フェージング環境下における正弦波パラメータ推定のベイズアプローチ

代表研究者　田　中　聡　久　東京農工大学 大学院工学研究院 先端電気電子部門

1 はじめに

観測したデータから，確率密度関数のパラメータを推定する問題は，信号処理や機械学習において重要である．また，確率密度関数のパラメータ推定法をスペクトルの当てはめ問題とみなすことで，正弦波の周波数推定が可能となる [5]．この密度関数のパラメータ推定には，最尤推定やベイズ推定が広く用いられている [1]．最尤推定が点推定の方法であるのに対し，ベイズ推定は，直接パラメータを推定するのではなく，パラメータの事後分布を求める手法である．最尤推定は，データ点数が少ない場合に過学習を起こしやすい [1]．これに対して，ベイズ推定は，より適切にパラメータを推定することができる．しかし，ベイズ推定で求める事後分布は，多くの場合解析的に求められない [1]．この場合に，事後分布を近似的に求める方法のひとつが変分ベイズ法である．

変分ベイズ法を上手く適用できるモデルのひとつは，工学的に応用範囲の広い混合 Gauss モデルである [1]．これに対して，離散信号のスペクトルからパラメータ推定をしたい場合，周波数は $-\pi$ から π に限られるため，円周上に分布する Gauss 分布である von Mises 分布を考慮する必要がある [3]．文献 [2] で混合 von Mises モデルに対する EM アルゴリズム，文献 [6] で変分ベイズ法が導出されている．しかし，文献 [2] の方法では，集中度の推定値を求める際に近似計算が必要となる．また，文献 [6] の方法では集中度の値が既知である必要がある．

そこで研究課題では，変分ベイズ法を用いて混合 von Mises モデルの全てのパラメータを推定する方法を構築し，混合正弦波のパラメータ推定を試みた．計算機実験により他の推定法と比較することで，得られたアルゴリズムの有効性を検証した．

2 ベイズ推定による密度関数のパラメータ推定

2.1 ベイズ推定

未知パラメータベクトル $\boldsymbol{\theta}$ をもつ密度関数 $p(\mathbf{x}|\boldsymbol{\theta})$ からなる分布によって独立に生成された，N 個のデータ集合 $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ が与えられたとする．

ベイズ推定では，パラメータの事後分布をもちいて密度関数のパラメータ $\boldsymbol{\theta}$ 求める方法である．ベイズの定理より，

$$p(\boldsymbol{\theta}|\mathbf{X}) = \frac{p(\mathbf{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{X})} = \frac{p(\mathbf{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int p(\mathbf{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}} \tag{1}$$

が成り立ち，これより $\boldsymbol{\theta}$ の事後分布を求める．パラメータの推定値 $\hat{\boldsymbol{\theta}}$ は，事後分布の期待値をとることで求められる．すなわち，

$$\hat{\boldsymbol{\theta}} = \mathbb{E}[\boldsymbol{\theta}|\mathbf{X}] = \int \boldsymbol{\theta} p(\boldsymbol{\theta}|\mathbf{X})\mathrm{d}\boldsymbol{\theta} \tag{2}$$

となる．ところが実際のモデルにおいては，事後分布を解析的に求めることが困難な場合が多い．

2.2　変分ベイズ法

変分ベイズ法 [1] は，事後分布 $p(\mathbf{Z}|\mathbf{X})$ が求められない場合に，代わりに別の密度関数 $q(\mathbf{Z})$ を求める方法である．一般に $\mathbf{Z}$ を，排反な確率変数 $\mathbf{Z}_i (i = 1, \ldots, M)$ に分割できると仮定し，密度関数のクラスを

$$q(\mathbf{Z}) = \prod_{n=1}^{M} q_i(\mathbf{Z}_i) \tag{3}$$

に制限する．変分ベイズ法の指導原理は，$p(\mathbf{Z}|\mathbf{X})$ との KL 情報量を最小にする $q(\mathbf{Z})$ を求めるというものである [1]．そのような $q(\mathbf{Z})$ は，それぞれの $q_i(\mathbf{Z}_i)$ に対して，

$$\ln q_j^{\star}(\mathbf{Z}_j) = \mathbb{E}_{i \neq j}[\ln p(\mathbf{X}, \mathbf{Z})] + \text{const} \tag{4}$$

で与えられる．このように近似して得られた事後分布は変分事後分布と呼ばれる．

3　混合 von Mises 分布

本研究課題が論ずる対象である von Mises 分布 [1, 3] は，円周上に分布する Gauss 分布であり，その密度関数は

$$\mathrm{M}(\theta|\theta_0, m) = \frac{1}{2\pi I_0(m)} \exp\{m \cos(\theta - \theta_0)\} \tag{5}$$

で与えられる．θ_0 は分布の平均を表し，m は集中度パラメータと呼ばれる，Gauss 分布の精度に相当するものである．また，$I_0(m)$ は 0 次の第 1 種変形 Bessel 関数であり，

$$I_0(m) = \frac{1}{2\pi} \int_0^{2\pi} \exp\{m \cos\theta\} \tag{6}$$

で表される．

平均が $\boldsymbol{\mu} = (\mu_1, \mu_2)$，$\mathbf{I}$ を 2×2 の単位行列として，共分散が $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$ である 2 変数 $\mathbf{x} = (x_1, x_2)$ 上の Gauss 分布

$$p(x_1, x_2) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2}{2\sigma^2}\right\} \tag{7}$$

を考える．ここで，

$$x_1 = r\cos\theta, \quad x_2 = r\sin\theta, \quad \mu_1 = r_0\cos\theta_0, \quad \mu_2 = r_0\sin\theta_0, \quad m = \frac{r_0}{\sigma^2}, \quad r = 1 \tag{8}$$

のように変換することで，式 (5) で表される von Mises 分布が得られる．

Gauss 分布に対して混合 Gauss モデルが定義できるように，von Mises 分布に対しても混合モデルを定義できる．混合 Gauss モデルに関しては，EM アルゴリズムを始めするパラメータ推定手法がある程度確立している．しかしながら，von Mises 分布に関しては，パラメータ推定が困難であった．以下詳しく述べる．

混合 von Mises モデルは，密度関数が

$$p(x) = \sum_{k=1}^{K} \pi_k \mathrm{M}(x_n|\theta_k, \tau_k) \tag{9}$$

で表される．この混合 von Mises モデルのパラメータ推定法として，EM 法 [2] と変分ベイズ法 [6] がある．この方法では，集中度の推定値を求める際に近似計算をする必要がある．また，変分ベイズ法の導出も試みられている [6] が，集中度を推定できない．以下では，文献 [6] にしたがって，この問題について詳しく見てゆく．

4　2 変数 Gauss 分布を介した変分ベイズ法の導出

本節では，変分ベイズ法を用いて，全てのパラメータを推定する方法を構築する．基本的な考え方は，von Mises 分布を等方性の Gauss 分布に変換し，混合 Gauss 分布の変分ベイズ法を導く点にある．これは，von Mises 分布が 2 変数 Gauss 分布から導かれることを用いている．ただし，von Mises から Gauss 分布への変換は一意に定まらないため，以下で述べるように導出には多少の工夫が必要になる．

混合 von Mises モデルの密度関数は式 (9) で表される混合 von Mises 分布で与えられるとする．ここで，von Mises 分布を等方性の 2 変数 Gauss 分布に変換する:

$$x_{n1} = r\cos x_n, \quad x_{n2} = r\sin x_n, \quad \mu_{k1} = r_0\cos\theta_k, \quad \mu_{k2} = r_0\sin\theta_k, \quad \kappa_k = r_0\tau_k, \quad r = 1. \tag{10}$$

さらに，$\mathbf{x}_n = (x_{n1}, x_{n2})^{\mathrm{T}}, \boldsymbol{\mu}_k = (\mu_{k1}, \mu_{k2})^{\mathrm{T}}$ を定義すると，等方性をもつ 2 変数 Gauss 分布は，

$$p(\mathbf{X}|\mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\kappa}) = \prod_{n=1}^{N}\prod_{k=1}^{K} \mathcal{N}\left(\mathbf{x}_n|\boldsymbol{\mu}_k, (\kappa_k\mathbf{I})^{-1}\right)^{z_{nk}} \tag{11}$$

と表現できる．ここで，$\mathbf{I}$ は 2×2 の単位行列，$\boldsymbol{\mu} = \{\boldsymbol{\mu}_k\}, \boldsymbol{\kappa} = \{\kappa_k\}$ である．これを用いて混合 Gauss モデルに対する変分ベイズ法 [1] と同様に，各パラメータの変分事後分布を求めることが可能になる．

パラメータの事前分布は，$p(\boldsymbol{\pi})$ に関しては Dirichlet 分布，また

$$p(\boldsymbol{\mu}, \boldsymbol{\kappa}) = p(\boldsymbol{\mu}|\boldsymbol{\kappa})p(\boldsymbol{\kappa}) = \prod_{k=1}^{K} \mathcal{N}\left(\boldsymbol{\mu}_k|\mathbf{m}_0, (\beta_0\kappa_k\mathbf{I})^{-1}\right)\mathcal{W}(\kappa_k\mathbf{I}|\mathbf{W}_0, \nu_0) \tag{12}$$

とする．ここで，$\boldsymbol{\kappa}$ の事前分布として導入した Wishart 分布は，

$$p(\boldsymbol{\kappa}) = \mathcal{W}(\kappa_k\mathbf{I}|\mathbf{W}_0, \nu_0) = B(\mathbf{W}_0, \nu_0)\kappa_k^{\nu_0-3}\exp\left\{-\frac{1}{2}\mathrm{Tr}\left(\kappa_k\mathbf{W}_0^{-1}\right)\right\} \tag{13}$$

となる．ここで，Gauss 分布は 2 変数なので，$D = 2$ である．

式 (13) において，$\mathrm{Tr}(\kappa_k\mathbf{W}_0^{-1})$ の値が等しいければ同一の分布となり，そのような $\mathbf{W}_0$ の値は無数に存在する．しかしながら，$\mathbf{W}_0$ は $\kappa_k\mathbf{I}$ のハイパーパラメータなので，単位行列の定数倍で表現されなくてはならない．この条件より，$\mathbf{W}_0 = \dfrac{1}{2b_0}\mathbf{I}, \nu_0 = 2a_0 + 1$ に制限することができ，式 (13) は，

$$p(\boldsymbol{\kappa}) = B\left(\frac{1}{2b_0}\mathbf{I}, 2a_0 + 1\right)\kappa_k^{2a_0-2}\exp\left\{-\frac{1}{2}\mathrm{Tr}\left(2b_0\kappa_k\mathbf{I}\right)\right\} = \mathrm{Gam}(\kappa_k|a_0, b_0)^2 \tag{14}$$

となり，Wishart 分布を 2 つの Gamma 分布の積として表すことができる．以上より，$\boldsymbol{\mu}$, $\boldsymbol{\kappa}$ の事前分布は

$$p(\boldsymbol{\mu}, \boldsymbol{\kappa}) = p(\boldsymbol{\mu}|\boldsymbol{\kappa})p(\boldsymbol{\kappa}) = \prod_{k=1}^{K} \mathcal{N}\left(\boldsymbol{\mu}_k|\mathbf{m}_0, (\beta_0\kappa_k\mathbf{I})^{-1}\right)\mathrm{Gam}(\kappa_k|a_0, b_0)^2 \tag{15}$$

となる．

ここで，変分事後分布 q を

$$q(\mathbf{Z}, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\kappa}) = q(\mathbf{Z})q(\boldsymbol{\pi})q(\boldsymbol{\mu}, \boldsymbol{\kappa}) \tag{16}$$

によって表現すると，式 (4) より，真の事後分布と分布 q の KL 情報情報量が最小となる $q^\star$ はそれぞれ，

$$\ln q^\star(\mathbf{Z}) = \mathbb{E}_{\boldsymbol{\pi},\boldsymbol{\mu},\boldsymbol{\kappa}}[\ln p(\mathbf{X},\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\boldsymbol{\kappa})] + \text{const}, \tag{17}$$

$$\ln q^\star(\boldsymbol{\pi}) = \mathbb{E}_{\mathbf{Z},\boldsymbol{\mu},\boldsymbol{\kappa}}[\ln p(\mathbf{X},\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\boldsymbol{\kappa})] + \text{const}, \tag{18}$$

$$\ln q^\star(\boldsymbol{\mu},\boldsymbol{\kappa}) = \mathbb{E}_{\mathbf{Z},\boldsymbol{\pi}}[\ln p(\mathbf{X},\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\boldsymbol{\kappa})] + \text{const} \tag{19}$$

によって求められる．同時密度関数は，

$$p(\mathbf{X},\mathbf{Z},\boldsymbol{\pi},\boldsymbol{\mu},\boldsymbol{\kappa}) = p(\mathbf{X}|\mathbf{Z},\boldsymbol{\mu},\boldsymbol{\kappa})p(\mathbf{Z}|\boldsymbol{\pi})p(\boldsymbol{\pi})p(\boldsymbol{\mu}|\boldsymbol{\kappa})p(\boldsymbol{\kappa}) \tag{20}$$

と表現できる．

5　計算機実験

フェージングを受けた混合正弦波信号の周波数推定問題への応用例を，数値実験とともに示す．実験では文献 [2] で導出された最尤推定法と，本研究課題で導出した変分ベイズ推定の 2 種類を用い，その結果を比較する．

フェージングは，携帯通信における受信局の移動や，レーダを用いた観測における対象物体の移動により起きる [4]．これらはドップラースプレッドと呼ばれ，フェージングの影響を受けた信号は複数の周波数帯域に広がったスペクトルをもつ [4]．

ここでは，K 個の正弦波が重畳し，フェージングの影響を受け，ノイズが加わった信号のモデル

$$x(n) = \sum_{k=1}^{K} a_k(n)\exp(j\omega_k n) + v(n) \tag{21}$$

を仮定する [4]．ここで，ω_k は正規化周波数，$v(n)$ は 0 平均で分散が σ_v^2 であるガウシアンノイズを表す．また，$a_k(n)$ はフェージングによる乗法性ノイズを表す．$a_k(n)$ は Gauss 過程であり，その自己相関関数は

$$R_a(\tau) = \sigma_a^2 \exp\left(-\frac{\omega_d^2 n^2}{4}\right) \tag{22}$$

と表されるとする．ここで ω_d は正規化ドップラー周波数を表す．

この仮定のもとでは，フェージングを受けた単一正弦波のパワースペクトル密度は Gauss 関数となる．そこでここでは，受信信号のスペクトルを混合 von Mises モデルにあてはめる．ここで推定対象である混合信号は密度関数ではないが，その周波数スペクトルを観測データのヒストグラムとみなし，対数尤度で表現することで，スペクトル推定問題は密度関数推定問題と等価となる [5]．

スペクトル推定問題を密度関数推定問題として扱うために，まず，観測データ $\mathbf{X}$ を N 点離散フーリエ変換し離散周波数スペクトル $|y(x_n)|$ を得る．ここで x_n は離散周波数である．このスペクトルはサンプル x_n が $y(x_n)$ 回観測されたデータと考えることができ，ヒストグラム $(x_n, y(x_n))$ とみなすことができる．

本実験では，3 つのモデルを作成した．それぞれのモデルの周波数を表 1 に示す．これらのデータに対して，最尤推定とベイズ推定の 2 種類の方法により周波数を推定し，比較する．最尤推定の際の初期値は，$\theta_1 = -0.5, \theta_2 = 0.5, m_1 = m_2 = 1, \pi_1 = \pi_2 = 0.5$ とし，ベイズ推定においては，ハイパパラメータは無情報事前分布となるように $\alpha_0 = \beta_0 = \theta_0 = a_0 = b_0 = 0$，変分事後分布のパラメータの初期値は $\theta_1 = -0.5, \theta_2 = 0.5, \alpha_1 = \alpha_2 = \beta_1 = \beta_2 = a_1 = a_2 = b_1 = b_2 = 1$ とした．

表 1: 周波数推定実験における，作成したデータと実験結果 (括弧内は標準偏差)

モデル	真値		最尤推定		ベイズ推定	
	ω_1	ω_2	θ_1^{EM}	θ_2^{EM}	θ_1^{VB}	θ_2^{VB}
A	-0.5π	0.5π	-0.50π	0.50π	-0.50π	0.50π
			(±0.00)	(±0.00)	(±0.00)	(±0.00)
B	-0.3π	0.3π	-0.31π	0.31π	-0.31π	0.31π
			(±0.00)	(±0.00)	(±0.00)	(±0.00)
C	-0.2π	0.2π	-0.04π	0.04π	$-\mathbf{0.22}\pi$	$\mathbf{0.22}\pi$
			(±0.01)	(±0.01)	(±0.00)	(±0.01)

実験結果を表 1 に示す．表 1 では，より真値に近い推定値を太字で示してある．モデル A，モデル B については推定精度に差がないが，モデル C についてはベイズ推定がより高い推定精度を示した．

6 結論

本研究課題では，変分ベイズ法を用いて，混合 von Mises モデルの全てのパラメータを推定する方法を構築し，正弦波推定問題に応用した．計算機実験の結果，導出した推定アルゴリズムは最尤推定よりも高い推定精度を示した．今後は，混合数も未知パラメータとして推定できるようにする必要がある．また，von Mises 分布の多次元版である von Mises-Fisher 分布に拡張することも課題である．

参考文献

[1] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[2] I. S. Dhillon and S. Sra, “Modeling data using directional distributions,” Technical Report TR-03-06, Department of Computer Sciences, The University of Texas, Feb. 2003.

[3] K. V. Mardia and P. E. Jupp, *Directional Statistics*. Wiley, 2000.

[4] M. Ghogho, A. Swami, and T. Durrani, “Frequency estimation in the presence of Doppler spread: Performance analysis,” *IEEE Trans. Signal Processing*, vol. 49, no. 4, pp. 777–789, 2001.

[5] 根智志, 田中聡久, “変分ベイズ法を用いた混合フェージング信号の周波数推定,” 2012 年電子情報通信学会総合大会論文集, A-4-33, 岡山大学, 2012.

[6] 田辺昭博, 福永健次, 大羽成征, 石井信, “混合 von Mises-Fisher 分布の変分ベイズ推定について,” 信学技報, vol. 14, no. 758, pp. 131–136, Mar. 2005.

〈発　表　資　料〉

題　　名	掲載誌・学会名等	発表年月
方向統計学のための変分ベイズ法	電子情報通信学会技術研究報告 SIP2013-51， pp. 113—118	2013 年 7 月