}{\left(\sum_{k'} t_{ik',n} v_{k'j,n}\right)^2}}{\sum_j \frac{1}{\sum_{k'} t_{ik',n} v_{k'j,n}} v_{kj,n}} \right]^{\frac{p}{p+2}}$$

指数に関して一般化された
擬似観測 $z_{ij,n}$ との板倉斎藤
擬距離NMFと等価

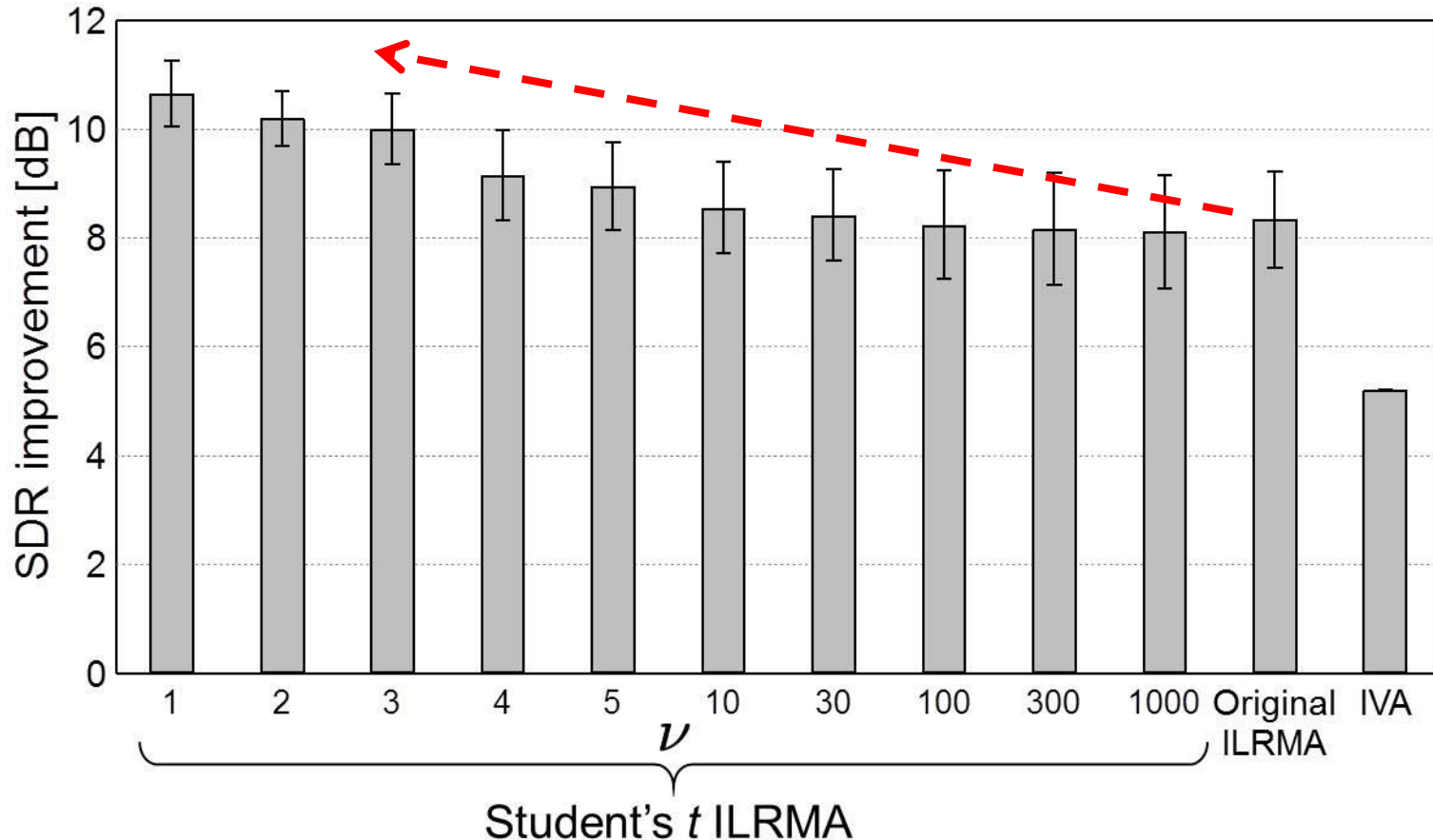
$$z_{ij,n} = \sigma_{ij,n}^{p-2} \left(\frac{\nu}{\nu+2} |y_{ij,n}|^{-2} + \frac{2}{\nu+2} \sigma_{ij,n}^{-2} \right)^{-1}$$

真の観測 y と低ランクモデル σ との ν : 2調和平均

⇒ ν を小さく (スパース音源に) すると低ランク性が強調される !

実験的比較例

- ν パラメータを変えて音声の分離を試みた



ν を小さく(スパース音源に)すると分離精度が向上する

事前分布・正則化の導入



■ 動機

- パラメータに事前分布を導入することにより音源分離精度を向上させる
- ILRMAのメインコストに正則化項を加算することによって実現される

■ 低ランク音源パラメータ TV については...

- ISNMF項にスパース正則化を付与する（L1正則化等）だけで実装可能

■ 分離行列 W については...

- ICAの時代は最急降下法だったので、メインコストと正則化項の偏微分を単純に足せばよかった
- **IPを使った補助関数法の場合、偏微分だけでは十分ではなく、「偏微分 = 0（停留点）」を解析的に解く必要がある**のでそれほど簡単ではない

正則化付き目的関数の変形



ILRMAの目的関数 (2次形式 + logdet)

$$\mathcal{J} = J \sum_i \left[\sum_n \mathbf{w}_{i,n}^H \mathbf{D}_{i,n} \mathbf{w}_{i,n} - \log |\det \mathbf{W}_i|^2 \right] + \mathcal{C}$$

ガウシアン事前分布による正則化 (2次形式 + **1次内積** + logdet)

$$\begin{aligned} \mathcal{J}_P &= J \sum_i \left[\sum_n \mathbf{w}_{i,n}^H \mathbf{D}_{i,n} \mathbf{w}_{i,n} - \log |\det \mathbf{W}_i|^2 \right] + \underbrace{J \sum_{i,n} \lambda_n \|\mathbf{w}_{i,n} - \hat{\mathbf{w}}_{i,n}\|^2}_{\text{Red dashed box}} + \mathcal{C} \\ &= J \sum_i \left[\sum_n \left(\mathbf{w}_{i,n}^H \hat{\mathbf{D}}_{i,n} \mathbf{w}_{i,n} - \lambda_n \hat{\mathbf{w}}_{i,n}^H \mathbf{w}_{i,n} - \lambda_n \mathbf{w}_{i,n}^H \hat{\mathbf{w}}_{i,n} \right) - \log |\det \mathbf{W}_i|^2 \right] + \mathcal{C}' \end{aligned}$$

$\hat{\mathbf{W}}_i = (\hat{\mathbf{w}}_{i,1}, \dots, \hat{\mathbf{w}}_{i,N})^H$: 事前情報から定まる目標分離行列

- 新たな目的関数は $\mathbf{w}_{i,n}$ の**1次項 (内積)** を含む → **IPが利用できない!**
- **1次項が加わった新たな目的関数の反復最適化法**が提案されている

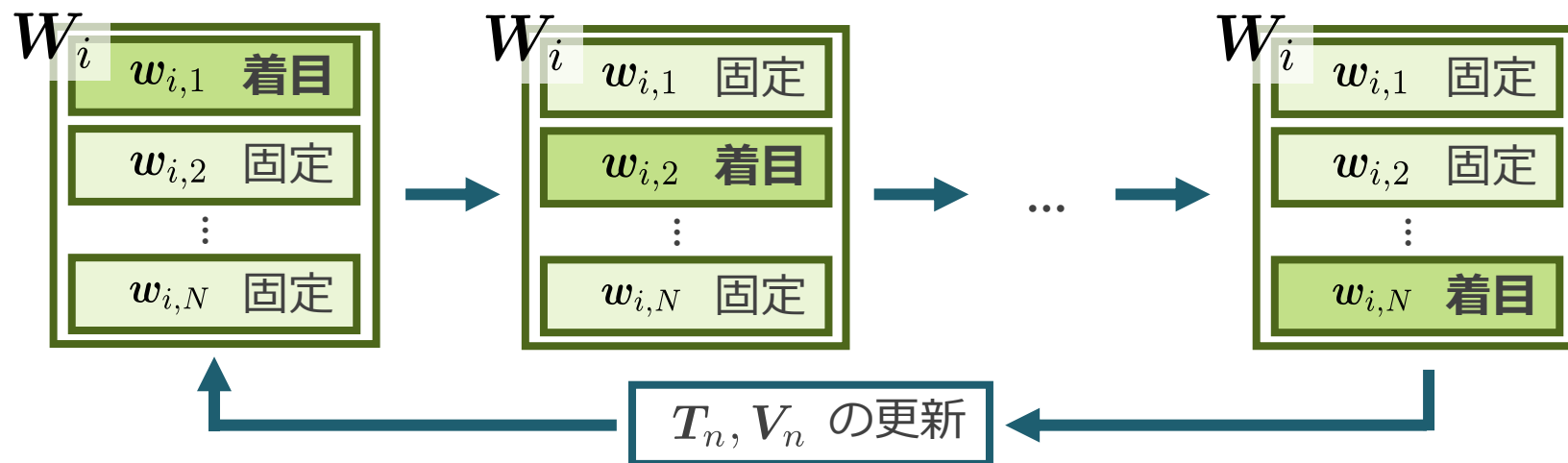
[三井、猿渡+, SIPシンポ2017]

目的関数最適化のアプローチ

正則化付きILRMAの目的関数

$$\mathcal{J}_P = J \sum_i \left[\sum_n \left(\mathbf{w}_{i,n}^H \hat{\mathbf{D}}_{i,n} \mathbf{w}_{i,n} - \lambda_n \hat{\mathbf{w}}_{i,n}^H \mathbf{w}_{i,n} - \lambda_n \mathbf{w}_{i,n}^H \hat{\mathbf{w}}_{i,n} \right) - \log |\det \mathbf{W}_i|^2 \right] + \mathcal{C}'$$

- \mathbf{W}_i に関する最適解を解析的に求めるのは困難 → **反復最適化**が必要
- **ブロック座標降下法**を利用して反復最適化を行う
 - あるベクトル $\mathbf{w}_{i,n}$ のみに着目すれば, 目的関数の停留点が代数的に求まる
 - 行ベクトルごとに更新 → **Vectorwise coordinate descent (VCD)** と呼ぶ



停留点の導出：余因子展開アプローチ

正則化付きILRMAの目的関数

$$\mathcal{J}_P = J \sum_i \left[\sum_n \left(\mathbf{w}_{i,n}^H \hat{\mathbf{D}}_{i,n} \mathbf{w}_{i,n} - \lambda_n \hat{\mathbf{w}}_{i,n}^H \mathbf{w}_{i,n} - \lambda_n \mathbf{w}_{i,n}^H \hat{\mathbf{w}}_{i,n} \right) - \log |\det \mathbf{W}_i|^2 \right] + C'$$

目的関数 \mathcal{J}_P の停留点に $\mathbf{w}_{i,n}$ を更新 … $\log |\det \mathbf{W}_i|$ の偏微分が厄介

- 余因子行列を次のように定義： $\mathbf{B}_i = (\mathbf{b}_{i,1}, \dots, \mathbf{b}_{i,N}) \stackrel{\text{def}}{=} (\det \mathbf{W}_i) \mathbf{W}_i^{-1}$
- logdet項は、次のように変形できる： $\log |\det \mathbf{W}_i|^2 = \log |\mathbf{w}_{i,n}^H \mathbf{b}_{i,n}|^2$
- $\mathbf{b}_{i,n}$ は $\mathbf{w}_{i,n}$ と関係なく定まるため、 $\mathbf{w}_{i,n}^*$ によるlogdetの偏微分は以下：

$$\frac{\partial \log |\mathbf{w}_{i,n}^H \mathbf{b}_{i,n}|^2}{\partial \mathbf{w}_{i,n}^*} = \frac{\mathbf{b}_{i,n}}{\mathbf{w}_{i,n}^H \mathbf{b}_{i,n}}$$

→ $\frac{\partial \mathcal{J}_P}{\partial \mathbf{w}_{i,n}^*} = 0$ の解へと $\mathbf{w}_{i,n}$ を更新すればよい！

「(目的関数の偏微分) = 0」の代数的求解

「(目的関数の偏微分) = 0」は、次の通り

$$\hat{D}_{i,n} \mathbf{w}_{i,n} - \lambda_n \hat{\mathbf{w}}_{i,n} - \frac{1}{\mathbf{w}_{i,n}^H \mathbf{b}_{i,n}} \mathbf{b}_{i,n} = 0$$

$\beta_{i,n} = 1 / \mathbf{w}_{i,n}^H \mathbf{b}_{i,n}$ において整理すると、以下の方程式を満たす $\beta_{i,n}$ を求める問題へと帰着可能

$$\mathbf{b}_{i,n}^H \hat{D}_{i,n}^{-1} \mathbf{b}_{i,n} |\beta_{i,n}|^2 + \lambda_n \hat{\mathbf{w}}_{i,n}^H \hat{D}_{i,n}^{-1} \mathbf{b}_{i,n} \beta_{i,n} - 1 = 0$$

この方程式の解は、次のように代数的に求まる:

$$\beta_{i,n} = e^{j\phi_{i,n}} / \sqrt{\mathbf{b}_{i,n}^H \hat{D}_{i,n}^{-1} \mathbf{b}_{i,n}} \quad (\lambda_n \hat{\mathbf{w}}_{i,n}^H \hat{D}_{i,n}^{-1} \mathbf{b}_{i,n} = 0)$$

$$\beta_{i,n} = \frac{\lambda_n \mathbf{b}_{i,n}^H \hat{D}_{i,n}^{-1} \hat{\mathbf{w}}_{i,n}}{2 \mathbf{b}_{i,n}^H \hat{D}_{i,n}^{-1} \mathbf{b}_{i,n}} \left[-1 \pm \sqrt{1 + \frac{4 \mathbf{b}_{i,n}^H \hat{D}_{i,n}^{-1} \mathbf{b}_{i,n}}{\lambda_n^2 |\mathbf{b}_{i,n}^H \hat{D}_{i,n}^{-1} \hat{\mathbf{w}}_{i,n}|^2}} \right] \quad (\lambda_n \hat{\mathbf{w}}_{i,n}^H \hat{D}_{i,n}^{-1} \mathbf{b}_{i,n} \neq 0)$$



最終的な反復更新式

$$\mathbf{u}_{i,n} = \hat{\mathbf{D}}_{i,n}^{-1} \mathbf{W}_i^{-1} \mathbf{e}_n$$

$$\hat{\mathbf{u}}_{i,n} = \lambda_n \hat{\mathbf{D}}_{i,n}^{-1} \hat{\mathbf{w}}_{i,n}$$

$$h_{i,n} = \mathbf{u}_{i,n}^H \hat{\mathbf{D}}_{i,n} \mathbf{u}_{i,n}$$

$$\hat{h}_{i,n} = \mathbf{u}_{i,n}^H \hat{\mathbf{D}}_{i,n} \hat{\mathbf{u}}_{i,n}$$

$$\mathbf{w}_{i,n} \leftarrow \begin{cases} \frac{\mathbf{u}_{i,n}}{\sqrt{h_{i,n}}} + \hat{\mathbf{u}}_{i,n} & (\text{if } \hat{h}_{i,n} = 0) \\ \frac{\hat{h}_{i,n}}{2h_{i,n}} \left[-1 + \sqrt{1 + \frac{4h_{i,n}}{|\hat{h}_{i,n}|^2}} \right] \mathbf{u}_{i,n} + \hat{\mathbf{u}}_{i,n} & (\text{otherwise}) \end{cases}$$

ベクトル毎の座標降下法に基づき $\mathbf{w}_{i,n}$ を順々に更新する

音源分離例 : Average SDRi of various BSS methods [dB]

(θ_1, θ_2)	AuxIVA	MNMF	t -MNMF	ILRMA (Method 1)	Regularized ILRMA (Method 3)
$(-40^\circ, +40^\circ)$	3.97	3.84	4.80	7.35	12.62
$(-40^\circ, +20^\circ)$	4.15	3.80	4.46	3.51	11.10

まとめ

- 優決定BSS問題における代表的なアルゴリズムとして、FDICA・IVAびILRMAを取り上げ、それらが何をモデリングしどのような数理アルゴリズムによって信号推定するのか考察した。
- 特に空間モデルの推定に関し、多チャンネルNMFに見られるような「生成モデル型」とBSSで用いられる「分離モデル型」の違い・得失についても解説を行い、最適化パラメータのドメイン変更によって性能が大きく改善されていった歴史を紹介した。
- スパース音源モデルの導入やパラメータ事前分布の導入に関して解説を行い、その最適化アルゴリズムを紹介した。

