ノンパラメトリックベイズモデルに基づく音響信号処理

吉井 和佳
産業技術総合研究所 情報技術研究部門
メディアインタラクション研究グループ
主任研究員

ノンパラメトリックベイズ

はじめに

- ノンパラメトリックベイズとは何か？
  - ベイズ推定における強力な方法論
    - 基礎理論 (部品) は1973年にすでに提案
      - T. Ferguson: A Bayesian analysis of some nonparametric problems, Annals of Statistics, Vol. 1, No.2, pp. 209-230, 1973.
    - 2000年前後からNIPSを中心に流行
      - 計算機の発展で計算量の大きいベイズ推定が可能に
- キーワード
  - ディリクレ過程
    - 棒折り過程 (Stick-Breaking Process: SBP)
    - 中華料理店過程 (Chinese Restaurant Process: CRP)
  - 無限モデル

観測データと確率モデル

- 観測データ
  - 我々が実際に観測できる量
  - 確率変数：何らかの確率分布に従って生成される
    - 例：サイコロをN回振って出た目の系列：離散分布に従う

$$X = \{x_1, x_2, \cdots, x_N\}$$

- 確率モデル
  - 観測データの生成過程の確率的な振る舞いを記述するもの
    - 確率モデルから生成されるデータは無数に考えられ
      観測データとはその１つの実現値 → 確率(密度)を与える

$$p(X \mid \Theta)$$

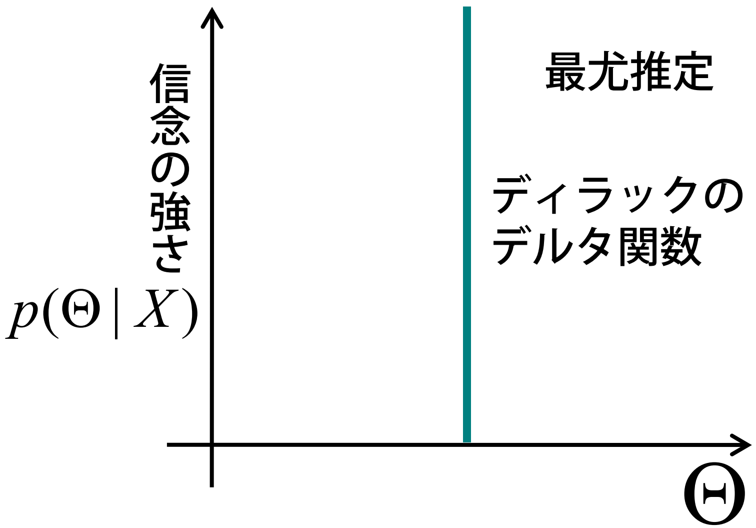
モデルパラメータ

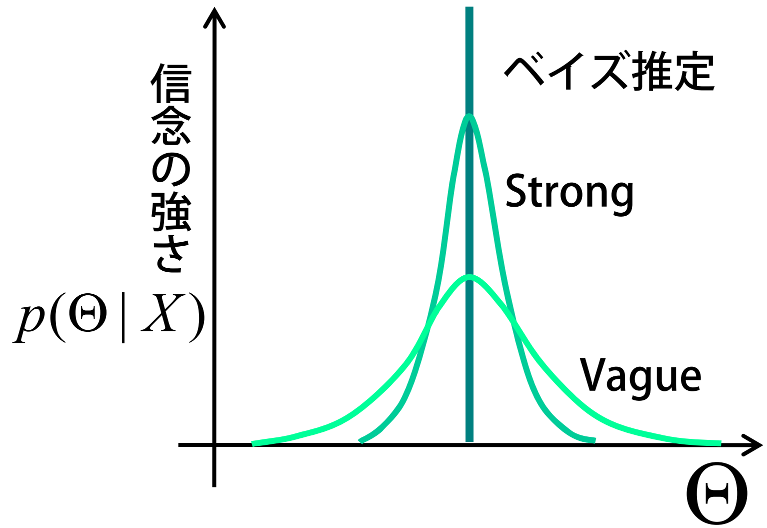
例: 離散分布のパラメータ
(各面の出る確率) $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}$

確率モデルの学習

- 与えられた観測データの生成確率が最大になるように確率モデルのパラメータを推定すること
  - 最尤推定
    - 最適なパラメータを点推定する (一意に決める)
  - ベイズ推定
    - パラメータの事後分布を推定する (信念の強さを反映する)

データが無限にあれば両者は一致





最尤推定

- 確率モデルが観測データに対してフィットしすぎると汎化能力（未知データに対する予測能力）が悪化

例：サイコロの各面が出る確率を最尤推定する

$$X = \{x_1, x_2, \cdots, x_N\}$$

$$\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}$$

$$p(X \mid \Theta) = \prod_{k=1}^{K} \theta_k^{n_k} \xrightarrow{\text{最大化}} \boxed{\theta_k = \frac{n_k}{N}}$$

N=10の観測データ

目 k	回数 n_k
1	4
2	1
3	0
4	0
5	2
6	3

最尤推定値 θ_k
0.4
0.1
0.0
0.0
0.2
0.3

各面の出る確率が
偏りすぎではないか？

x_{N+1} はどうなるのだろうか？
3,4が出る確率はゼロ！

最尤推定 (詳細)

- 観測データだけから最適なパラメータを点推定

例：サイコロの各面が出る確率を最尤推定する

$$X = \{x_1, x_2, \cdots, x_N\}$$

$$\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}$$

最大化したい目的関数

$$p(X \mid \Theta) = \prod_{k=1}^{6} \theta_k^{n_k}$$

拘束条件

$$\sum_{k=1}^{6} \theta_k = 1$$

1. 対数をとって凹関数化：これを最大化

$$\log p(X \mid \Theta) = \sum_{k=1}^{6} n_k \log \theta_k$$

2. 拘束条件付きの最適化：ラグランジュの未定乗数法

$$F = \log p(X \mid \Theta) + \lambda \left(1 - \sum_{k=1}^{6} \theta_k \right)$$

3. 偏微分して 0 とおく

$$\frac{\partial F}{\partial \theta_k} = \frac{n_k}{\theta_k} - \lambda \equiv 0 \quad \Rightarrow \quad \theta_k = \frac{n_k}{\lambda}$$

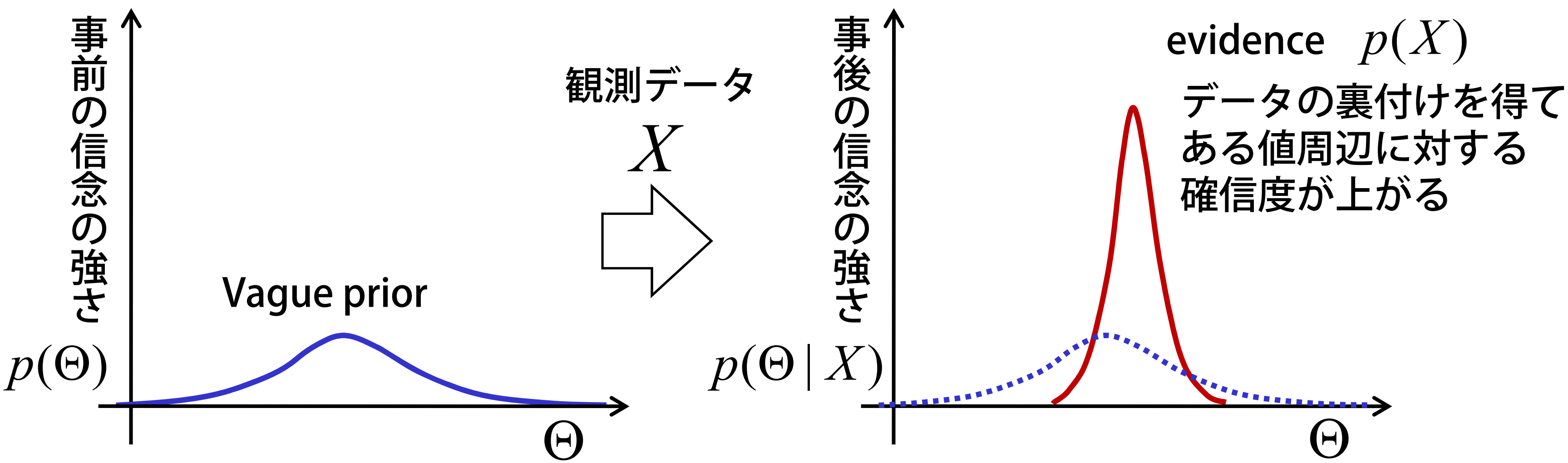
4. 拘束条件に代入することで未定乗数を計算

$$\lambda = N$$

$$\boxed{\theta_k = \frac{n_k}{N}}$$

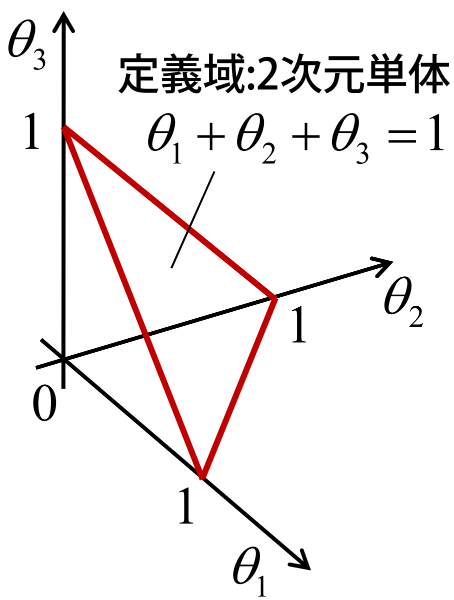
信念の強さの確率表現

- 未知パラメータの値のあらゆる可能性を考えておけば極端な推定結果は回避可能
  - 事前分布：「このあたりがそれっぽい」という予断
    - 事前分布のパラメータ(ハイパーパラメータ)を変えることで**事前の信念の強さ**を反映することができる
  - 事後分布：観測データを見たあとでの判断
    - 観測データが増えるたびに**事後の信念の強さ**が変化

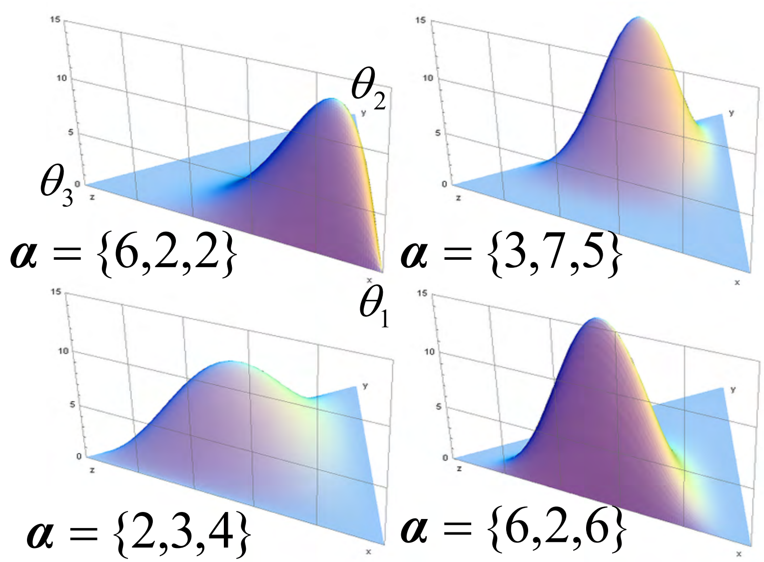


事前分布

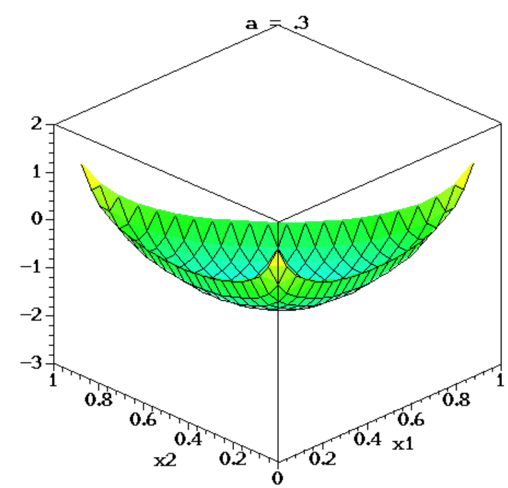
- 事前分布とは確率分布上の確率分布
  - 確率分布を記述するパラメータが従う確率分布とも言える
  - 共役事前分布の利用が便利
    - 事前分布 $p(\Theta)$ と事後分布 $p(\Theta \mid X)$ が同じ形になる
      - 離散分布に対する共役事前分布：ディリクレ分布



$\Theta = \{\theta_1, \theta_2, \theta_3\}$ の分布 $\mathrm{Dir}(\Theta \mid \boldsymbol{\alpha})$



$0 < \boldsymbol{\alpha} < 2$ における変化



ベイズ推定

- 適切な事前分布を与えることで過学習が抑制
  - ディリクレ事前分布 = 下駄をはかせる

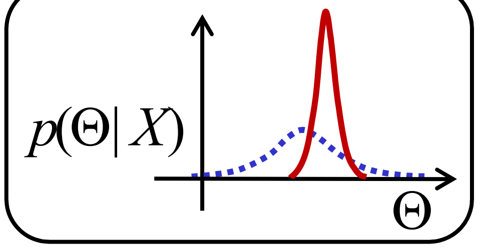
本当に
これだけ！

ディリクレ分布のパラメータ

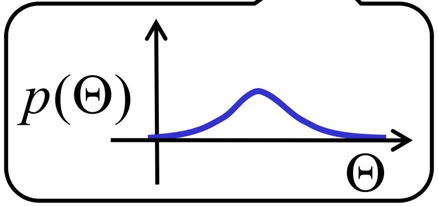
k	事前の観測回数 α_k	実際の観測回数 n_k	事後の観測回数 $\alpha_k + n_k$
1	3	4	7
2	3	1	4
3	3	0	3
4	3	0	3
5	3	2	5
6	3	3	6

ベイズ推定による
不確実性の保持

$\mathrm{Dir}(\Theta \mid \boldsymbol{\alpha} + \boldsymbol{n})$



$\mathrm{Dir}(\Theta \mid \boldsymbol{\alpha})$



点推定

最尤推定値

$\Theta = \{0.4, 0.1, 0.0, 0.0, 0.2, 0.3\}$

MAP推定値

$\hat{\Theta} = \{0.25, 0.14, 0.11, 0.11, 0.18, 0.21\}$

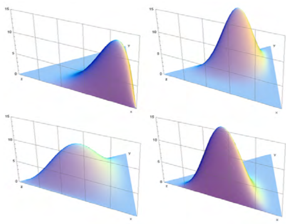
ベイズ推定 (詳細)

- 共役事前分布を用いて事後分布を計算

例：サイコロの各面が出る確率をベイズ推定する

$$X = \{x_1, x_2, \cdots, x_N\} \quad \Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}$$

$$\boxed{p(\Theta) = \mathrm{Dir}(\Theta \mid \boldsymbol{\alpha})} = C(\boldsymbol{\alpha}) \prod_{k=1}^{6} \theta_k^{\alpha_k - 1}$$



a_k：仮想的な観測回数に相当 (事前の信念の強さ)

ある程度真っ当なサイコロだと信じるなら例えば $a_k = 3$

信念が変化

$$p(X \mid \Theta) = \prod_{k=1}^{6} \theta_k^{n_k}$$

ベイズ推定ではベイズの定理がすべて！

$$p(\Theta \mid X) = \frac{p(X \mid \Theta) p(\Theta)}{p(X)} \propto p(X \mid \Theta) p(\Theta) = C(\boldsymbol{\alpha}) \prod_{k=1}^{6} \theta_k^{\alpha_k + n_k - 1}$$

$$\boxed{p(\Theta \mid X) = C(\boldsymbol{\alpha} + \boldsymbol{n}) \prod_{k=1}^{6} \theta_k^{\alpha_k + n_k - 1} = \mathrm{Dir}(\Theta \mid \boldsymbol{\alpha} + \boldsymbol{n})}$$

正規化係数Cは明示的に
考えなくても大丈夫

ベイズモデル再考

- これまで確率モデルの複雑さは既知と仮定
  - 複雑さは「観測データに合わせて」手動で指定
    - 例：サイコロの面数 (複雑さ) は６と指定
      ＝ ６次元ディリクレ分布を事前分布として利用

疑問：サイコロの面数が不明な場合はどうすべきか？

観測データ $X = \{1, 2, 2, 1, 2, 1, 3, 2\}$

観測データに対しては３面サイコロを考えれば良さそうに思えるが
このサイコロを振ると次に「4」や「5」が出るかもしれない

⇩

４次元ディリクレ分布を仮定すべき？ → 将来「5」が出たら困る
５次元ディリクレ分布を仮定すべき？ → 将来「6」が出たら困る
・・・

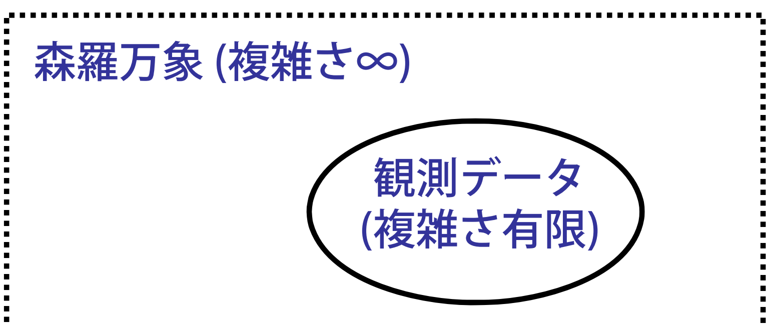
⇩

無限次元ディリクレ分布 ＝ ディリクレ過程 (DP)

ノンパラメトリックベイズモデル

- 無限の複雑さを持つベイズモデル
  - 「ノンパラメトリック」＝「無限個のパラメータからなる」
    - 「パラメータがない」という意味ではない
  - 観測データに限らない森羅万象を考慮
    - 本質的に汎化性能に優れている
    - 無限集合である「森羅万象」は無限のバリエーションに富む
      - 無限個のデータがあれば無限個のパラメータが必要
    - 有限集合である「観測データ」はそのごく一部
      - 有限個のデータであれば有限個のパラメータで十分

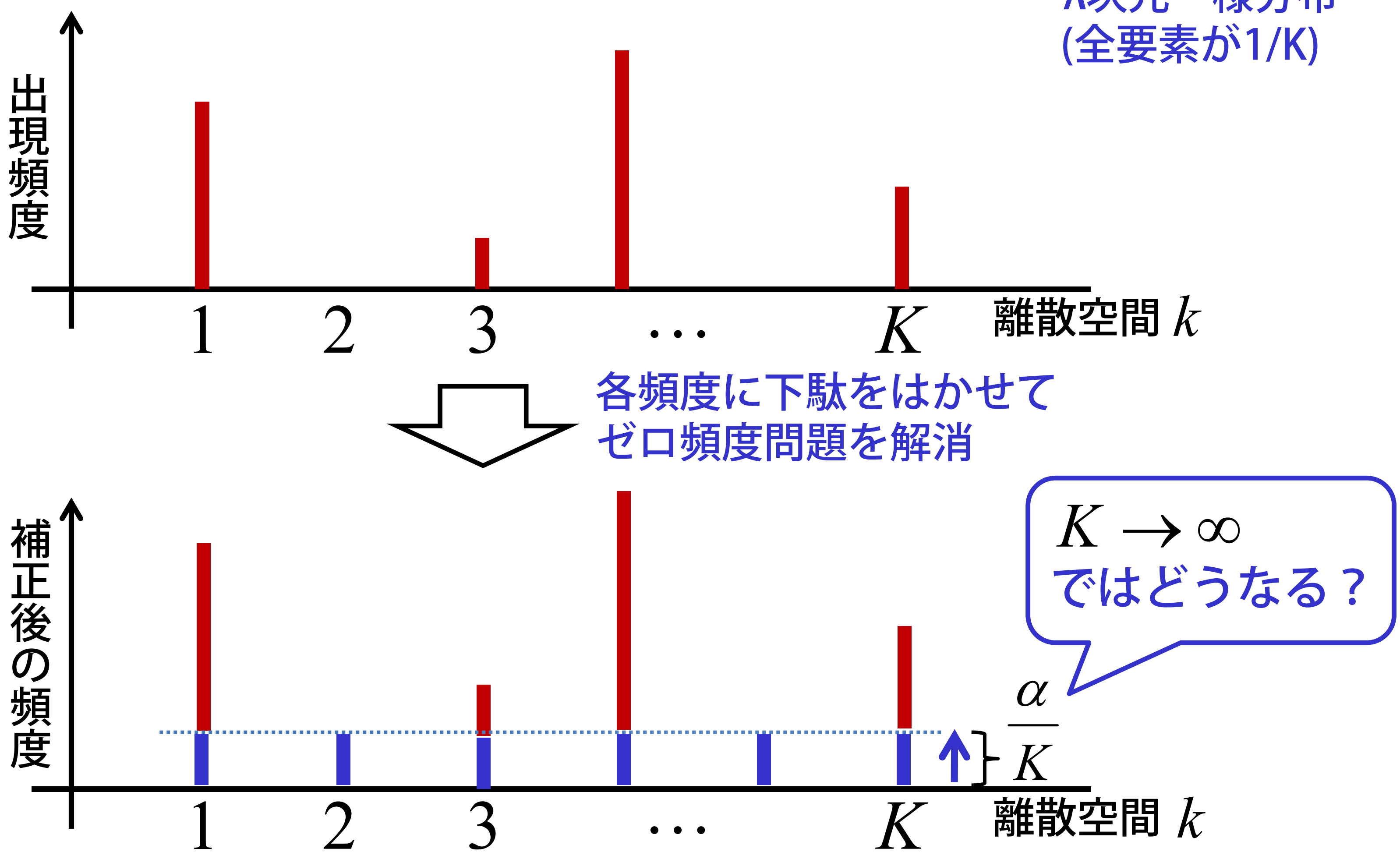
→ 計算機で実現可能！



*K*次元ディリクレ分布

- ディリクレ事前分布＝かさ上げスムージング $\mathrm{Dir}\left(\Theta \middle| \alpha \mathbf{1}\right)$
  - パラメータ：はかせる下駄の高さ

*K*次元一様分布
(全要素が1/K)

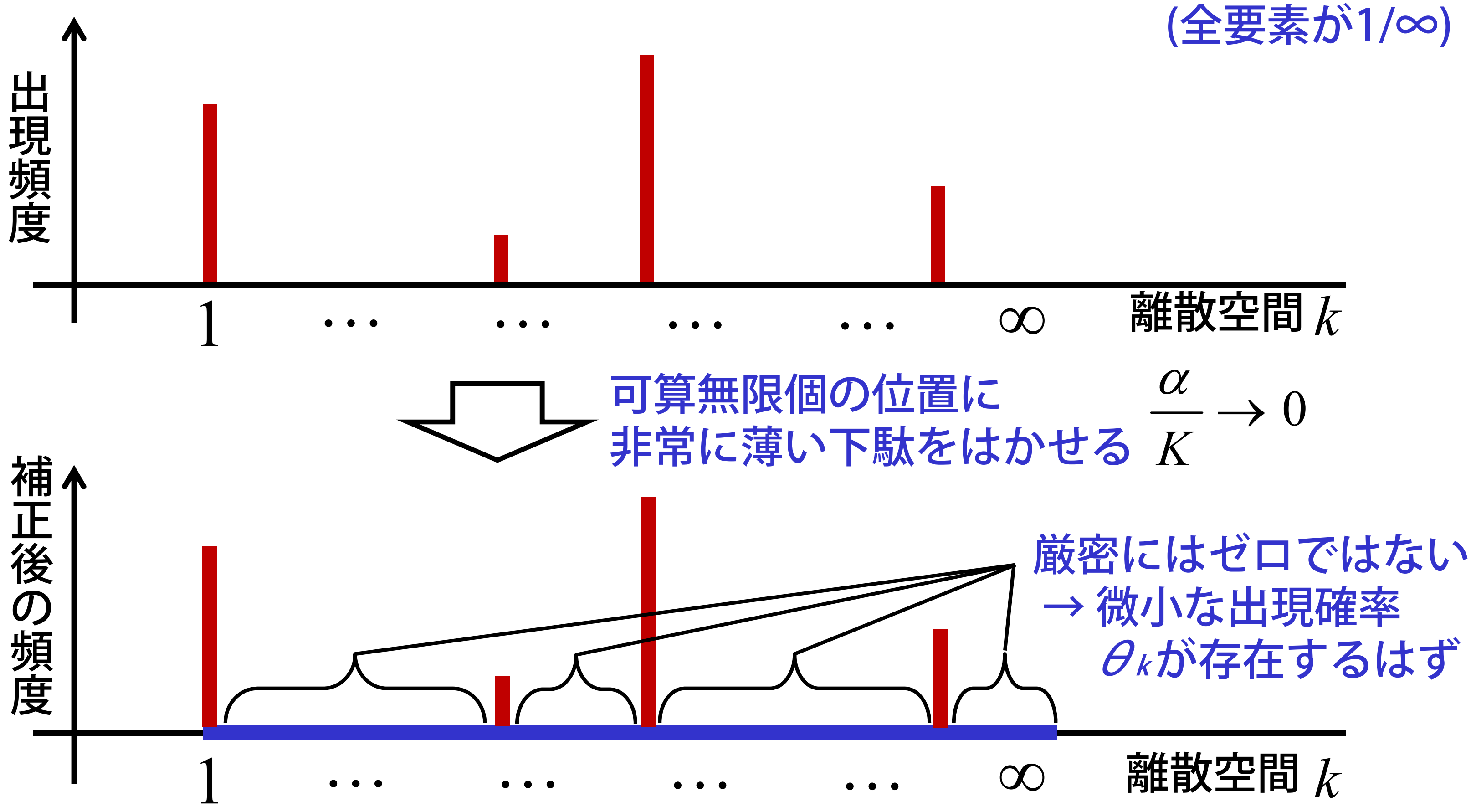


無限次元ディリクレ分布

- ディリクレ分布の次元 K を無限に発散させる
  - 下駄の高さが非常に薄くなる
  - 無限個ある面の出現確率は厳密には非零

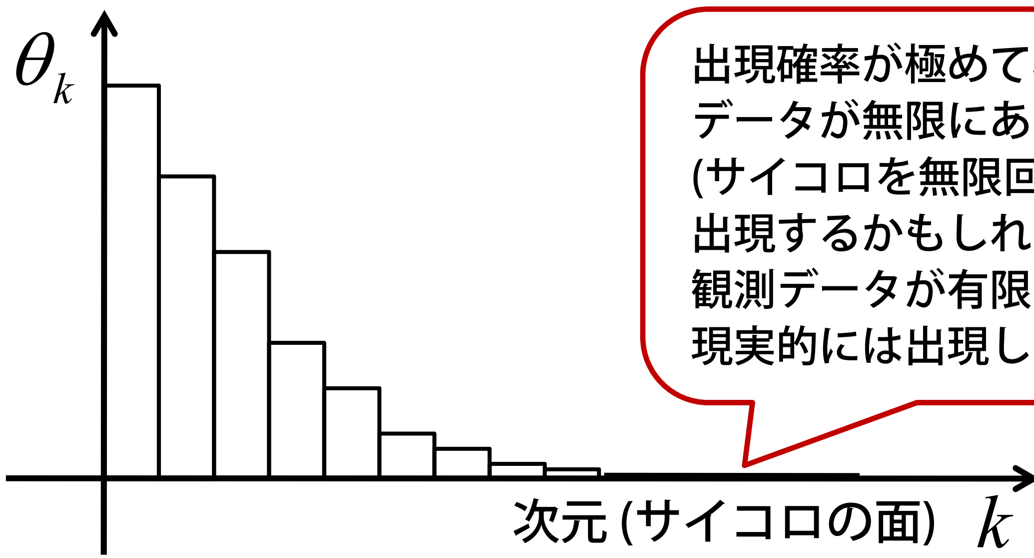
$$\mathrm{Dir}\left(\Theta \middle| \alpha \underline{\boldsymbol{1}}\right) \quad K \to \infty$$

∞次元一様分布
(全要素が1/∞)



ディリクレ過程 (DP)

- 無限次元ディリクレ分布と等価 $\Theta \sim \mathrm{DP}(\alpha, \boldsymbol{I}) = \mathrm{Dir}(\alpha \boldsymbol{I})$
  - 無限次元離散分布に対する確率分布 $\Theta = \{\theta_1, \theta_2, \theta_3, \cdots, \theta_\infty\}$
    - 無限個のパラメータの和は 1
    - ほとんどのパラメータはほとんどゼロ (ゼロではない)



出現確率が極めて小さいので
データが無限にあれば
(サイコロを無限回振れば)
出現するかもしれないが
観測データが有限であれば
現実的には出現しない

- 無限次元離散分布を直接モデル化 (棒折り過程)
- 無限次元離散分布から得られるサンプルをモデル化 (中華料理店過程)

棒折り過程 (SBP)

DPの一般化

Pitman-Yor過程
$\mathrm{Beta}(1-d, \alpha+dk)$
Beta two-parameter過程
$\mathrm{Beta}(\alpha, \beta)$

- 無限次元の離散分布を直接表現
  - 長さ1の棒を無限回折りとっていく
  - どこで折りとるかは確率的に決まる

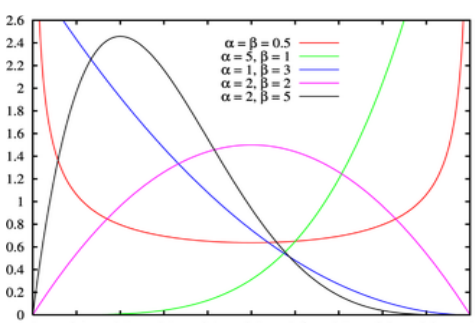
$$\tilde{\theta}_k \sim \mathrm{Beta}(1, \alpha)$$

平均的には
1:αで折る

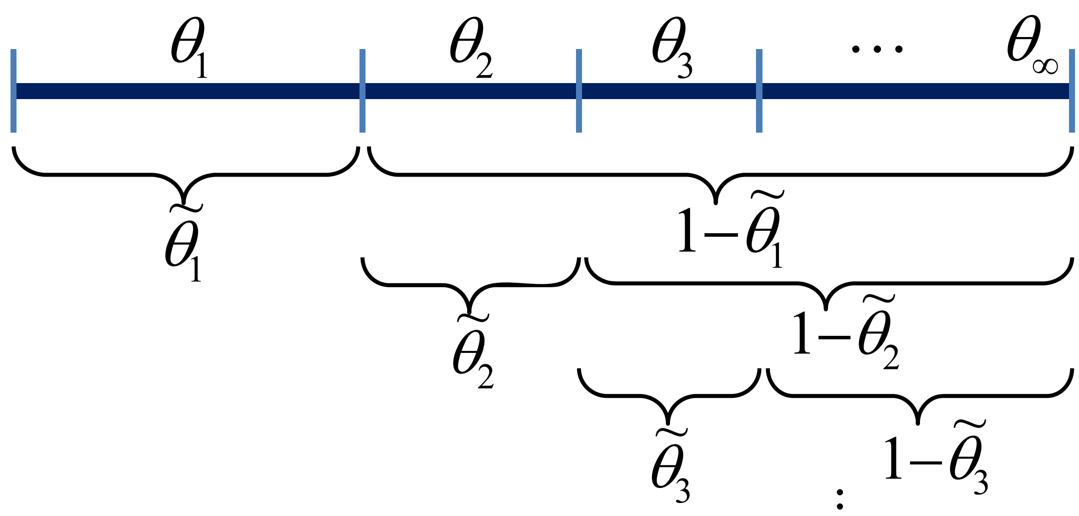
$$\theta_k = \tilde{\theta}_k \prod_{l=1}^{k-1} (1-\tilde{\theta}_l)$$

$\Theta \sim \mathrm{GEM}(\alpha)$ と表記

$\mathrm{Beta}(\alpha, \beta)$

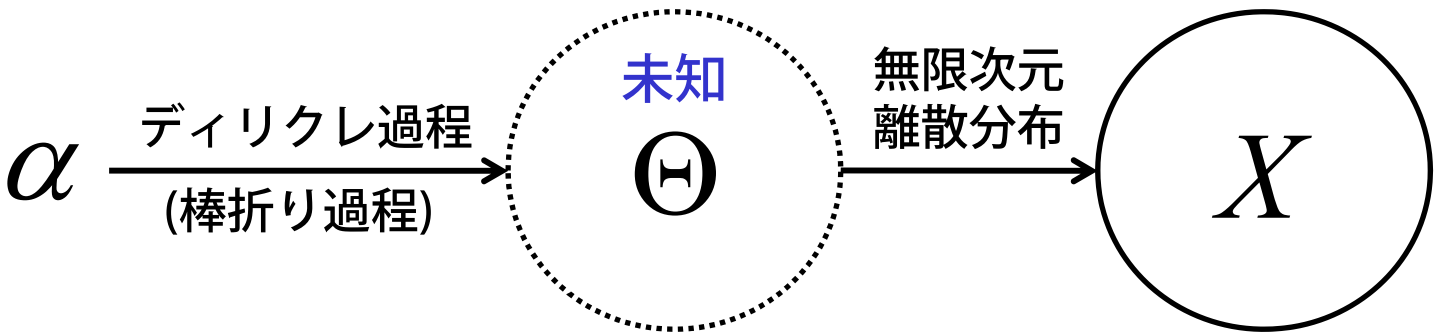


(0,1)上の確率分布



「無限」の取り扱い

- 計算機で無限個のパラメータの値を取り扱うのは困難
  - サイコロの目の出方自体に着目してみる
  - Vapnikの原理
    - ある問題を解くとき、その問題よりも難しい問題を途中段階で解いてはならない



$\{\theta_1, \theta_2, \theta_3, \cdots \theta_\infty\}$　　$\{1, 2, 2, 1, 2, 1, 3, 2\}$

いきなりサイコロの目が分かる！　Θを積分消去（あらゆる可能性を考える）

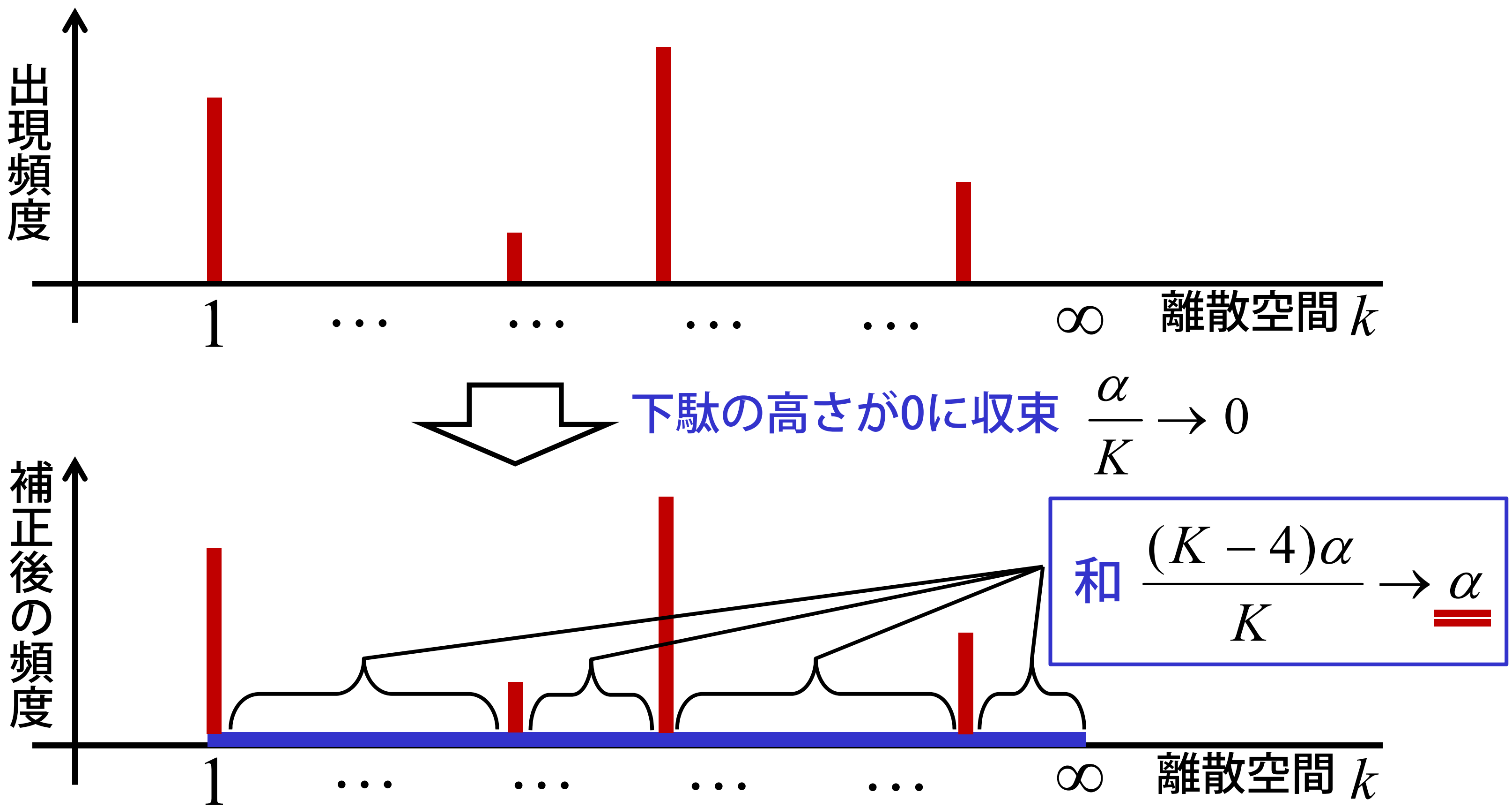
$$\underbrace{p(X|\alpha)}_{\text{中華料理店過程}} = \int p(X|\Theta)\underbrace{p(\Theta|\alpha)}_{\text{棒折り過程}}d\Theta$$

無限次元ディリクレ分布再考

- ディリクレ分布の次元 K を無限に発散
  - 下駄の高さはゼロに収束
  - 未観測の部分の下駄の高さ合計はゼロではない！

$$\mathrm{Dir}\left(\Theta \middle| \alpha \underline{\boldsymbol{1}}\right) \quad K \to \infty$$

∞次元一様分布
(全要素が1/∞)



中華料理店過程 (CRP)

- 無限次元離散分布からの<u>サンプル (頻度)</u> に着目
  - 「The rich get richer」の法則
    - 次の目の出方は過去に出た目の頻度に比例する
  - N回の試行なら高々N種類しか出現しない

新たなクラスのために
小さな確率をリザーブ
(これまでの客数が多く
なるほど出にくくなる)

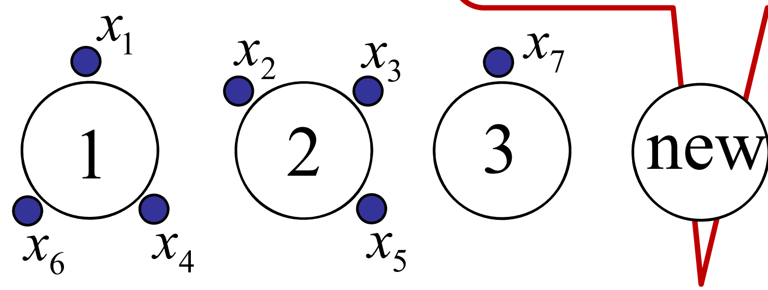
観測データ $X = \{1, 2, 2, 1, 2, 1, 3, 2\}$

パラメータ $\Theta = \{\theta_1, \theta_2, \theta_3, \cdots \theta_\infty\}$

確率モデル

$$p(\Theta) = \mathrm{DP}(\alpha, \boldsymbol{1})$$

$$p(X \mid \Theta) = \prod_{k=1}^{\infty} \theta_k^{n_k}$$



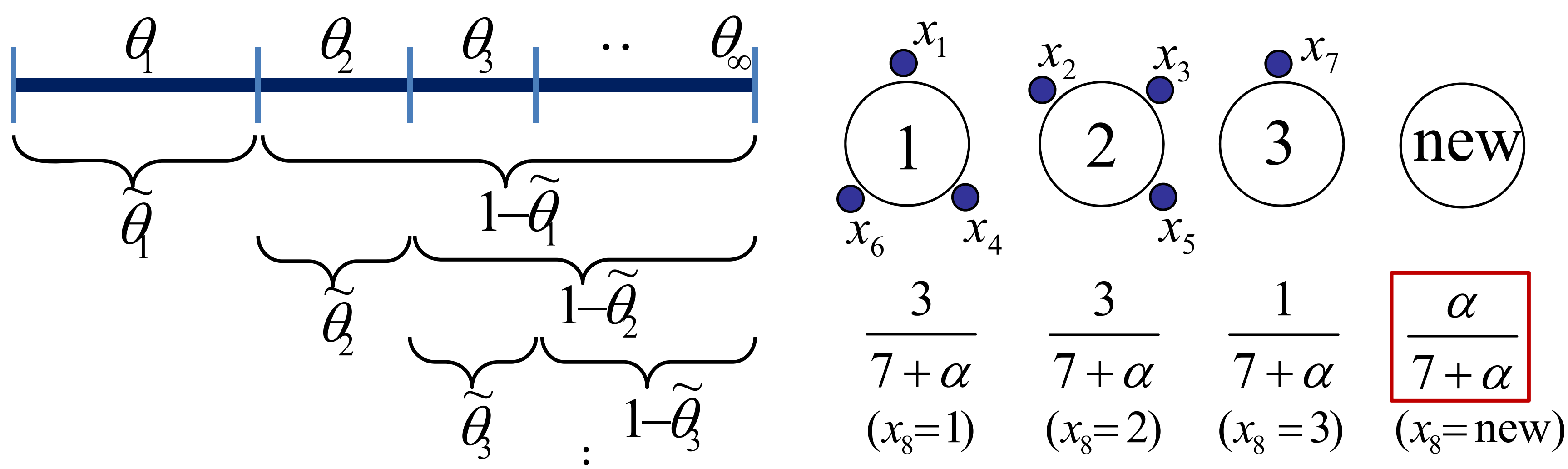
$$p(x_8 \mid x_{1:7}) = \underset{(x_8=1)}{\frac{3}{7+\alpha}} \quad \underset{(x_8=2)}{\frac{3}{7+\alpha}} \quad \underset{(x_8=3)}{\frac{1}{7+\alpha}} \quad \underset{(x_8=\mathrm{new})}{\boxed{\frac{\alpha}{7+\alpha}}}$$

$$p(X \mid \alpha) = p(x_1) p(x_2 \mid x_1) p(x_3 \mid x_{1:2}) \cdots p(x_N \mid x_{1:N-1})$$

$X \sim \mathrm{CRP}(\alpha)$ と表記

ディリクレ過程の表現方法

- 棒折過程 (Stick-Breaking Process: SBP) $\Theta \sim \mathrm{GEM}(\alpha)$
  - 無限次元のパラメータを生成する仕組み
    - 無限面のサイコロ自体を考える
- 中華料理店過程 (Chinese Restaurant Process: CRP)
  - 観測データを直接生成できる仕組み $X \sim \mathrm{CRP}(\alpha)$
    - 無限面のサイコロを振って出た目を考える



潜在変数モデル

- 観測データだけでなく非観測データも考える
  - 機械学習における中心的な確率モデル (例: 混合モデル)

例：性別ラベルが分からない身長データ

観測変数X：身長 (ガウス分布に従う)

潜在変数Z：男 or 女 (2 次元離散分布に従う)

パラメータΘ：2つのガウス分布の平均と分散・混合比

⇩ クラス数が未知であれば？

例：クラスラベルが分からない特徴量データ

観測変数X：特徴量 (ガウス分布に従う)

潜在変数Z：クラス (無限次元離散分布に従う)

パラメータΘ：無限個のガウス分布の平均と分散・混合比

潜在変数Zに対する事前分布としてDPが利用可能！
→ 無限混合ベイズモデル (iGMM)

ディリクレ過程の応用

- 無限混合ベイズモデルをどう定式化するか
  - ディリクレ過程の基底測度を要素分布に合わせて設計する

例：クラスラベルが分からない特徴量データ

観測変数X：特徴ベクトル (ガウス分布に従う)

潜在変数Z：クラス (無限次元離散分布に従う)

パラメータΘ：無限個のガウス分布の平均と分散・混合比

$$p(x \mid \Theta) = \sum_{k=1}^{\infty} \pi_k N(x \mid \mu_k, \sigma_k^2)$$

無限個の混合比は
棒折り過程でOK

無限個のガウス分布は
どうやって生成するの？

→ ガウス・ウィシャート分布
(ガウス分布に対する共役事前分布)

$$p(\Theta) = \mathrm{DP}(\alpha, G_0)$$

基底測度：合わせて表記する

中華料理店過程表現

- テーブルの料理を生成する機構 ＝ 基底測度 G_0

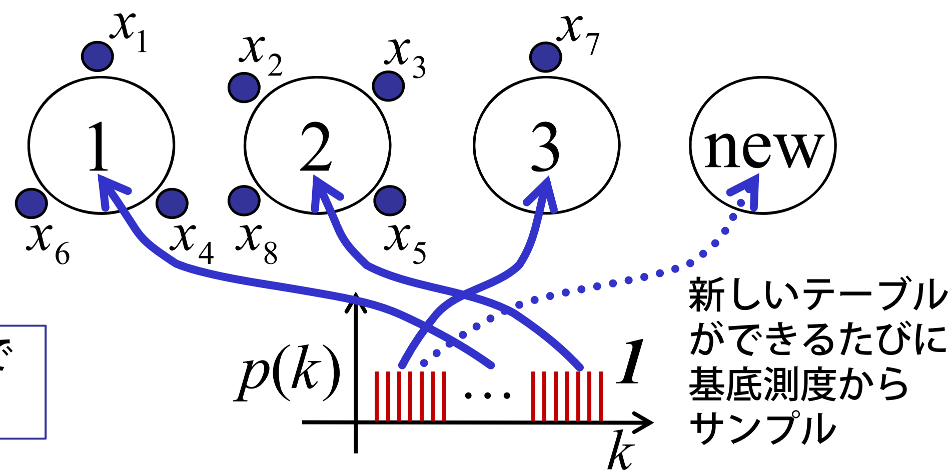
無限面サイコロの場合

$\mathrm{DP}(\alpha, \mathbf{1})$

観測変数：サイコロの目

$X = \{1, 2, 2, 1, 2, 1, 3, 2\}$

正整数に対するラベル付けは任意なので
出現した目の種類の順に1,2,3…とする



無限混合モデルの場合

$\mathrm{DP}(\alpha, G_0)$

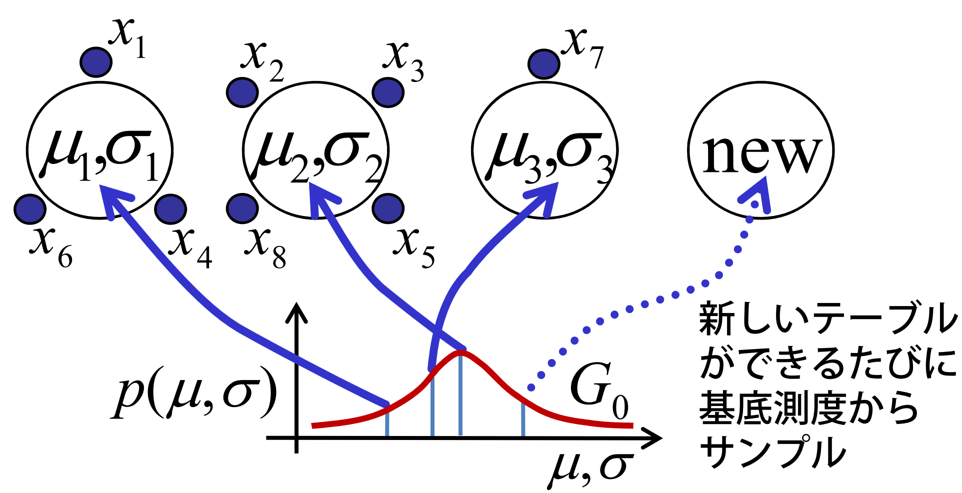
潜在変数：クラス

$Z = \{1, 2, 2, 1, 2, 1, 3, 2\}$

対応するガウス分布から生成

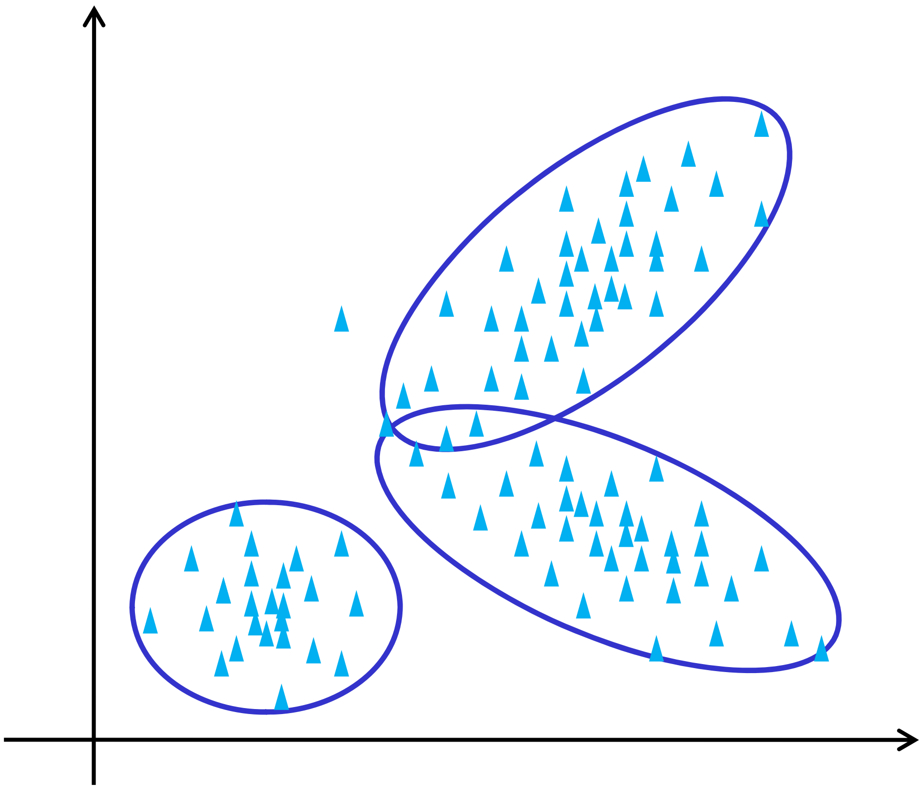
$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$

観測変数：特徴量



Infinite GMM

- 可算無限個のクラスを許容する
  ベイズ混合ガウスモデル [Rasmussen2000]
  - ディリクレ過程 (DP) を利用



有限混合モデル：
1. クラス数Kを変化させて
   大量のモデルを学習
2. AICやBICなどの指標を用いて
   適切なものをあとから選択

無限混合モデル：
一度の学習でクラス数の
事後分布が求まる
(必要ならば事後確率の高い
クラス数を求めることも可能)

Infinite HMM

- 可算無限個の状態(と出力シンボル)を許容するベイズ隠れマルコフモデル
  - 階層ディリクレ過程 (HDP) を利用
  - ギブスサンプリング [Beal2001]
    - 各時刻の隠れ状態を反復的にサンプル
      - 十分時間をかければ真の事後分布に収束
      - HMMは隣同士の相関が強くて収束が非常に遅い
  - ビームサンプリング [VanGael2007]
    - 問題：状態数が無限なので動的計画法が実行できない
      - かといって打ち切り近似はしたくない・・・
    - 解決法：スライスサンプリングを組み合わせる
      - ある閾値以下の確率の遷移を無視する
      - 閾値も確率的にサンプルすることで真の事後分布に収束

確率的文脈自由文法 (PCFG)

- シンボルの導出規則に確率が付与されたもの
  - 予めチョムスキー標準系に変換しておく

S → NP VP	1.0
PP → P NP	1.0
VP → V NP	0.7
VP → VP PP	0.3
P → *with*	1.0
V → *saw*	1.0

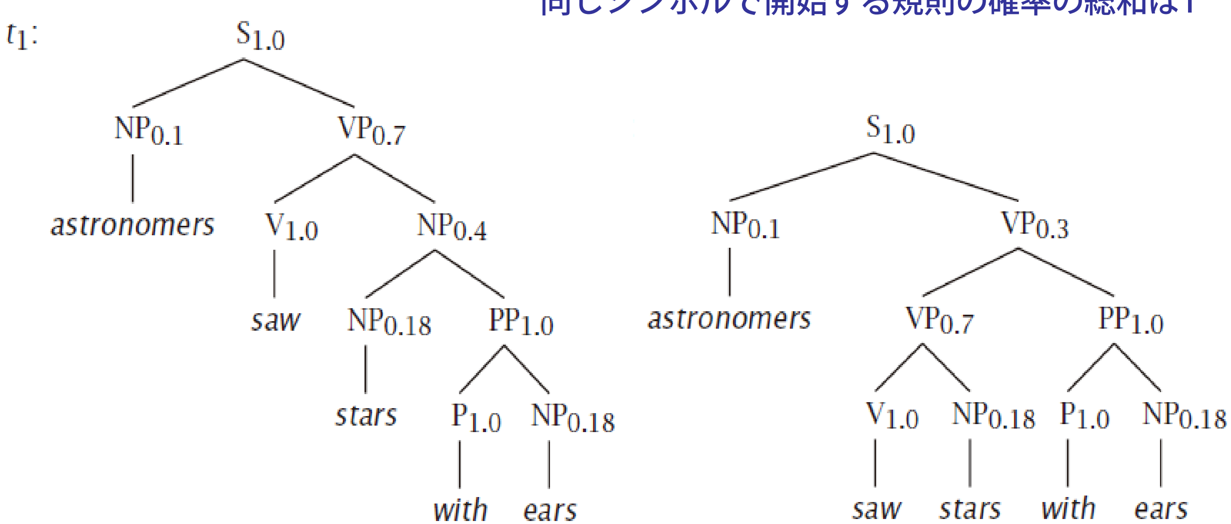
NP → NP PP	0.4
NP → *astronomers*	0.1
NP → *ears*	0.18
NP → *saw*	0.04
NP → *stars*	0.18
NP → *telescopes*	0.1

同じシンボルで開始する規則の確率の総和は1

PCFGの最尤学習：
単語列が与えられた時に
各導出規則の確率を求める

PCFGのベイズ学習：
各シンボルごとに導出確率の
ディリクレ事後分布を求める

[栗原2004]



$$
\begin{aligned}
P(t_1) &= 1.0 \times 0.1 \times 0.7 \times 1.0 \times 0.4 \\
&\quad \times 0.18 \times 1.0 \times 1.0 \times 0.18 \\
&= 0.0009072 \\
P(t_2) &= 1.0 \times 0.1 \times 0.3 \times 0.7 \times 1.0 \\
&\quad \times 0.18 \times 1.0 \times 1.0 \times 0.18 \\
&= 0.0006804 \\
P(w_{15}) &= P(t_1) + P(t_2) = 0.0015876
\end{aligned}
$$

Infinite PCFG

- 可算無限個のシンボルおよび導出規則を許容するベイズ文脈自由文法 [Liang2007]
  - 階層ディリクレ過程 (HDP) を利用
  - 手動で導出規則を与えなくてよい
    - 必要なシンボル・必要な導出規則が自動的に生成

有限モデル

S → NP VP	1.0	NP → NP PP	0.4
PP → P NP	1.0	NP → *astronomers*	0.1
VP → V NP	0.7	NP → *ears*	0.18
VP → VP PP	0.3	NP → *saw*	0.04
P → *with*	1.0	NP → *stars*	0.18
V → *saw*	1.0	NP → *telescopes*	0.1

無限モデル

S1 → S1 S2

S1 → S1 S5

S2 → S2 S3

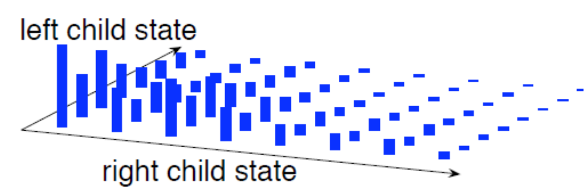
⋮

1. 可算無限個のシンボルを生成 (DP)

$\beta \sim \mathrm{GEM}(\alpha)$

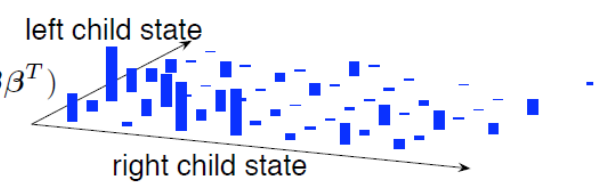
2. シンボルの組に対する基底測度を準備

$\beta\beta^T$



3. 導出規則の右側の分布を生成 (DP)

$\phi_z^B \sim \mathrm{DP}(\beta\beta^T)$



最近のトレンド

- ベータ過程 (BP) を用いた研究が増加
  - 無限種類の因子 (factor/feature) を考える
    - 潜在変数：サンプルごとの各因子の有無（無限次元）
    - 因子数がKであれば、2のK乗の状態が表現可能
  - 利用しやすいサンプリングスキームが存在
    - ディリクレ過程 (DP) → 中華料理店過程 (CRP)
    - ベータ過程 (BP) → インド料理過程 (IBP)

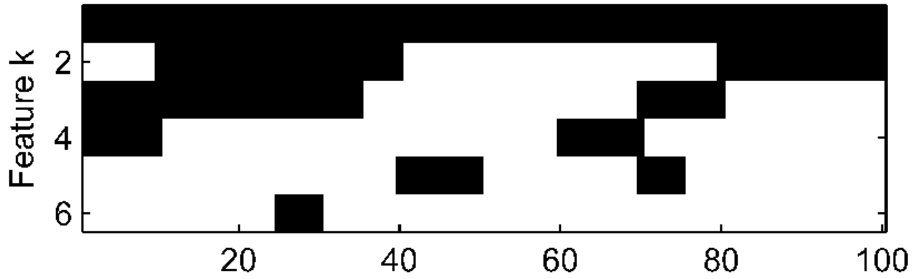
混合モデル (Latent Class Model)
観測データ中の各サンプルはどれか一つのクラスから生成

因子モデル (Latent Feature Model)
観測データ中の各サンプルは複数の因子から生成
例：主成分分析 (PCA)　因子分析 (FA)　独立成分分析 (ICA)
　　確率的行列分解 (PMF)　非負値行列分解 (NMF)

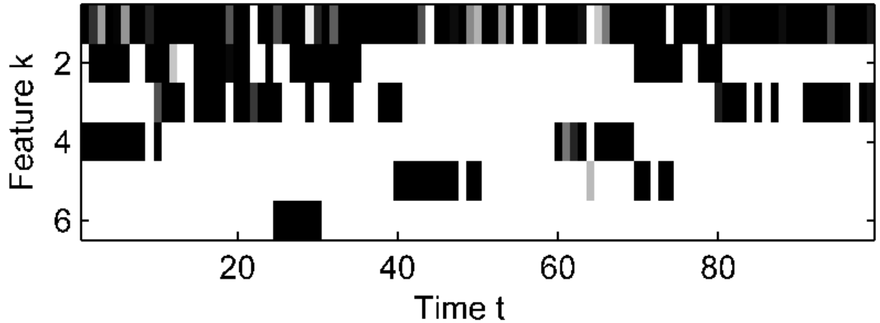
Infinite ICA/SFA

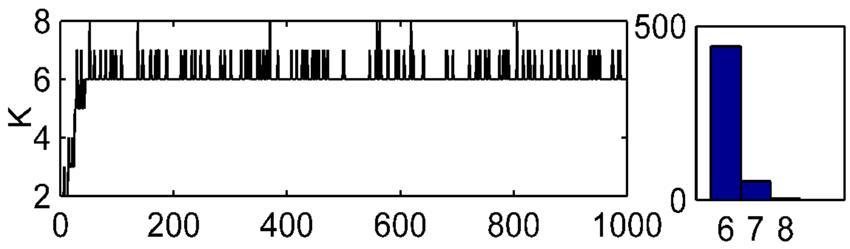
- 可算無限個の信号源を許容する独立成分分析およびスパース因子分析 [Knowles2007]
  - インド料理過程 (IBP) を利用
    - ある客はそれまでの人気に比例した確率で料理を**複数**選ぶ
    - さらに新しい料理にも確率的にチャレンジする

正解



推定結果



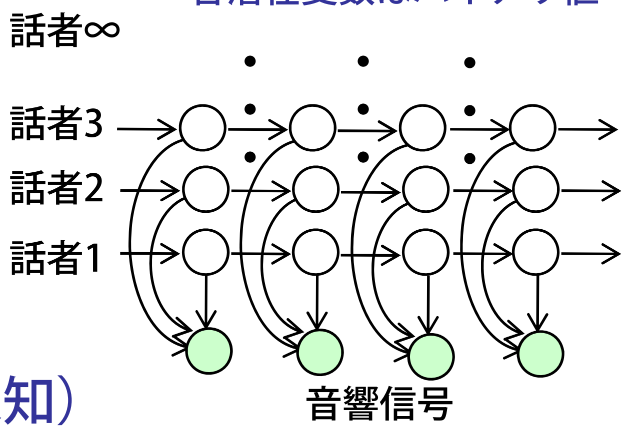


MCMCによる推論：
必要な信号源の数 (料理の種類数) を
増減させながら事後分布空間を探索

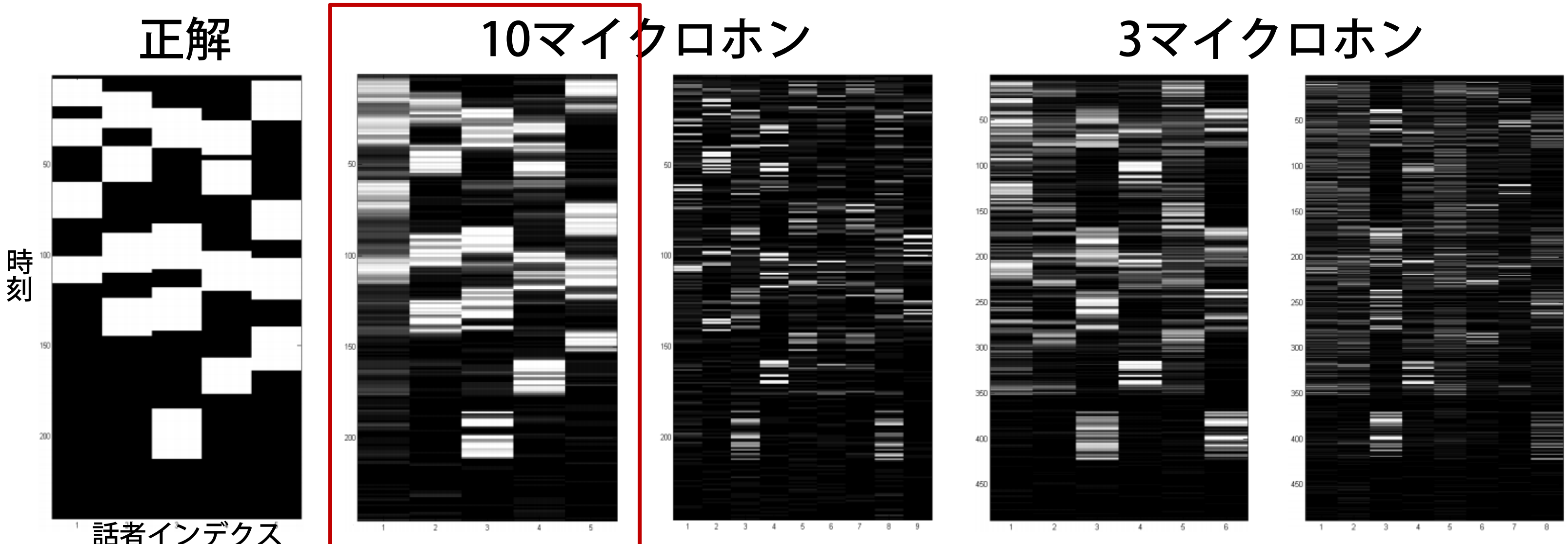
Infinite Factorial HMM

- 可算無限個の隠れマルコフチェインを許容するHMM [VanGael2009]
  - マルコフインド料理過程 (mIBP)
  - ICAと組み合わせることが可能
    - <u>理論的にはNMFとの組み合せも可能</u>

各潜在変数はバイナリ値



混合音からの話者区間推定実験（話者数未知）



(a) Ground Truth (b) ICA iFHMM (c) iICA (d) ICA iFHMM (e) iICA

音楽情報処理への応用

私の問題意識：教師なし音楽理解

- 音楽的に訓練されていない人であっても音楽を楽しめる
  - 混合音中に含まれる複数の音符を聴き分けることができる
    - 音符の個数は事前には分からない
  - 音符の組み合わせが調和するかしないか直感的に判断できる
    - コード名は知ってなくてもよい
- 音楽的な感覚を獲得できる計算機をつくりたい
  - 確率モデルに基づく構成論的アプローチ
  - 音符上の構造パターンの教師なし学習
  - 膨大な音楽鑑賞経験に基づく構造学習
    - 入力：音楽音響信号
    - 出力：音符配置・構造解釈

「次世代型」計算音楽学を確立したい

計算言語学との類似点

- どうやってテキストを分かち書きするか？
  - 普通の発想：Mecab・Chasenなどを使う
    - 有限の語彙セットが必要
    - 学習データとして分かち書きされたテキストが必要
    - 未知語は解析できない！ (twitter・しょこたんブログなど)

	MeCab	ChaSen	JUMAN	KAKASI
解析モデル	bi-gram マルコフモデル	可変長マルコフモデル	bi-gram マルコフモデル	最長一致
コスト推定	コーパスから学習	コーパスから学習	人手	コストという概念無し
学習モデル	CRF (識別モデル)	HMM (生成モデル)		
辞書引きアルゴリズム	Double Array	Double Array	パトリシア木	Hash?
解探索アルゴリズム	Viterbi	Viterbi	Viterbi	決定的?
連接表の実装	2次元 Table	オートマトン	2次元 Table?	連接表無し?
品詞の階層	無制限多階層品詞	無制限多階層品詞	2段階固定	品詞という概念無し?
未知語処理	字種 (動作定義を変更可能)	字種 (変更不可能)	字種 (変更不可能)	
制約つき解析	可能	2.4.0で可能	不可能	不可能
N-best解	可能	不可能	不可能	不可能

Mecabサイト (http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html) より抜粋

私が感銘を受けた研究：教師なし単語分割

- 英文「Alice in Wonderland」を解析

first,shedreamedoflittlealiceherself,andonceagainthetinyhandswereclaspeduponher
knee,andthebrighteagereyeswerelookingupintohersshecouldheartheverytonesofher
voice,andseethatqueerlittletossofherheadtokeepbackthewanderinghairthatwouldal
waysgetintohereyesandstillasshelistened,orseemedtolisten,thewholeplacearoundhe
rbecamealivethestrangecreaturesofherlittlesister'sdream.thelonggrassrustledatherfe
EtasthewhiterabbithurriedBythefrightenedmousesplashedhiswaythroughtheneighb
ouringpoolshecouldheartherattle…

ノンパラメトリックベイズ教師なし単語分割 [持橋 2009]

first, she dream ed of little alice herself ,and once again the tiny hand s were clasped upon her knee ,and the bright eager eyes were looking up into hers shecould hearthe very tone s of her voice , and see that queer little toss of herhead to keep back the wandering hair that would always get into hereyes and still as she listened , or seemed to listen , thewhole place a round her became alive the strange creatures of her little sister 'sdream. thelong grass rustled ather feet as thewhitera bbit hurried by the frightened mouse splashed his way through the neighbour ing pool shecould hearthe rattle…

私が感銘を受けた研究：教師なし単語分割

- 古文「源氏物語」を解析

しばしは夢かとのみたどられしを、やうやう思ひしづまるにしも、さむべき方なくたへがたきは、いかにすべきわざにかとも、問ひあはすべき人だになきを、忍びては参りたまひなんや。若宮の、いとおぼつかなく、露けき中に過ぐしたまふも、心苦しう思さるるを、とく参りたまへ』など、はかばかしうも、のたまはせやらず、むせかへらせたまひつつ、かつは人も心弱く見たてまつるらむと、思しつつまぬにしもあらぬ御気色の…

ノンパラメトリックベイズ教師なし単語分割 [持橋 2009]

しばし/は/夢 か/と/のみ/たど られ/し/を/、/やうやう/思ひ/しづま/る/に/しも/、/さむ/べき/方/な/く/たへ/がた/き/は/、/いかに/す/べき/わざ/に/かと/も/、/問ひ/あは/す/べき/人/だに/な/き/を/、/忍びて/は/参り/たまひ/なん や/。/若 宮/の/、/いと/おぼつかな/く/、/露け/き/中に/過ぐし/たまふ/も/、/心/苦し/う/思さる/る/を/、/とく/参り/たまへ/など/、/はかばかしう/も/、/のたまはせ/やら/ず/、/むせ かへ/ら/せ/たまひ/つつ/、/かつ/は/人も心 弱/く/見/たてまつ/る/ら/む/と/、/思しつつ/ま/ぬ/に/しも/あら/ぬ/御気色/の/…

教師あり学習

- なんらかの事前知識や教師データが必要
  - 自動採譜・音源分離
    - 音符の個数は既知と仮定
    - あらかじめ楽器音や歌声の特徴を学習しておく手法も
  - コード認識 (分類タスク)
    - 分類すべきコード名の語彙をあらかじめ準備
      (C maj, D min, E dim, F aug, G maj7, A minmaj7など)

学習フェーズ

ラベル付きの音響信号から
確率モデルを学習

予測フェーズ

ラベルが付与されていない
音響信号に対して
最尤のラベルを付与

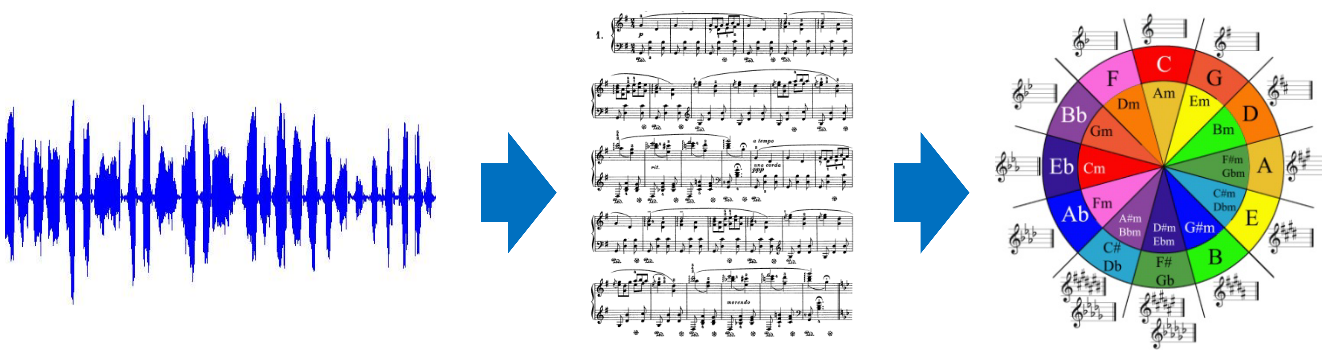
未知語問題！

教師なし学習

- 音響信号だけから含まれている音符を発見し (自動採譜)
  音符上の構造を推論したい (文法獲得)
  - 事前知識は必要なし
    - 音符数もコード語彙も必要ない
  - 観測データは混合音のみ
    - 音響信号から適切な個数の音符を見つける
    - コードの概念 (典型的な音符の組み合わせ) を獲得する

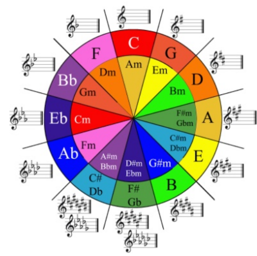
未知語問題がない！

  - 音響信号の最も良い解釈を見つけたい：確率に基づく評価！
    - 階層構造：信号 → 楽譜 → 文法



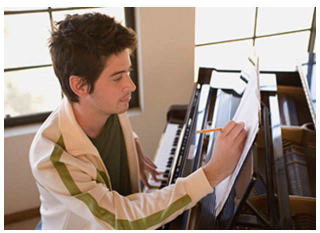
音楽の生成過程

- 教師なし音楽理解 ＝ 逆問題に対する解
  - 音楽音響信号が与えられたときに生成過程を推論したい



文法
(音楽理論)

潜在変数



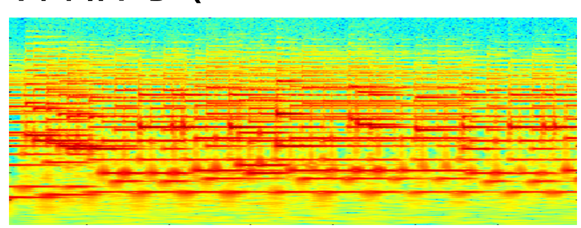
作曲

潜在変数

楽譜 (音符配置)

演奏

音響信号 (スペクトログラム)



生成過程

構造学習

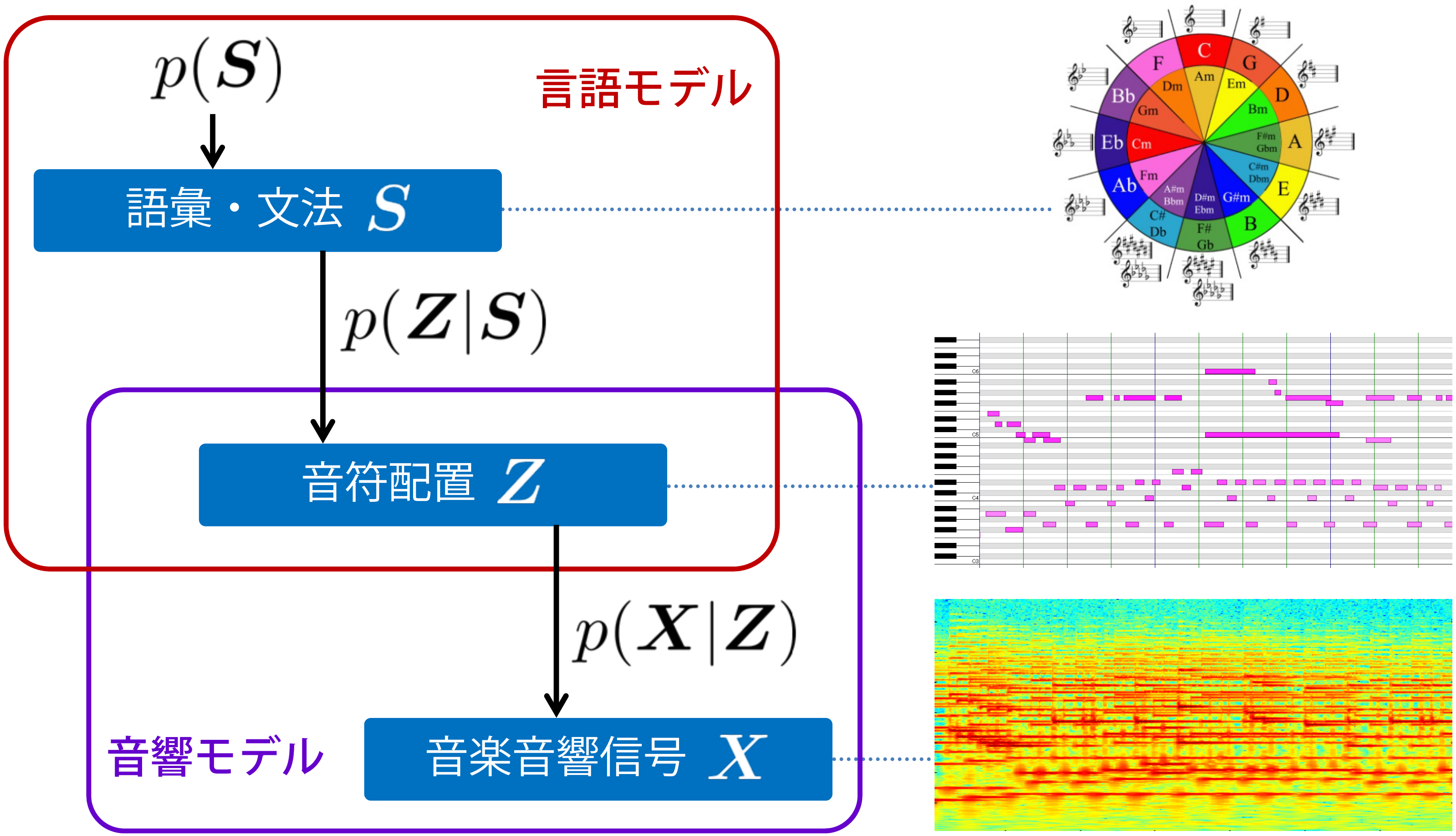
構造学習

事後推論

確率的な定式化

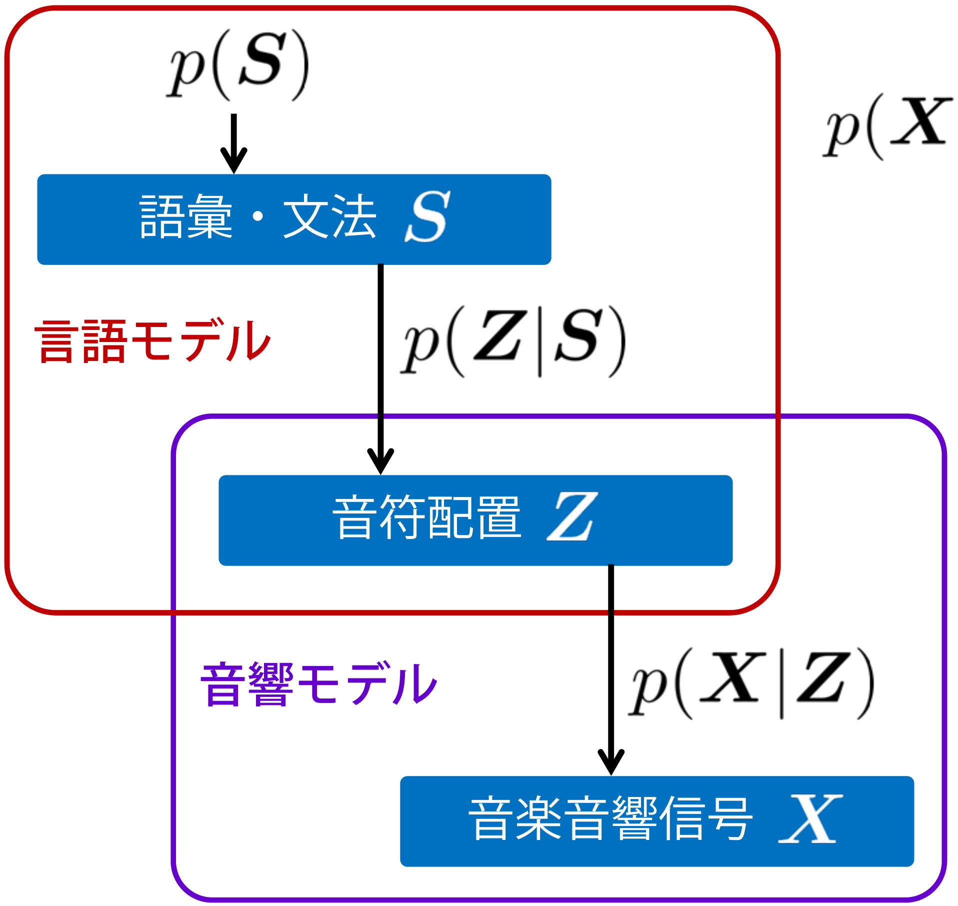
- 階層ベイズモデルによる統一的記述
  - 音響モデルと言語モデルとの統合モデル

音声認識と同様の枠組み
ただし語彙や文法は未知！



グランドチャレンジ

- 目標：階層ベイズモデルの事後分布計算 = 構造学習
  - 潜在変数 (文法・音符) のあらゆる可能性を考慮に入れる



同時分布

$$p(\boldsymbol{X},\boldsymbol{Z},\boldsymbol{S}) = p(\boldsymbol{X}|\boldsymbol{Z})p(\boldsymbol{Z}|\boldsymbol{S})p(\boldsymbol{S})$$

事後分布

$$p(\boldsymbol{Z},\boldsymbol{S}|\boldsymbol{X}) = \frac{p(\boldsymbol{X},\boldsymbol{Z},\boldsymbol{S})}{p(\boldsymbol{X})}$$

周辺尤度 (エビデンス)

$$p(\boldsymbol{X}) = \int p(\boldsymbol{X},\boldsymbol{Z},\boldsymbol{S})d\boldsymbol{Z}d\boldsymbol{S}$$

エビデンスを最大化するような
モデルをどう設計すればよいか？

モデル選択

- エビデンスを最大化するモデルの複雑さを求めたい
  - 音符を何個に設定すればいいか？
  - コード名を何個考慮すればいいか？

組み合わせ最適化における
グリッドサーチは
計算量的に非現実的！

音響モデルの複雑さ

音符の個数

言語モデルの複雑さ

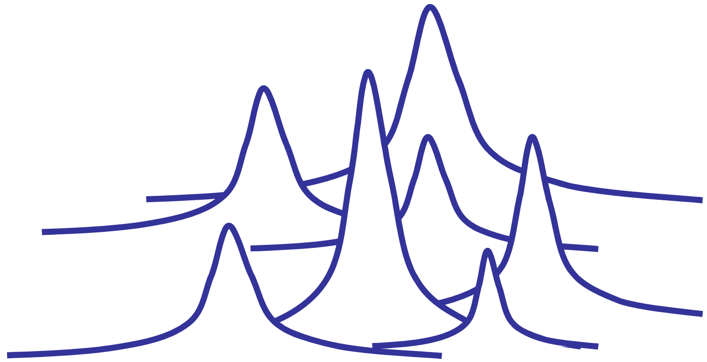
コード名の種類数

	10	20	30	40	…
10	-20,000	-9,000	-8,000	-8,000	
20	-10,000	-8,500	-7,000	-7,500	
30	-9,000	-8,300	-7,500	-7,900	
40	-8,800	-8,400	-8,000	-8,500	
…					

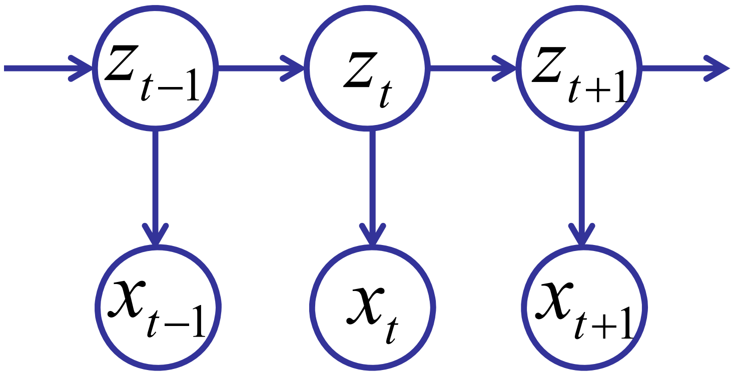
ノンパラメトリックベイズとは？

- 観測データに合わせてモデルの複雑さを自動調節可能
  - 無限ガウス混合モデル (iGMM)
    - 混合数を自動的に調節可能
  - 無限隠れマルコフモデル (iHMM)
    - 状態数を自動的に調節可能
  - 無限非負値行列分解 (iNMF)
    - 基底数を自動的に調節可能

構造学習への応用では
極めて強力な武器！



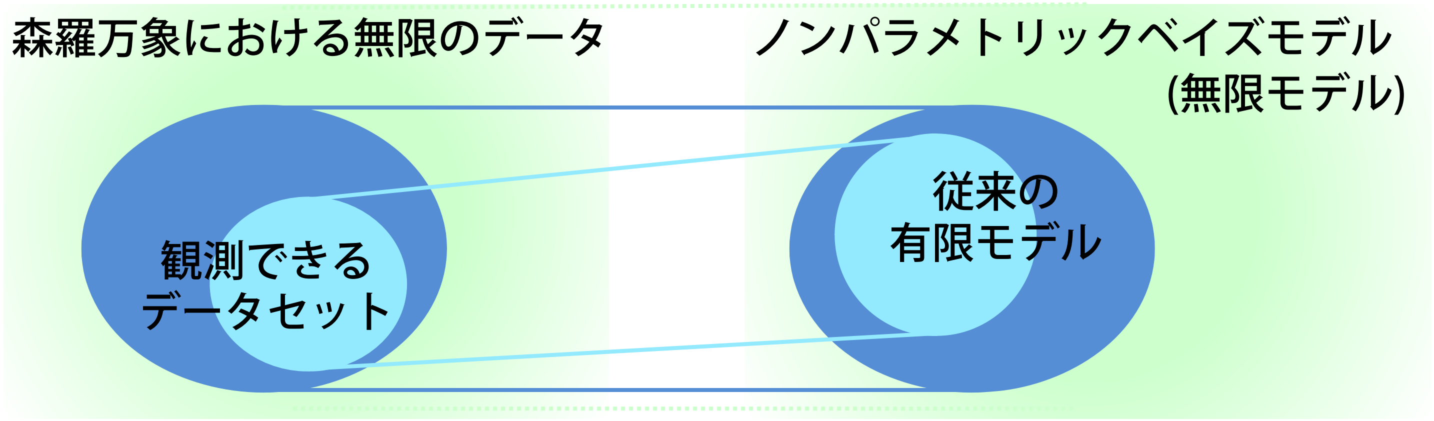
GMM



HMM

なぜノンパラメトリックベイズか？

- 構造学習に対する理論的に確立されたアプローチ
  - 組み合わせ最適化をしなくてよい！
    - モデルの構造とパラメータの同時学習が可能
    - 無限次元空間におけるスパース学習
  - モデルの「実効的」複雑さが自動的に調節される
    - 無限にデータがあれば無限の複雑さが必要 (「真の」複雑さ)
    - 手持ちの観測データが有限であれば、一部だけでよい



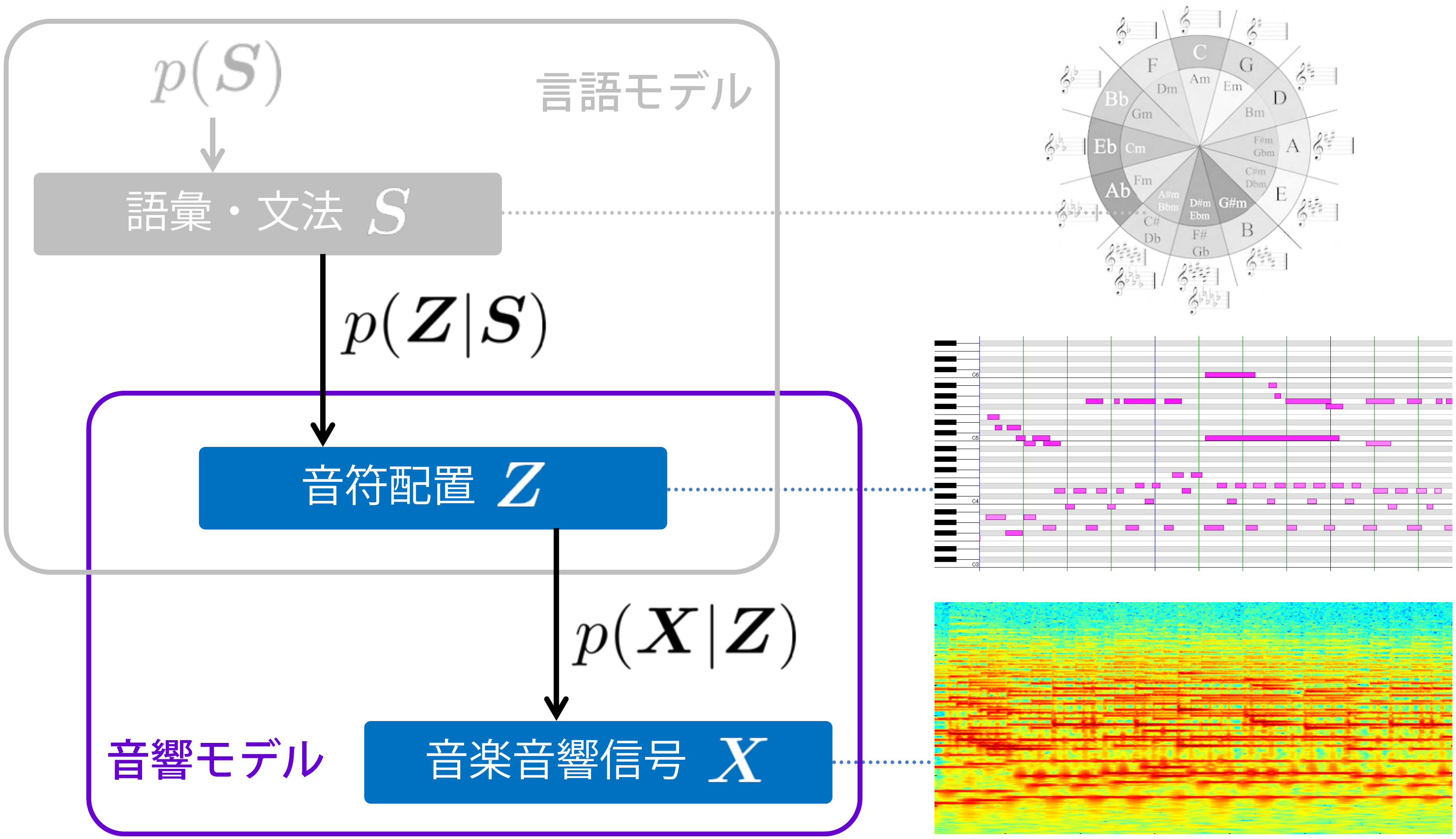
これまでの研究

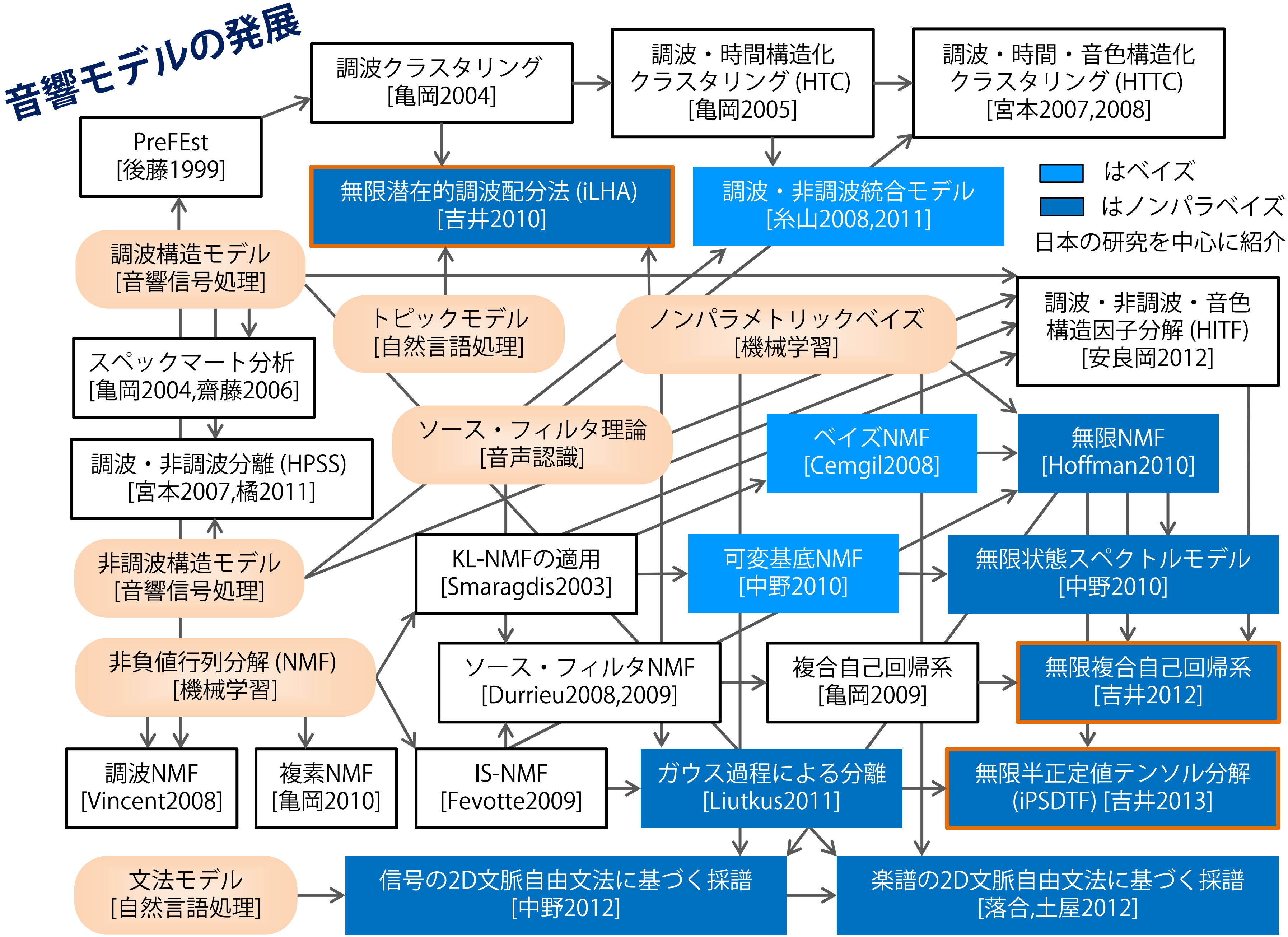
- ノンパラメトリックベイズ音響・言語モデルの開発
  - 音響データに対する構造学習
    - 無限潜在的調波配分法 [Yoshii@IEEE Trans. 2012 (ISMIR 2010)]
    - 無限複合自己回帰モデル [Yoshii@ISMIR 2012]
    - 無限半正定値テンソル分解 [Yoshii@ICML 2013 (ISMIR 2013)]
  - 楽譜データに対する構造学習
    - 語彙フリー無限グラムモデル [Yoshii@ISMIR 2011]

音響モデル / 言語モデル	混合モデル (例: GMM)	因子モデル (例: NMF)
連鎖モデル (例: Nグラムモデル)	本研究	本研究
木構造モデル (例: 文脈自由文法)	未提案	Nakano@ICASSP 2012 Kameoka@ISMIR 2012

Big Picture

- 階層ベイズモデルによる統一的記述
  - 音響モデルと言語モデルとの統合モデル

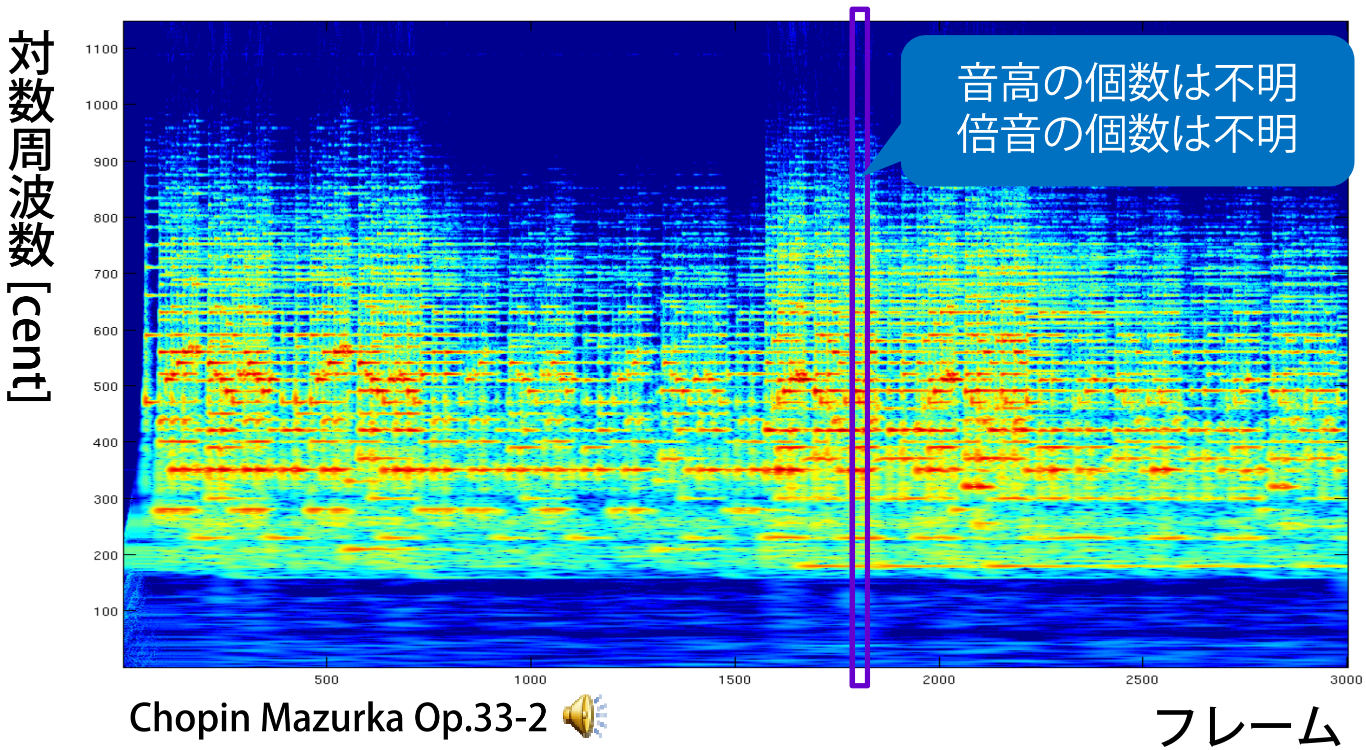



音響モデルの発展
調波クラスタリング [亀岡2004]
調波・時間構造化 クラスタリング (HTC) [亀岡2005]
調波・時間・音色構造化 クラスタリング (HTTC) [宮本2007,2008]
PreFEst [後藤1999]
無限潜在的調波配分法 (iLHA) [吉井2010]
調波・非調波統合モデル [糸山2008,2011]
はベイズ
はノンパラベイズ
日本の研究を中心に紹介
調波構造モデル [音響信号処理]
トピックモデル [自然言語処理]
ノンパラメトリックベイズ [機械学習]
調波・非調波・音色 構造因子分解 (HITF) [安良岡2012]
スペックマート分析 [亀岡2004,齋藤2006]
ソース・フィルタ理論 [音声認識]
ベイズNMF [Cemgil2008]
無限NMF [Hoffman2010]
調波・非調波分離 (HPSS) [宮本2007,橘2011]
非調波構造モデル [音響信号処理]
KL-NMFの適用 [Smaragdis2003]
可変基底NMF [中野2010]
無限状態スペクトルモデル [中野2010]
非負値行列分解 (NMF) [機械学習]
ソース・フィルタNMF [Durrieu2008,2009]
複合自己回帰系 [亀岡2009]
無限複合自己回帰系 [吉井2012]
調波NMF [Vincent2008]
複素NMF [亀岡2010]
IS-NMF [Fevotte2009]
ガウス過程による分離 [Liutkus2011]
無限半正定値テンソル分解 (iPSDTF) [吉井2013]
文法モデル [自然言語処理]
信号の2D文脈自由文法に基づく採譜 [中野2012]
楽譜の2D文脈自由文法に基づく採譜 [落合,土屋2012]

ノンパラメトリックベイズ音響モデル

- 目標：多重基本周波数推定
  - 混合音中の各フレームごとに発音中の音高を全て列挙する
    - 音高の個数は不明であることに注意

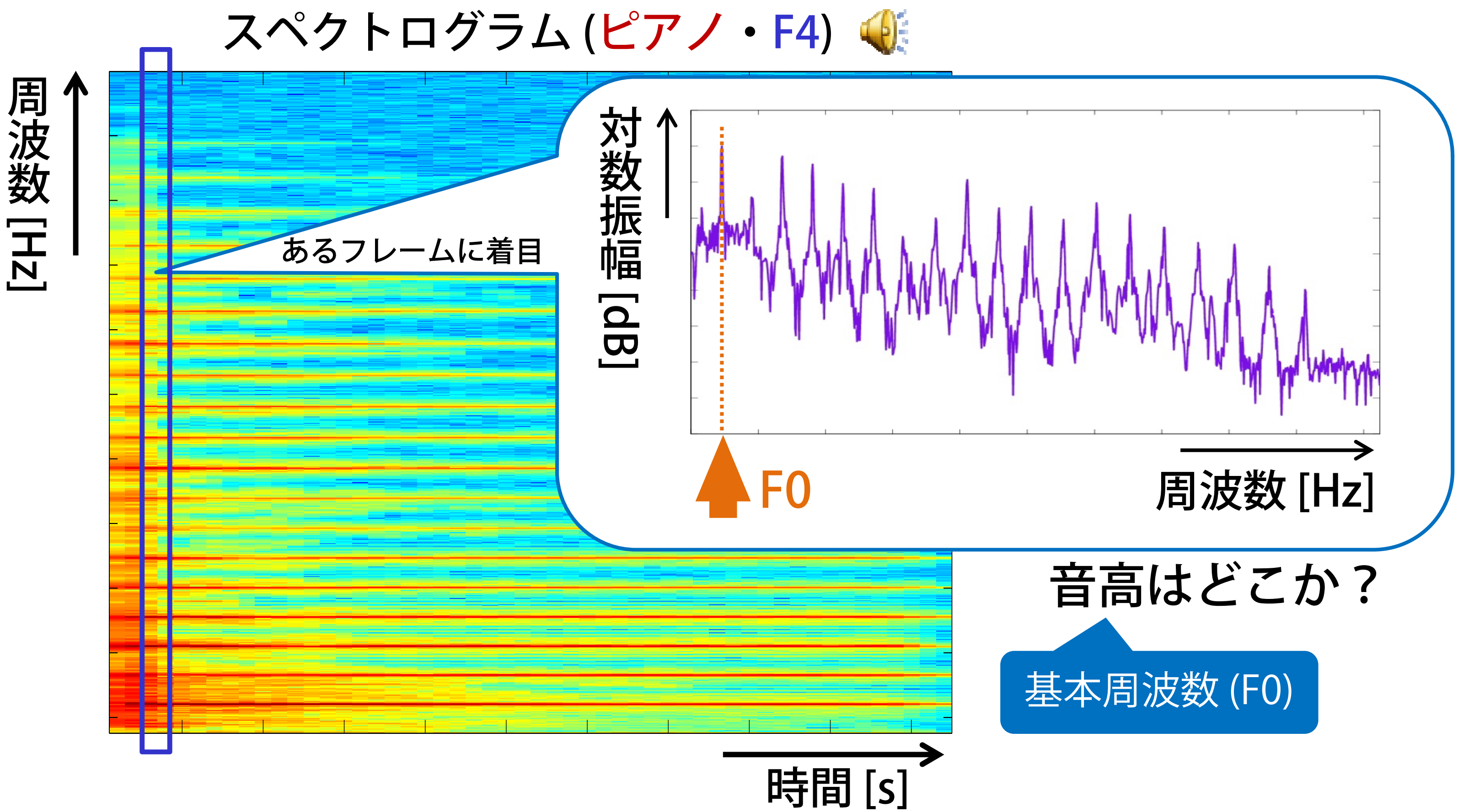
観測データ (ウェーブレットスペクトログラム)



Chopin Mazurka Op.33-2

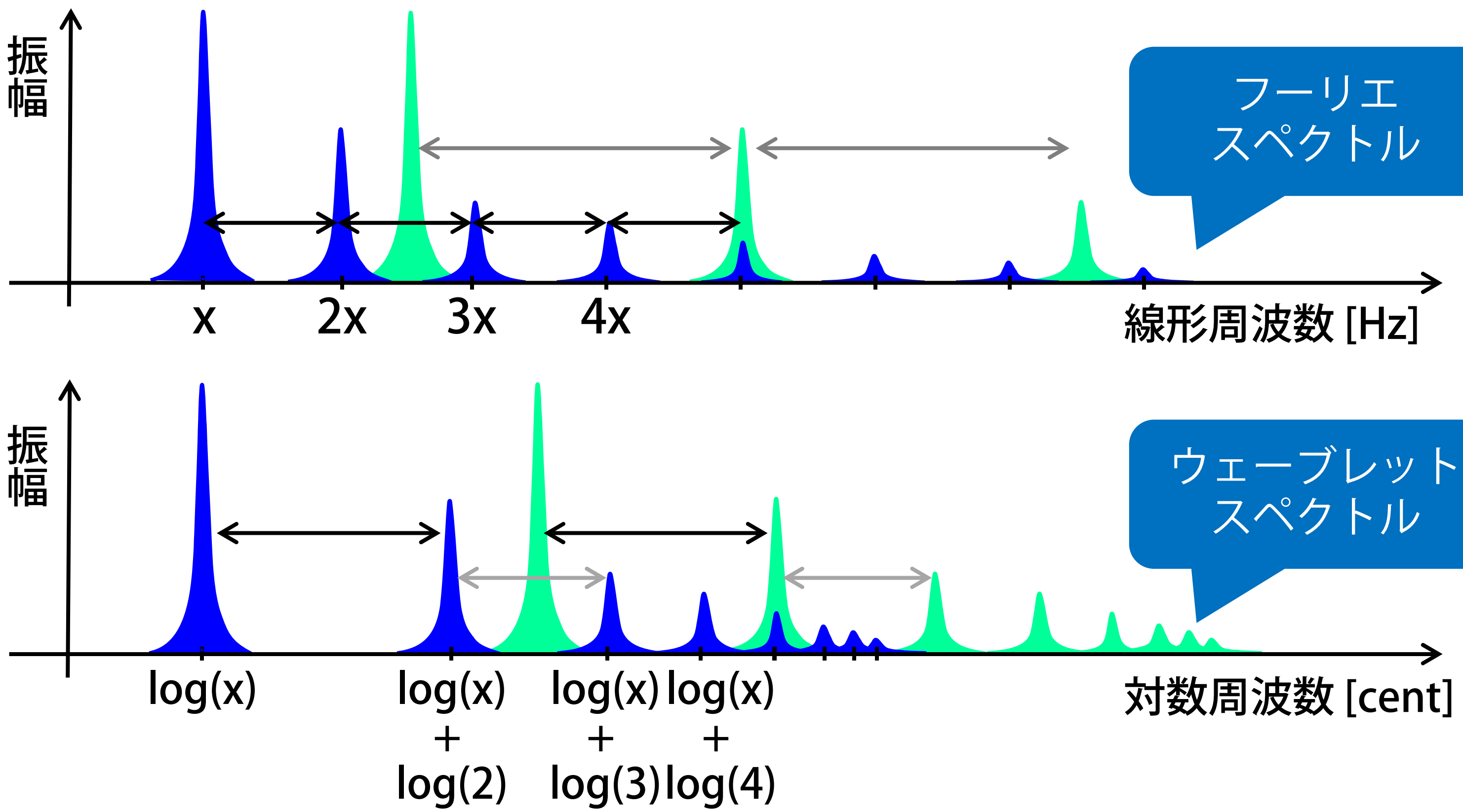
楽器音のスペクトル (1/2)

- 音高をもつ楽器音は調波構造をもつ
  - 周波数方向に等間隔に倍音ピークが存在 (くし型スペクトル)



楽器音のスペクトル (2/2)

- 周波数軸の選択によって着目すべき構造が異なる
  - 線形周波数軸：となりあう倍音間の間隔はF0に等しい
  - 対数周波数軸：F0にかかわらず倍音の相対的な位置は一定



混合モデルに基づくアプローチ

- 観測スペクトルに対するネスト型混合ガウスモデル

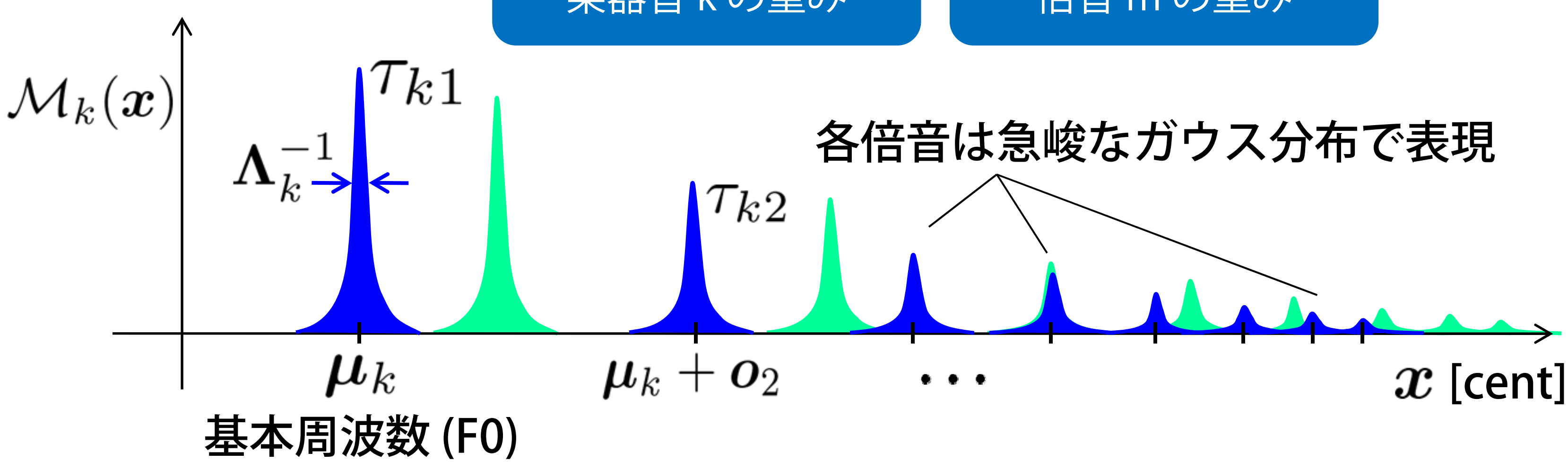
[Goto 2000,2004]
[Kameoka 2004]

  - 各楽器音はM個の倍音から構成
  - 観測スペクトルはK個の楽器音から構成

$$\mathcal{M}_d(\boldsymbol{x}) = \sum_{k=1}^{K} \pi_{dk} \sum_{m=1}^{M} \tau_{km} \mathcal{N}(\boldsymbol{x} | \boldsymbol{\mu}_k + \boldsymbol{o}_m, \boldsymbol{\Lambda}_k^{-1})$$

フレーム d における
楽器音 k の重み

楽器音 k における
倍音 m の重み



提案法：無限潜在的調波配分法

- 無限個の楽器音・倍音から構成されるネスト型GMM [Yoshii 2010,2012]
  - 適切な個数を一挙に推定可能！
    - 楽器音の個数K ＋倍音の個数M

ネスト型GMM

$$\mathcal{M}_d(\boldsymbol{x}) = \sum_{k=1}^{K} \pi_{dk} \sum_{m=1}^{M} \tau_{km} \mathcal{N}\left(\boldsymbol{x} \middle| \boldsymbol{\mu}_k + \boldsymbol{o}_m, \boldsymbol{\Lambda}_k^{-1}\right)$$

KおよびMともに無限に発散した極限を考える

ネスト型無限GMM

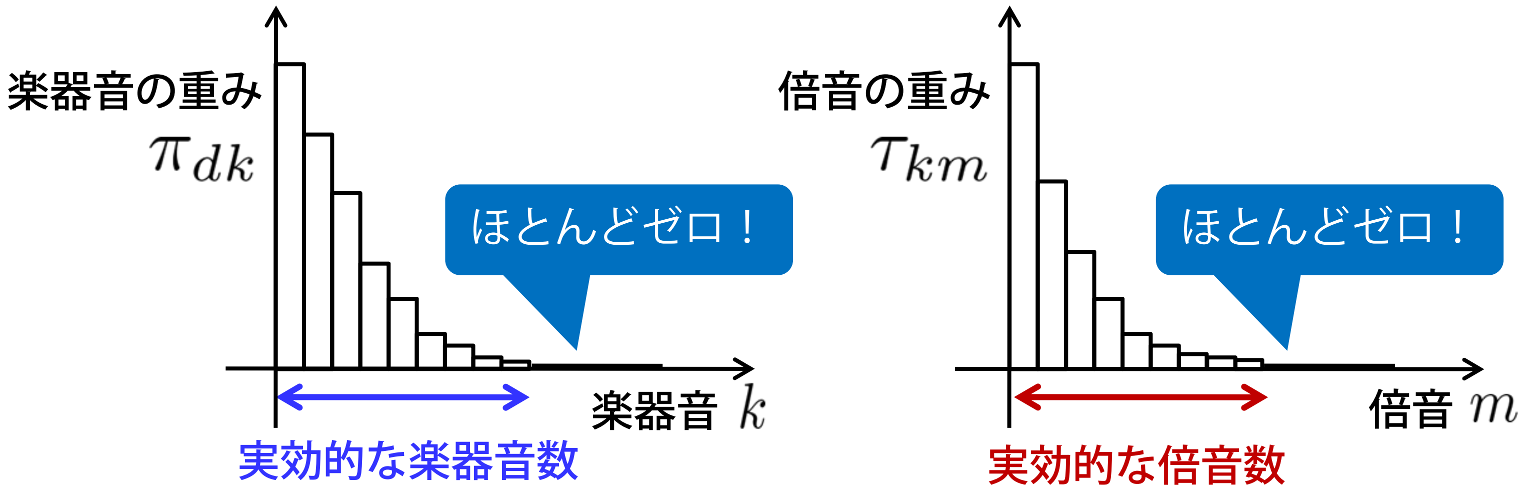
$$\mathcal{M}_d(\boldsymbol{x}) = \sum_{k=1}^{\infty} \pi_{dk} \sum_{m=1}^{\infty} \tau_{km} \mathcal{N}\left(\boldsymbol{x} \middle| \boldsymbol{\mu}_k + \boldsymbol{o}_m, \boldsymbol{\Lambda}_k^{-1}\right)$$

どうやって実効的な複雑さを決めることができるか
→ベイズモデルに基づくスパース学習

スパース学習

- (階層) ディリクレ過程を用いたノンパラベイズ推論 [Teh 2006]
  - DP&HDP：無限次元の指数減衰する重みを確率的に生成可能
    - 全ての重みを足すと1に正規化されている
    - ほとんど全ての重みがほぼゼロに近い

$$\mathcal{M}_d(\boldsymbol{x}) = \sum_{k=1}^{\infty} \pi_{dk} \sum_{m=1}^{\infty} \tau_{km} \mathcal{N}\left(\boldsymbol{x} \middle| \boldsymbol{\mu}_k + \boldsymbol{o}_m, \boldsymbol{\Lambda}_k^{-1}\right)$$



自動採譜実験結果

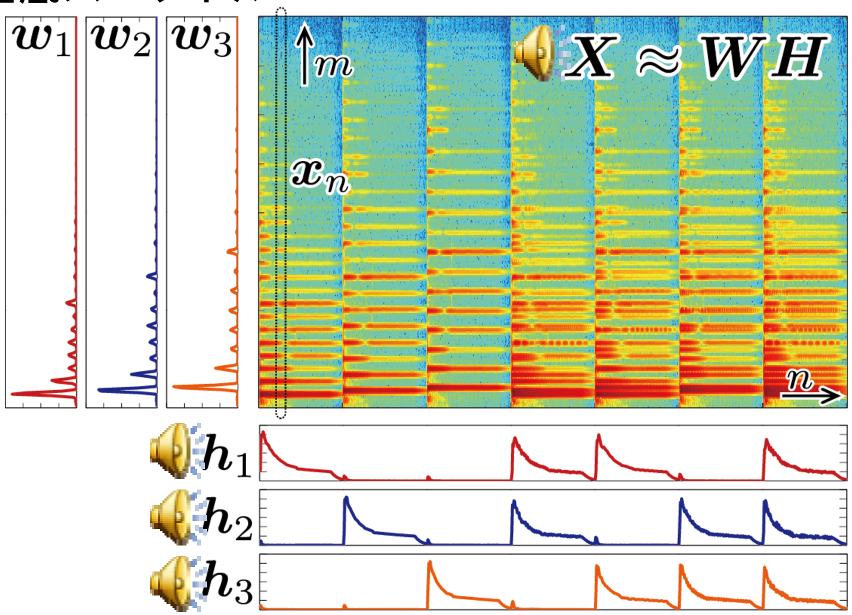
- 提案手法 (iLHA) が安定して良好な精度を達成
  - 人手で最適チューンした従来手法と同等程度の性能を達成
  - 時間方向のモデル化で改善の余地あり (HTCと同様)

	PreFEst	HTC	LHA	iLHA
Jazz No.1	75.8	79.0	70.7	82.2
Jazz No.2	78.5	78.0	69.1	77.9
Jazz No.6	70.4	78.3	49.8	71.2
Jazz No.7	83.0	86.0	70.2	85.5
Jazz No.8	85.7	84.4	55.9	84.6
Jazz No.9	85.9	89.5	68.9	84.7
Classic No.30	76.0	83.6	81.4	81.6
Classic No.35	72.8	76.0	58.9	79.6
平均精度	79.4	82.0	65.8	81.7
	[Goto 2000]	[Kameoka 2007]	有限モデル	無限モデル

因子モデルに基づくアプローチ

- 非負値行列分解 (NMF) に基づく手法が主流 [Lee 2000] [Smaragdis 2003]
  - 音源分離と非常に親和性が高い
  - スペクトログラムを非負値行列とみなして分解

基底スペクトル



音量ベクトル

観測ベクトル群

$$\boldsymbol{X} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N] \in \mathbb{R}^{M \times N}$$

基底ベクトル群

$$\boldsymbol{W} = [\boldsymbol{w}_1, \cdots, \boldsymbol{w}_K] \in \mathbb{R}^{M \times K}$$

音量ベクトル群

$$\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K]^T \in \mathbb{R}^{K \times N}$$

低ランク近似

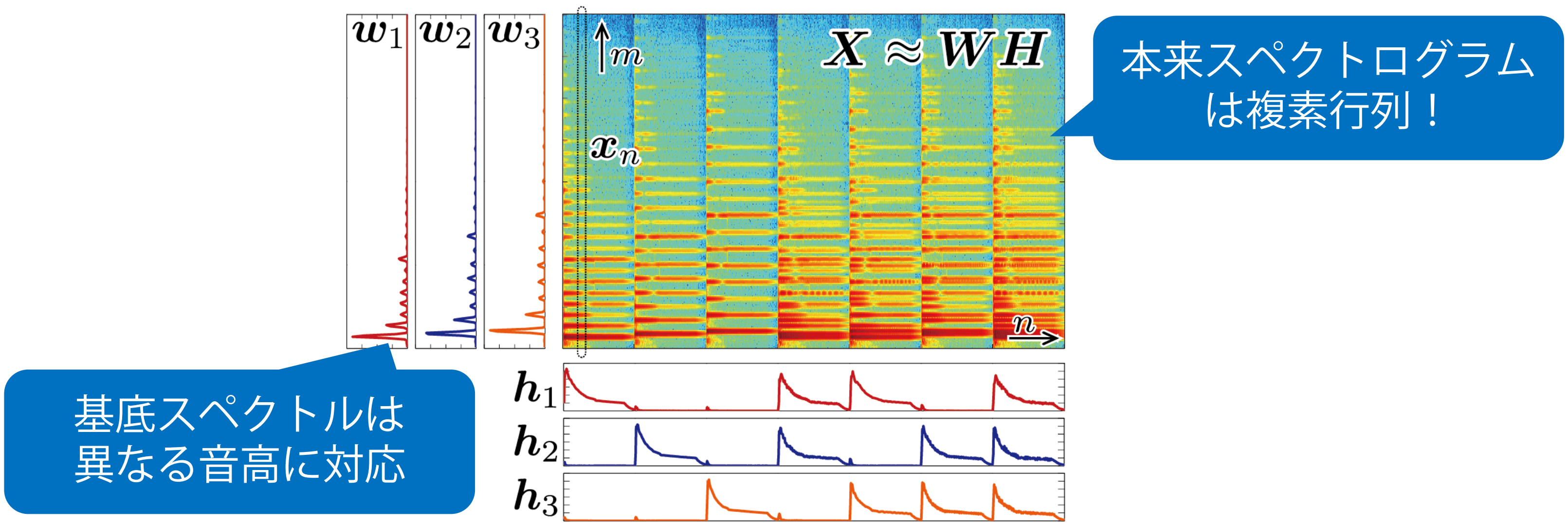
$$\boldsymbol{x}_n \approx \sum_{k=1}^{K} \boldsymbol{w}_k h_{kn} \stackrel{\mathrm{def}}{=} \boldsymbol{y}_n$$

非負ベクトルを
非負ベクトルの和で近似

音響信号処理分野で独自の目覚ましい発展！

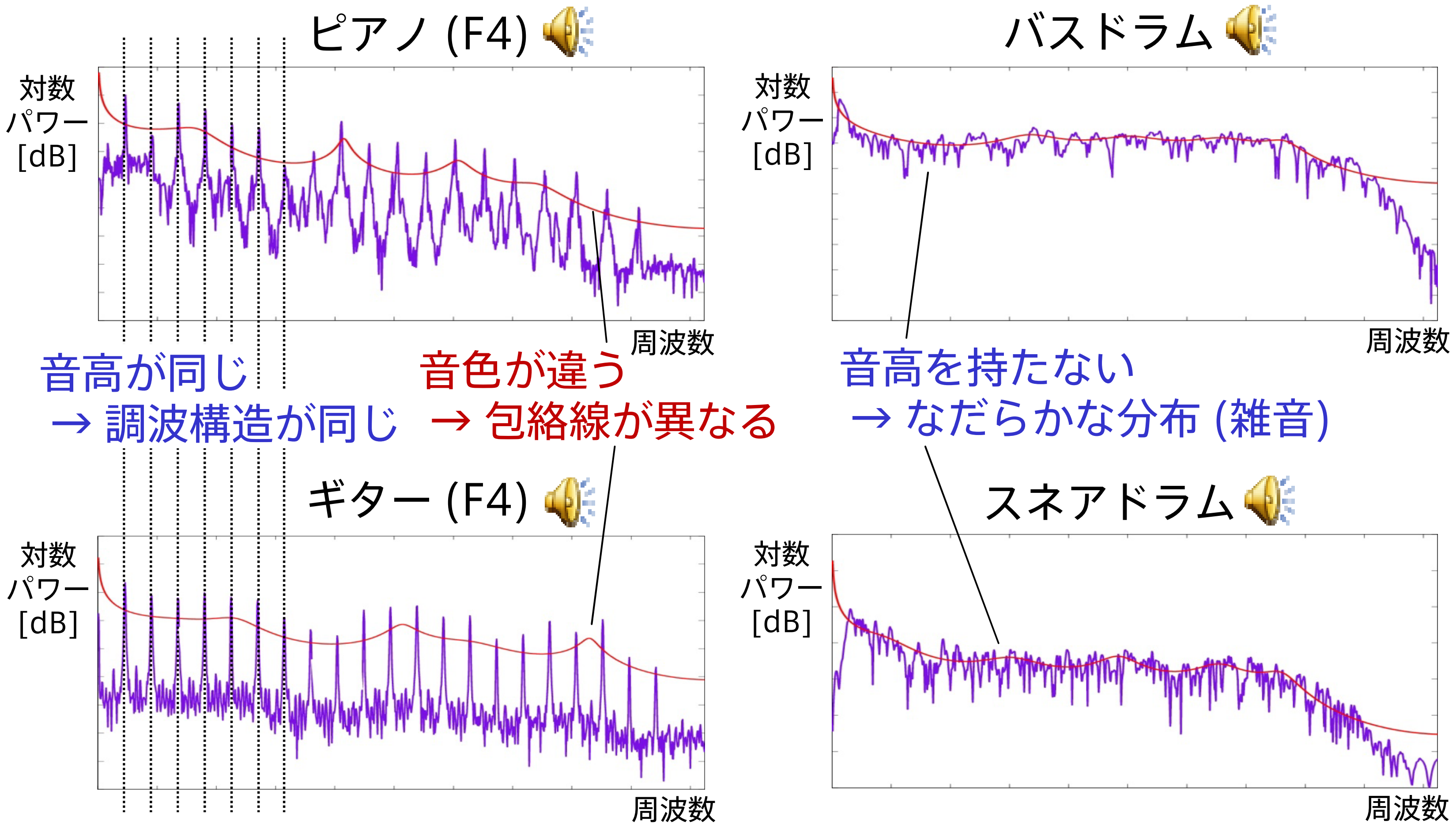
非負値行列分解の問題点

- 混合音を異なる音高ごとにしか分離できない
  - 本当は「音高」ではなく「音色」ごとに分離したい
  - 音源信号をあとからクラスタリングするのは容易ではない
- 音源分離品質に限界がある
  - 非負値行列を得るため、スペクトルの位相は捨てられる
  - 音源信号の位相復元は困難



楽器音の生成機構

- ソース・フィルタ理論による説明
  - スペクトル = 微細構造 + 包絡構造



音楽の成り立ち

- 音楽音響信号を「パーツ」に分解したい
  - 楽器音：2種類の要素「音高」と「音色」の組み合わせ

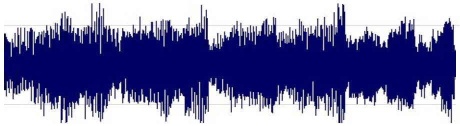
音高 (ソース)

様々な
楽器音を生成

⊗

重ね合わせ

音色 (フィルタ)

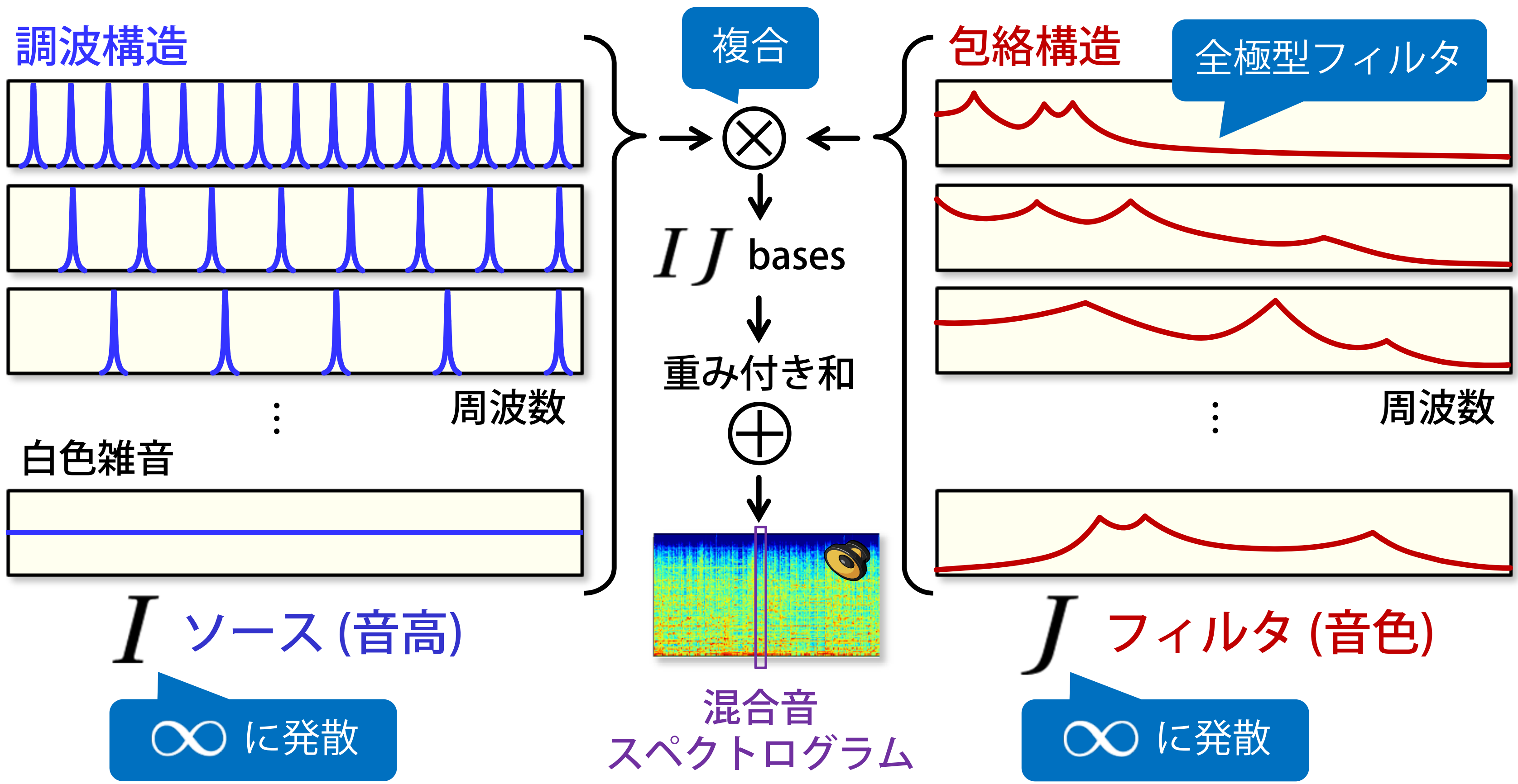


混合音

提案法：無限複合自己回帰モデル

- 無限個のソース・フィルタで構成されるNMF [Yoshii 2012]
  - 適切な個数を一挙に推定可能！
    - 音高数 I ＋音色数 J

複合自己回帰モデル [Kameoka 2009]のノンパラメトリックベイズ拡張



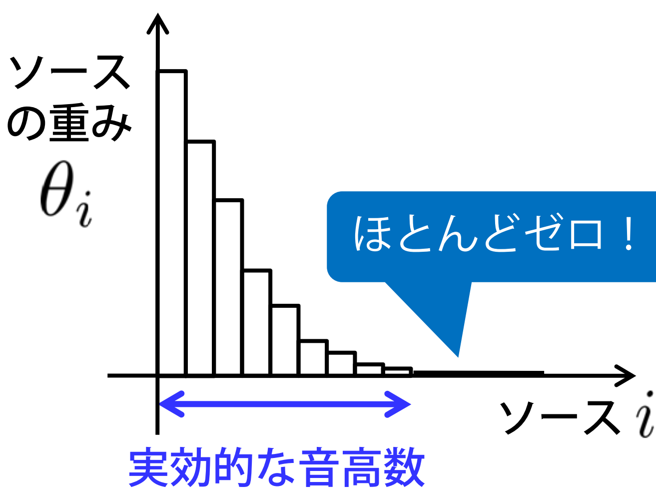
スパース学習

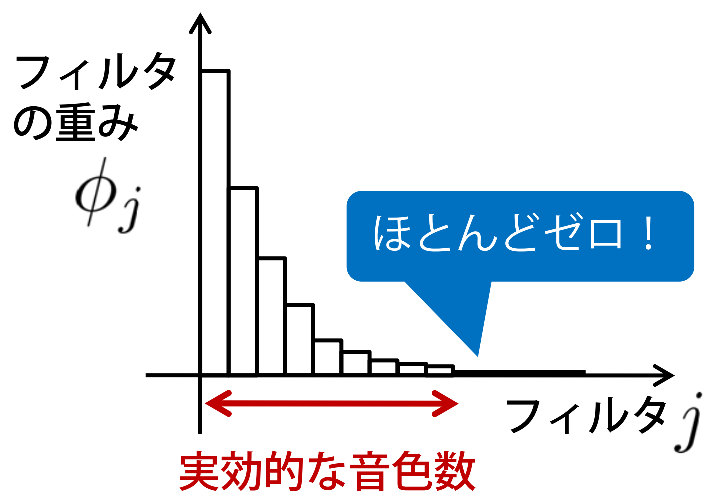
- ガンマ過程を用いたノンパラベイズ推論 [Hoffman 2010]
  - GaP：無限次元の指数減衰する非負重みを確率的に生成可能
    - ほとんど全ての重みがほぼゼロに近い

$$X_{mn} = \sum_{i}^{\infty}\sum_{j}^{\infty} \theta_i \phi_j \overbrace{\underbrace{W_{im}}_{\text{基底}} \underbrace{A_{jm}}_{\text{基底}}}^{\text{複合}} H_{ijn}$$

音量

ソース フィルタ

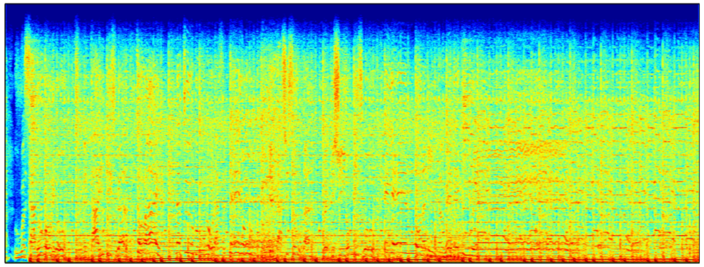




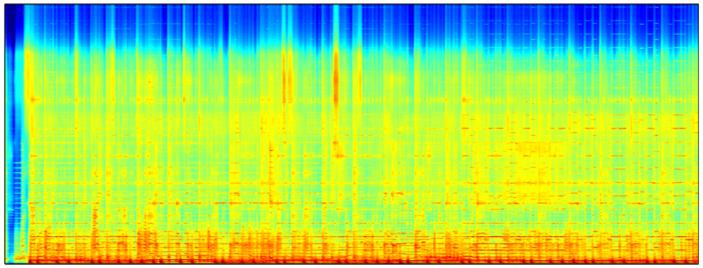
音源分離結果: RWC-MDB-P-2001 No. 1

- 混合音を音色 (フィルタ) ごとに分離することが可能

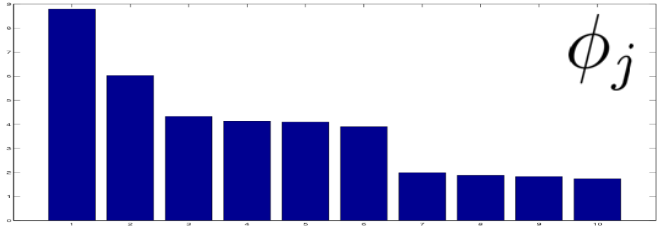
観測スペクトログラム



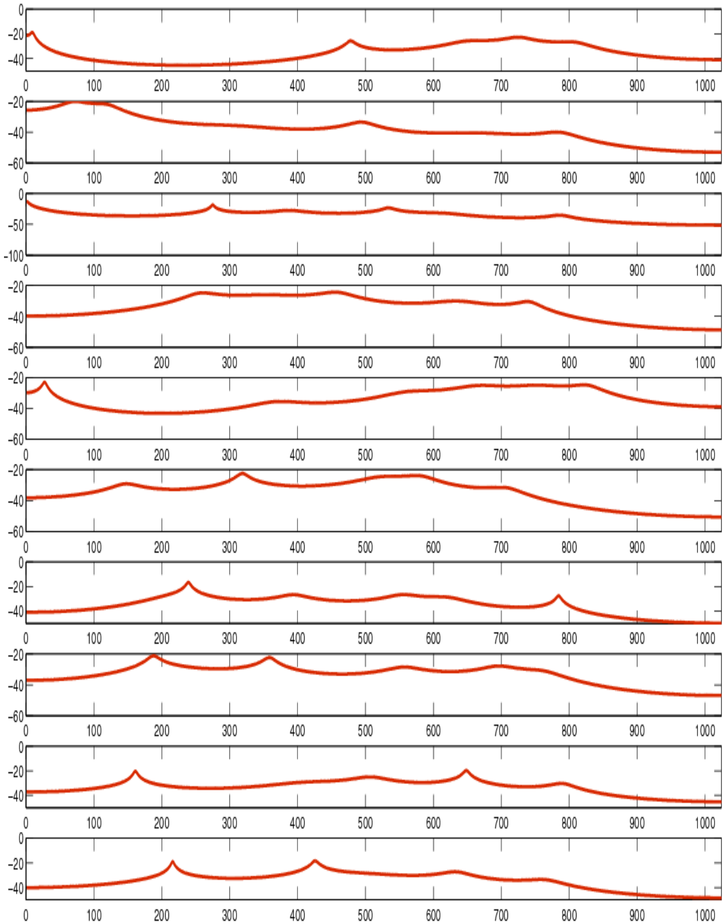
再構成スペクトログラム



フィルタの重み



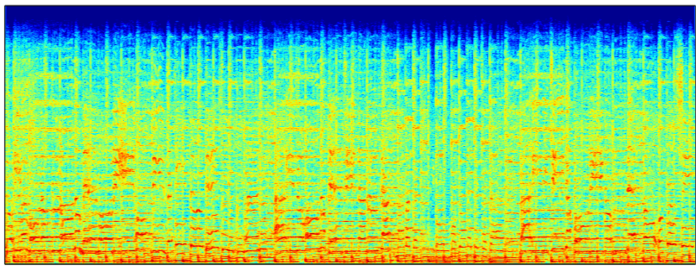
教師なし学習されたフィルタ



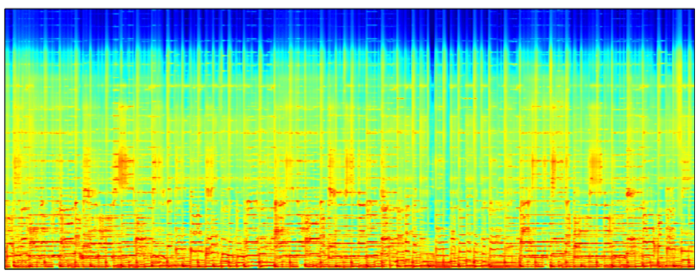
音源分離結果: RWC-MDB-P-2001 No. 5

- 混合音を音色 (フィルタ) ごとに分離することが可能

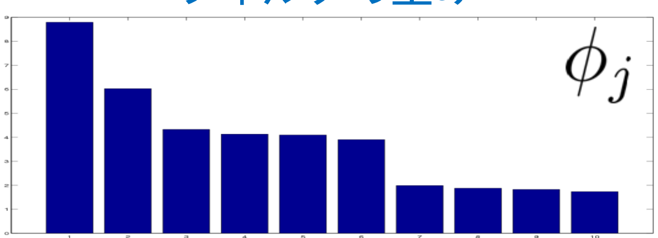
観測スペクトログラム



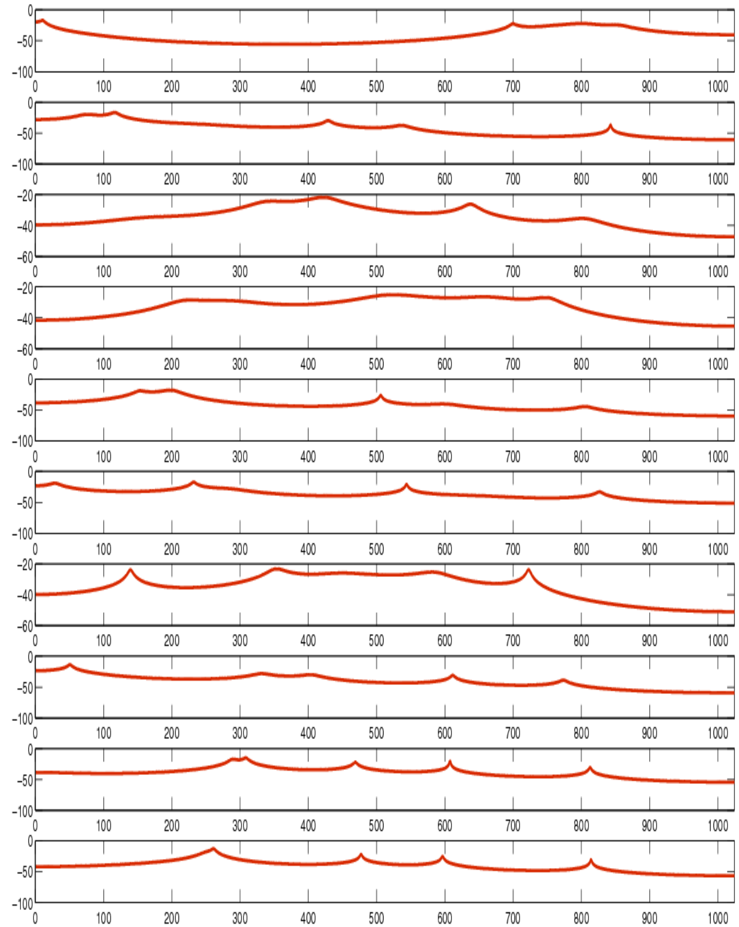
再構成スペクトログラム



フィルタの重み



教師なし学習されたフィルタ



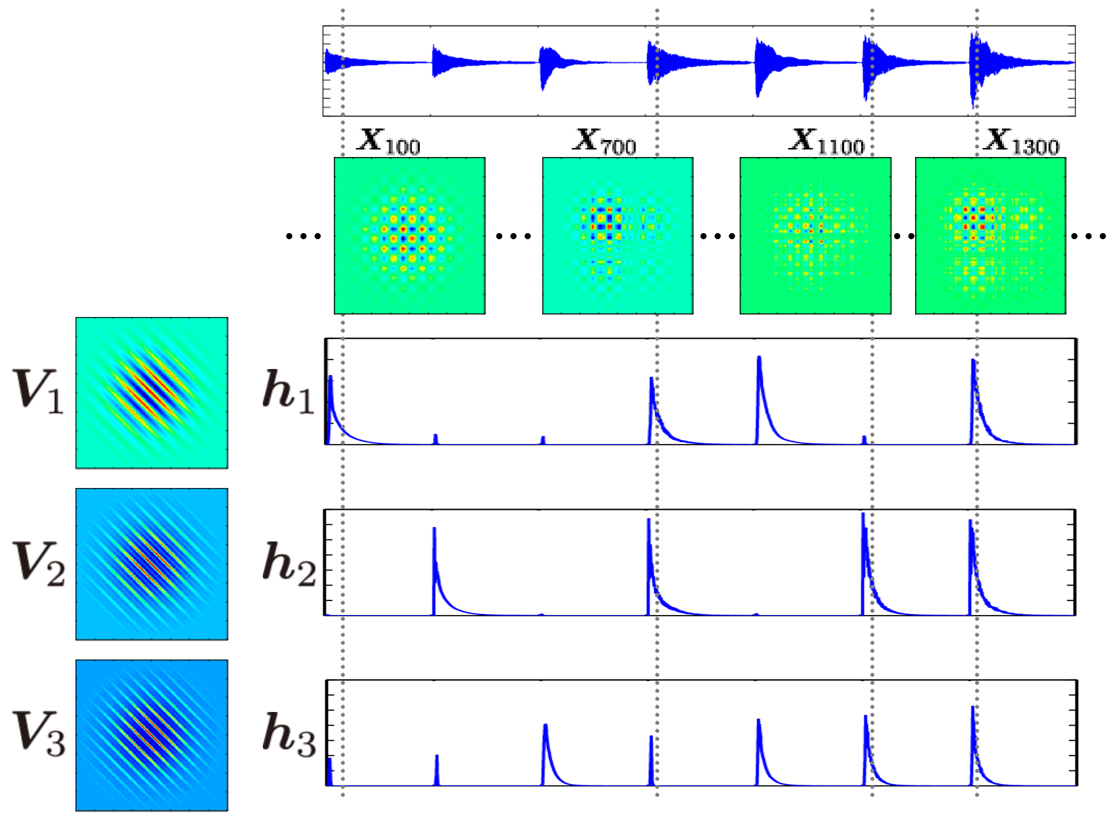
音源分離の高精度化

- 音響信号の持つ情報を全部活用したい
  - 周波数領域に変換 → 複素スペクトログラム
  - NMFを使いたいがために位相を捨てるのはもったいない
    - 複素NMF [Kameoka 2009] がひとつの解決策
- そもそも時間領域で音源分離すればよさそう
  - 時間領域での音響信号のモデル化がヒント
    - 自己回帰過程 (AR) / 線形予測分析 (LPC) [Itakura 1968]
    - 板倉-齋藤距離NMF [Fevotte 2009]
  - 入力信号としてガウス性白色雑音以外も考えたい
    - マルチカーネルLPC [亀岡 2010, Yoshii 2013]

音源信号がガウス過程に従うと仮定して
そのカーネルを直接求めることはできないか？

提案法：無限半正定値テンソル分解

- **時間領域で直接混合音を分離可能** [Yoshii 2013]
  - 周波数領域で複素スペクトログラムの分解を行うことと等価
  - 各音源信号の従うガウス過程のカーネル (基底行列) を直接推定



観測行列群

$$\boldsymbol{X} = [\boldsymbol{X}_1, \cdots, \boldsymbol{X}_N] \in \mathbb{R}^{M\times M\times N}$$

基底行列群

$$\boldsymbol{V} = [\boldsymbol{V}_1, \cdots, \boldsymbol{V}_K] \in \mathbb{R}^{M\times M\times K}$$

音量ベクトル群

$$\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K]^T \in \mathbb{R}^{K\times N}$$

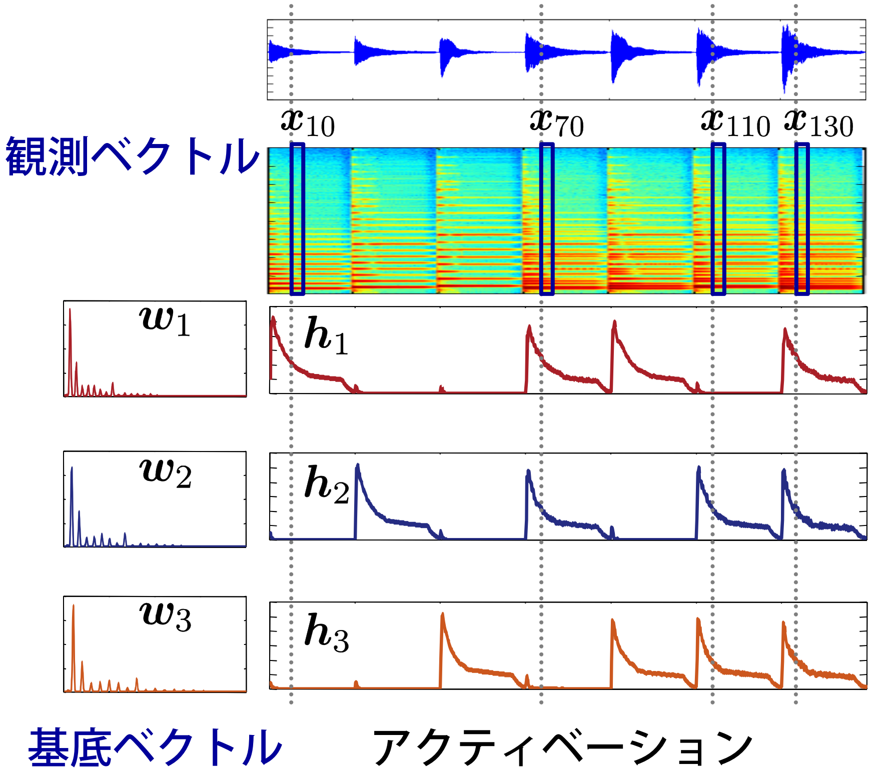
低ランク近似

$$\boldsymbol{X}_n \approx \sum_{k=1}^{K} \boldsymbol{V}_k h_{kn} \stackrel{\mathrm{def}}{=} \boldsymbol{Y}_n$$

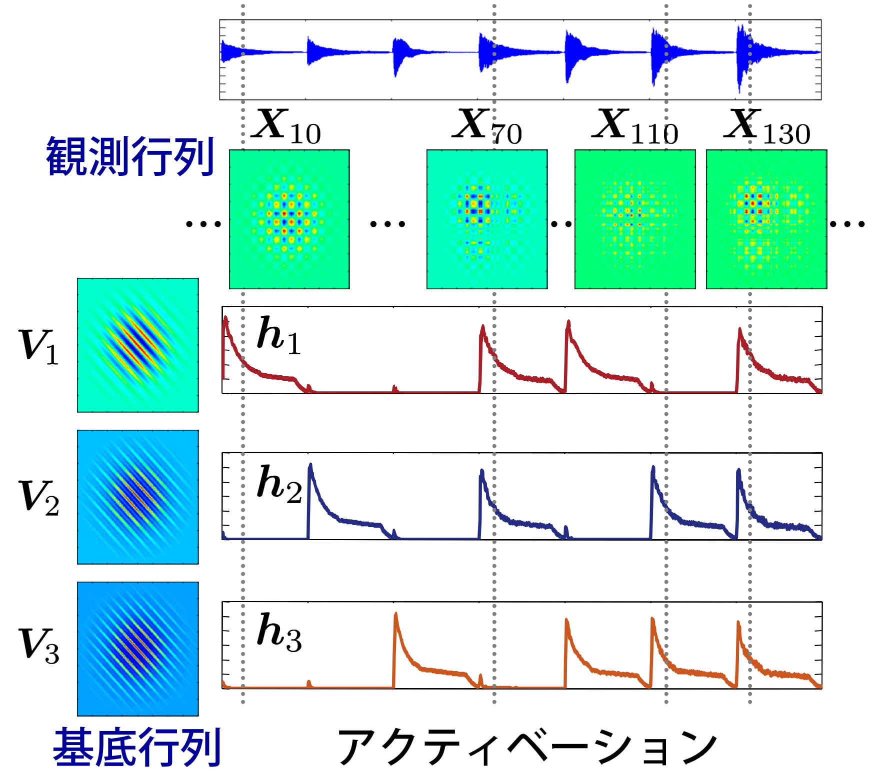
半正定値行列を
半正定値行列の和で近似

PSDTF: NMFの正統な後継者

NMF: 非負値ベクトルを
少数の非負値ベクトルの凸結合で近似



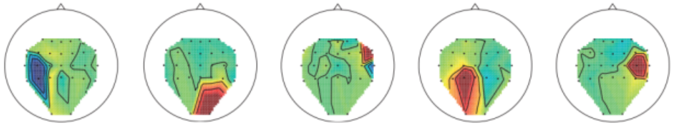
PSDTF: 半正定値行列を
少数の半正定値行列の凸結合で近似



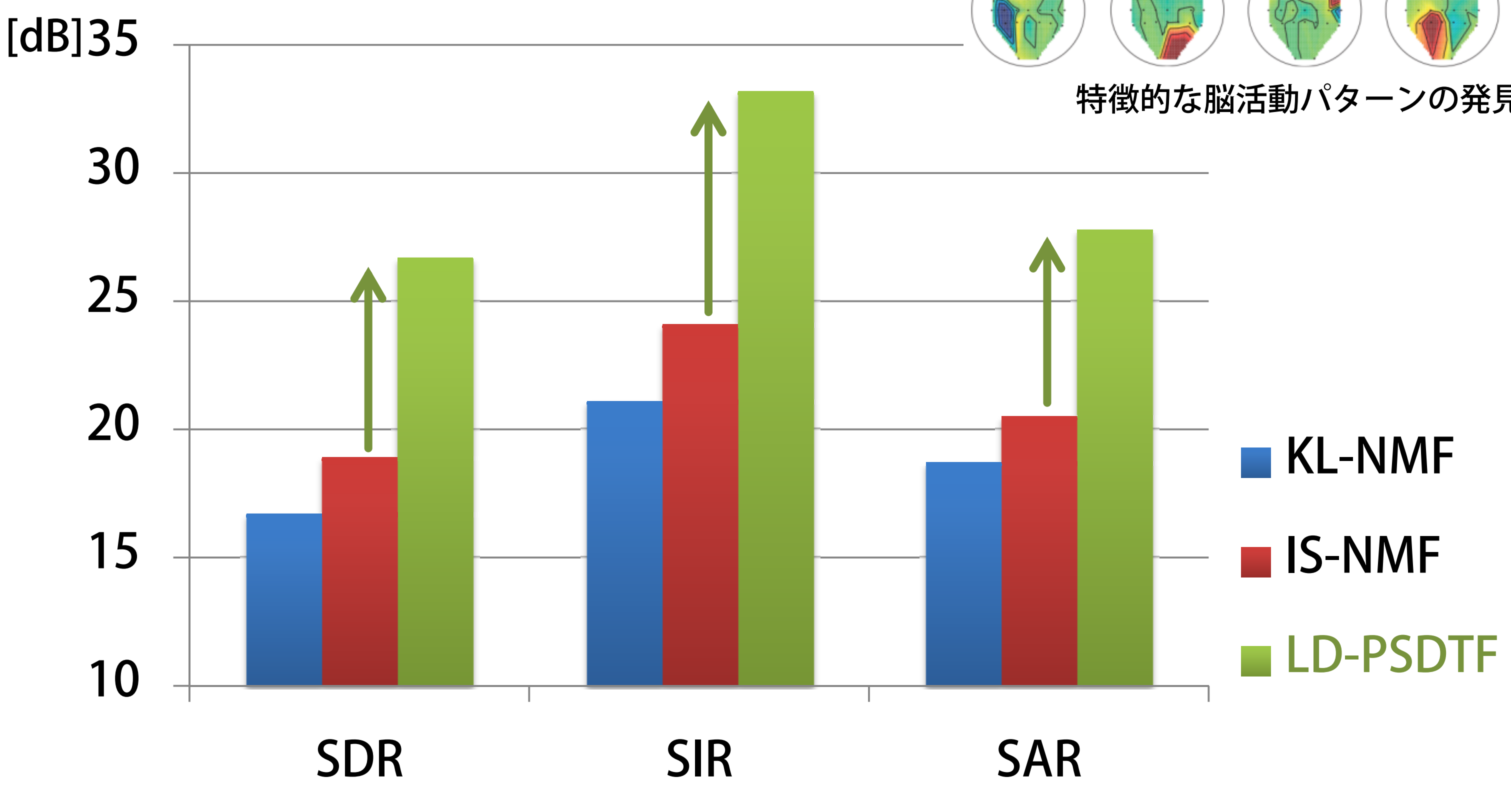
ガンマ過程に基づく基底数の無限化も可能 (ノンパラメトリックベイズ)

モノラル音響信号分離結果

- LD-PSDTF が KL-NMF や IS-NMF の性能をはるかに凌駕
  - MIDIによる3混合の簡単なピアノの和音で評価
  - NMFの極めて本質的な拡張 → 他分野への応用 (例: EEG信号解析)



特徴的な脳活動パターンの発見



これまでの研究

- ノンパラメトリックベイズ音響・言語モデルの開発
  - 音響データに対する構造学習
    - 無限潜在的調波配分法 [Yoshii@IEEE Trans. 2012 (ISMIR 2010)]
    - 無限複合自己回帰モデル [Yoshii@ISMIR 2012]
    - 無限半正定値テンソル分解 [Yoshii@ICML 2013 (ISMIR 2013)]
  - 楽譜データに対する構造学習
    - 語彙フリー無限グラムモデル [Yoshii@ISMIR 2011]

音響モデル / 言語モデル	混合モデル (例: GMM)	因子モデル (例: NMF)
連鎖モデル (例: Nグラムモデル)	本研究	本研究
木構造モデル (例: 文脈自由文法)	未提案	Nakano@ICASSP 2012 Kameoka@ISMIR 2012

Big Picture

- 階層ベイズモデルによる統一的記述
  - 音響モデルと言語モデルとの統合モデル

$p(\boldsymbol{S})$

言語モデル

語彙・文法 $\boldsymbol{S}$

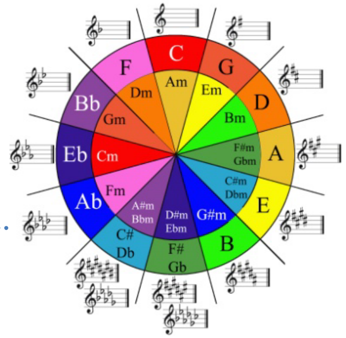
$p(\boldsymbol{Z}|\boldsymbol{S})$

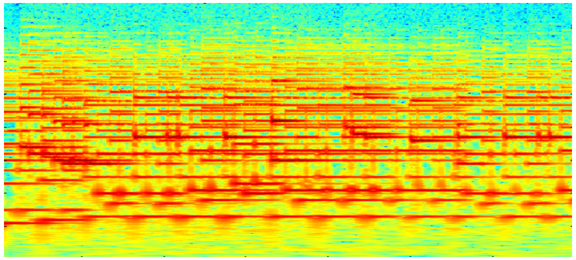
音符配置 $\boldsymbol{Z}$

$p(\boldsymbol{X}|\boldsymbol{Z})$

音響モデル

音楽音響信号 $\boldsymbol{X}$

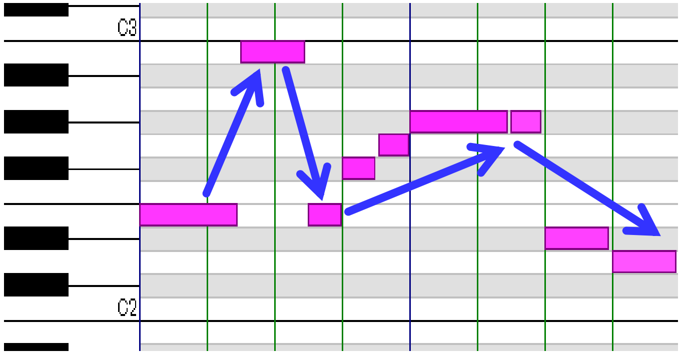


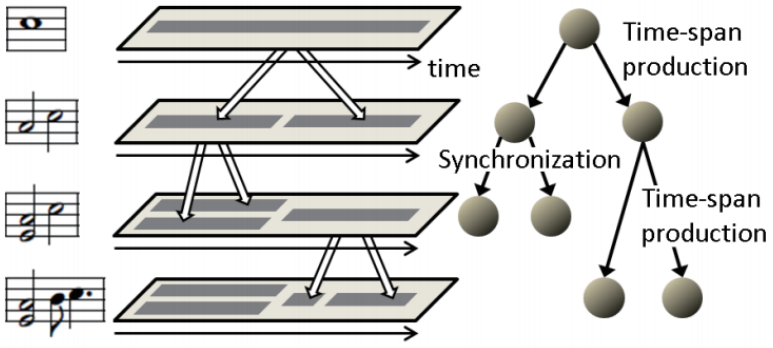


ノンパラメトリックベイズ言語モデル

- 目標：音楽文法推論 (grammar induction)
  - 自然言語処理分野における言語モデルの技術が役立つ
    - 連鎖モデル
      - Nグラムモデル (HPY拡張)
    - 木構造モデル
      - 文脈自由文法(PCFG) (HDP拡張)
    - トピックモデル
      - 潜在的ディリクレ配分法 (LDA) (HDP拡張)

音符は時間・周波数平面上に分布しているため安直な応用は不可能





連鎖モデルに基づくアプローチ

- 音符の配置に対してＮグラムモデルが使えないか？
  - コードシンボル系列対しては直接適用可能

… C:maj F:maj G:maj C:maj C:maj F:maj G:maj …

3グラム

3グラム

3グラム

  - コード名を与えなくても音符を直接扱いたい

教師なし音楽理解！

… C+E+G F+A+C G+B+D C+E+G C+E+G F+A+C G+B+D …

3グラム

4グラム

2グラム

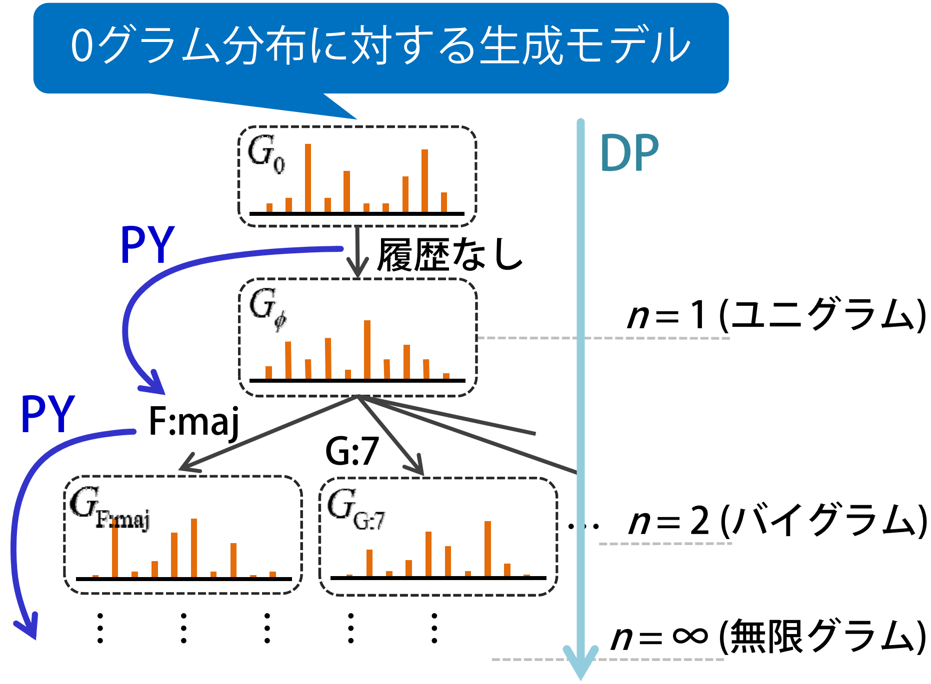
あらゆる音高の
組み合わせを考慮

あらゆる履歴長の
可能性を考慮

提案法：語彙フリー無限グラムモデル

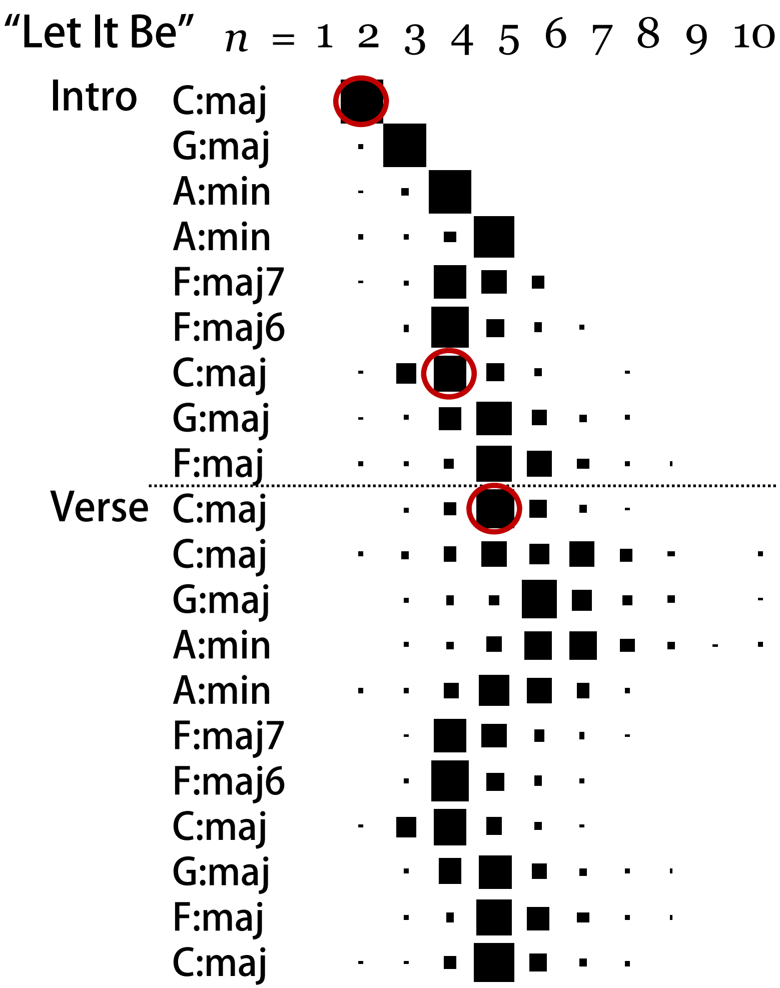
- 音高の組み合わせ系列に対するベイズモデル [Yoshii 2011]
  - 階層Pitman-Yor過程 (HPY) に基づくスムージング [Teh 2006]
  - ディリクレ過程 (DP) に基づく履歴長の無限化 [Mochihashi 2007]
  - コードの生成モデルに基づく語彙の無限化
    - 組み合わせ最適化をする必要がない！

ラベル表記	コンポーネント表記
C:maj	C:100010010000
D:min	D:100100010000
⋮	⋮
D:maj add4	D:100011010000
⋮	⋮
N/A	C:101010010000
N/A	C:100110010000



コード進行モデル学習結果

- ビートルズ180曲から特徴的なコードパターンを学習



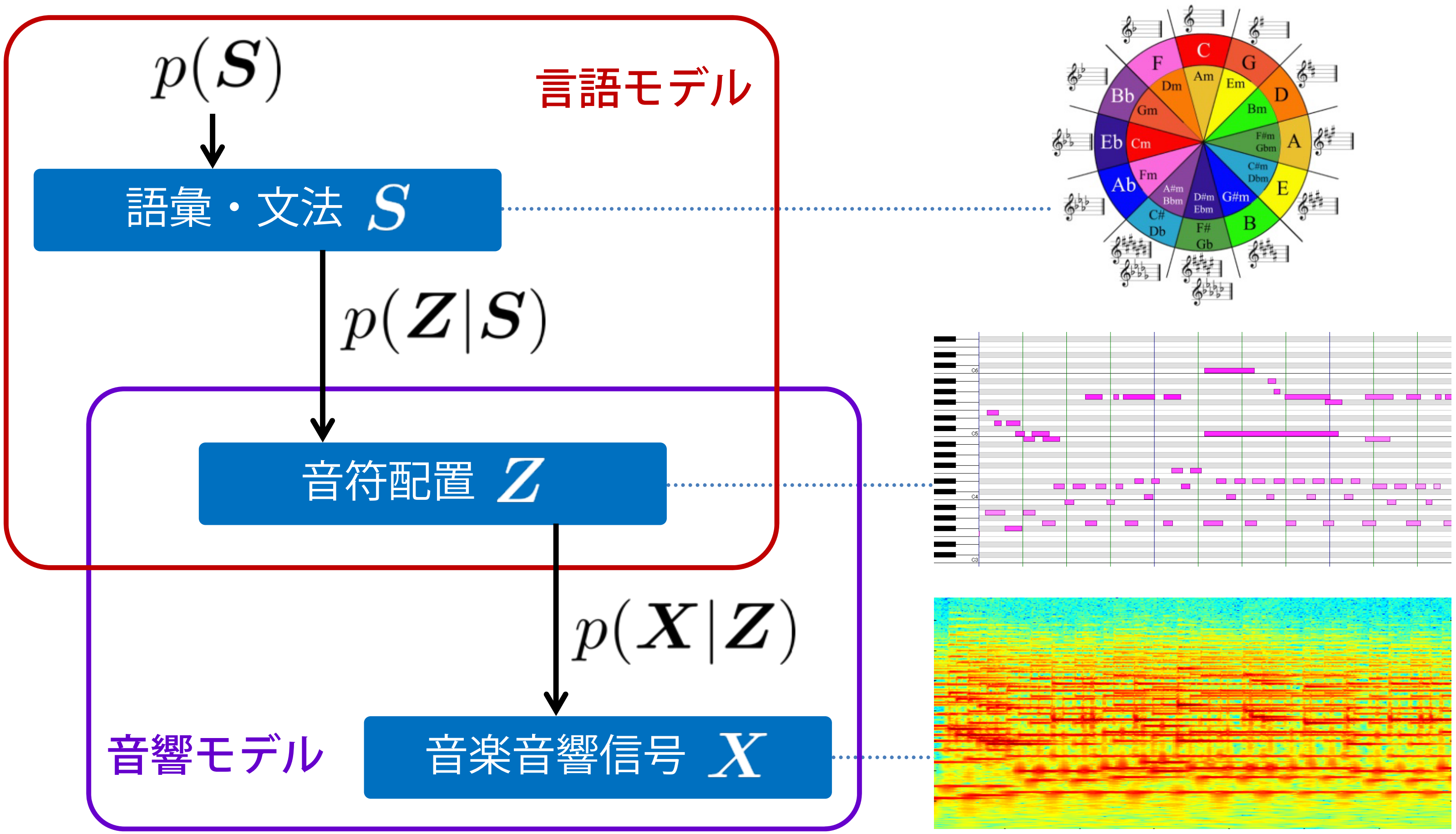
分かりやすさのためコード名で表記

確率	n	確率句 (典型的コードパターン)
0.701	3	C:7 F:7 C:7
0.682	3	B:maj F:maj G:maj
0.656	3	A:min C:7 F:maj
0.647	3	F:min G:maj C:maj
0.645	4	F:maj F:maj G:maj C:maj
0.632	3	E:min C:7 F:maj
0.630	3	C:maj7 D:min7 E:min7
0.623	4	B:maj F:maj G:maj C:maj
0.622	3	D:min7 G:sus4 G:maj
0.620	5	D:min G:maj C:maj F:maj C:maj

Big Picture

- 階層ベイズモデルによる統一的記述
  - 音響モデルと言語モデルとの統合モデル

音声認識と同様の枠組み
ただし<u>語彙や文法は未知</u>！



まとめ：教師なし音楽理解

- ノンパラメトリックベイズ音響・言語モデルの開発
  - 音響データに対する構造学習
    - 無限潜在的調波配分法 [ISMIR 2010] [IEEE Trans. 2012]
      - ノンパラメトリックベイズF0推定法
    - 無限複合自己回帰モデル [ISMIR 2012]
      - ノンパラメトリックベイズ楽器パート分離法
    - 無限半正定値テンソル分解 [ISMIR 2013] [ICML 2013]
      - ノンパラメトリックベイズ時間領域モノラル信号分解法
  - 楽譜データに対する構造学習
    - 語彙フリー無限グラムモデル [ISMIR2011]
      - 音符の組み合わせに対する確率モデル

高度な統計的機械学習技術を駆使した先駆的な研究を推進