



HAL
open science

Diffuse noise robust multiple source localization based on matrix completion via trace norm minimization

Nobutaka Ito, Emmanuel Vincent, Nobutaka Ono, Rémi Gribonval, Shigeki Sagayama

► **To cite this version:**

Nobutaka Ito, Emmanuel Vincent, Nobutaka Ono, Rémi Gribonval, Shigeki Sagayama. Diffuse noise robust multiple source localization based on matrix completion via trace norm minimization. ASJ Spring Meeting, Mar 2011, Tokyo, Japan. inria-00596138

HAL Id: inria-00596138

<https://inria.hal.science/inria-00596138v1>

Submitted on 30 Sep 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

トレースノルム最小化による行列補完に基づく 拡散性雑音に頑健な複数音源定位*

伊藤 信貴 (東大院情報理工) エマニュエル・ヴァンソン (INRIA)

小野 順貴 (東大院情報理工) レミ・グリボンヴァル (INRIA) 嵯峨山 茂樹 (東大院情報理工)

1 はじめに

我々は、様々な方向から到来する拡散性雑音に頑健な複数音源定位法を研究している。拡散性雑音の共分散行列の部分空間モデルと、方向性信号の共分散行列の低ランク性を用いて、雑音の寄与を含む観測共分散行列から、方向性信号のノイズフリーな共分散行列を推定し、これに対して MUSIC (Multiple Signal Classification) を適用する。論文 [2] では方向性信号の共分散行列のランクを与えて、最尤法により推定を行う方法を提案したが、本稿ではトレースノルムの最小化に基づく方法を提案する。

2 背景

2.1 問題設定：拡散性雑音の存在下の複数音源定位

信号を STFT 領域における零平均の確率変数で表し、 τ, ω でフレーム番号、角周波数を表す。 $\alpha(\tau, \omega)$ のパワースペクトログラム (分散) を $\phi_{\alpha\alpha}(\tau, \omega)$ で表し、 $\alpha(\tau, \omega)$ の共分散行列を $\Phi_{\alpha\alpha}(\tau, \omega)$ で表す。

M 素子のアレイが拡散性雑音の存在下で、未知の方位角から水平に平面波として到来する L ($L < M$) 個の信号を受信するとする。 $s_l(\tau, \omega)$ を l 番目の信号、 $x(\tau, \omega) \in \mathbb{C}^M$ を観測信号、 $v(\tau, \omega) \in \mathbb{C}^M$ を拡散性雑音とする。方位角 θ に対応するステアリングベクトルを $d(\omega; \theta)$ とし、 l 番目の音源の方位角を θ_l とすると、観測信号は次のようにモデル化できる。

$$x(\tau, \omega) \triangleq \sum_{l=1}^L s_l(\tau, \omega) d(\omega; \theta_l) + v(\tau, \omega) \quad (1)$$

音源と雑音、および音源同士が無相関であるとする、共分散行列について次の関係式が成立する。

$$\Phi_{xx}(\tau, \omega) = \sum_{l=1}^L \phi_{s_l s_l}(\tau, \omega) d_l(\omega) d_l^H(\omega) + \Phi_{vv}(\tau, \omega) \quad (2)$$

ここで、 $d_l(\omega) \triangleq d(\omega; \theta_l)$ である。

本稿の問題設定は、 $x(\tau, \omega)$ が与えられたときに、 $\{\theta_l | l = 1, \dots, L\}$ を推定することである。

2.2 MUSIC とその課題

複数音源定位の基本的な手法である MUSIC は、MUSIC スペクトルと呼ばれる θ の関数

$$f(\omega; \theta) \triangleq \frac{1}{\|V^H(\omega) d(\omega; \theta)\|_2^2} \quad (3)$$

が $\theta = \theta_1, \theta_2, \dots, \theta_L$ でピーク値をとるという性質に基づく。ここで、 $V(\omega) \in \mathbb{C}^{M \times (M-L)}$ は雑音部分空間

$$\mathcal{N}(\omega) \triangleq \{\alpha \in \mathbb{C}^M | \alpha^H d_l(\omega) = 0, \forall l\} \quad (4)$$

の正規直交基底を列とする行列である。

$f(\omega; \theta)$ の計算においては、 $\mathcal{N}(\omega)$ の正確な同定が重要であるが、拡散性雑音の存在下でこれを正確に同定することが課題であった。 $\mathcal{N}(\omega)$ は方向性信号の共分散行列 [式 (2) の右辺第一項] の零空間に一致する

から、共分散行列の固有空間に影響しない空間的に白色な雑音の場合には、観測された共分散行列の最小固有値に属する固有空間として $\mathcal{N}(\omega)$ を求めることができる。しかし、共分散行列の固有空間を変える拡散性雑音の場合には、観測された共分散行列から直接 $\mathcal{N}(\omega)$ を正確に同定することは困難である。

3 共分散行列の雑音除去のための提案法

3.1 概要

提案法では、拡散性雑音に対して頑健に複数音源を定位するために、観測された共分散行列から方向性信号の共分散行列を復元し、その零空間として同定された $\mathcal{N}(\omega)$ に対して MUSIC を適用する。拡散性雑音に対しては、空間的な無相関性 [1] など、何らかのモデルが仮定できる場合も多い。本研究では、拡散性雑音に対する従来のモデルの多くが、行列空間内のある部分空間 $\mathcal{V}(\omega)$ とみなせることに着目し、観測された共分散行列を直交補空間 $\mathcal{V}^\perp(\omega)$ に射影することにより、方向性信号のノイズフリーな $\mathcal{V}^\perp(\omega)$ 成分を得る。一方、方向性信号の共分散行列の $\mathcal{V}(\omega)$ 成分はこのような方法では得られないので、方向性信号の共分散行列が理想的には低ランクである (ランクが音源数に一致) ことに着目し、低ランク行列の要素 (もしくはそれらの線形結合) の一部分から行列全体を復元する行列補完と呼ばれる数理的技術を用いて、方向性信号の共分散行列を推定する。

3.2 拡散性雑音の共分散行列の部分空間モデル [2]

本稿では、拡散性雑音の共分散行列が、 $M \times M$ エルミート行列の全体がなす \mathbb{R} 上のベクトル空間 $\mathcal{H} \subset \mathbb{C}^{M \times M}$ の部分空間 $\mathcal{V}(\omega)$ に属すと仮定する。 \mathcal{H} には Euclid 内積と Euclid ノルムが付随する。例えば、球面等方雑音 [3] のように、コヒーレンス行列が既知 $[\Gamma(\omega)]$ とする] の雑音の共分散行列は、

$$\mathcal{V}(\omega) = \{k\Gamma(\omega) | k \in \mathbb{R}\} \quad (5)$$

に属す。このとき、行列 $A \in \mathcal{H}$ の $\mathcal{V}(\omega)$ および $\mathcal{V}^\perp(\omega)$ への射影演算子は次式で与えられる。

$$\mathcal{P}(\omega)[A] = \frac{\text{tr}(A\Gamma(\omega))}{\text{tr}(\Gamma^2(\omega))} \Gamma(\omega) \quad (6)$$

$$\mathcal{P}^\perp(\omega)[A] = A - \mathcal{P}(\omega)[A] \quad (7)$$

3.3 方向性信号の共分散行列の復元アルゴリズム

推定すべき方向性信号の共分散行列を Θ とすると、3.1 節に述べたように、下式でモデル化できる。

$$\mathcal{P}^\perp[\Theta] = \mathcal{P}^\perp[\Phi_{xx}] \quad (8)$$

(本節では引数 τ, ω を省略する。) 上式を満たす Θ は無数に存在するが、ランク関数の凸緩和である [4] トレースノルム (特異値の和) $\|\Theta\|_*$ を用いて

$$\min_{\Theta \in \Psi} \frac{1}{2} \|\mathcal{P}^\perp[\Theta] - \mathcal{P}^\perp[\Phi_{xx}]\|_F^2 + \mu \|\Theta\|_* \quad (9)$$

* Diffuse Noise Robust Multiple Source Localization Based on Matrix Completion via Trace Norm Minimization by ITO, Nobutaka (The University of Tokyo), VINCENT, Emmanuel (INRIA), ONO, Nobutaka (The University of Tokyo), GRIBONVAL Rémi (INRIA), and SAGAYAMA, Shigeki (The University of Tokyo)

なる最適化問題を定義し、その解として Θ を推定する。ここで、 $\|\cdot\|_F$ は Frobenius ノルムであり、 Ψ は $M \times M$ の半正定値行列全体の集合を表す。

以下にこの最適化問題を解くアルゴリズムを示す。これは、制約条件が無い場合に対する Toh らのアルゴリズム [5] を修正して、半正定値制約 $\Theta \in \Psi$ を伴う本問題に適用したものである。

- $\hat{\Theta}^{(0)} = \hat{\Theta}^{(-1)} = \Phi_{xx}$; $t^{(0)} = t^{(-1)} = 1$
- $k = 0, 1, 2, \dots$ に対し、次を収束条件が満たされるまで反復する。
 - $Y^{(k)} = \hat{\Theta}^{(k)} + \frac{t^{(k-1)} - 1}{t^{(k)}} (\hat{\Theta}^{(k)} - \hat{\Theta}^{(k-1)})$
 - $G^{(k)} = \mathcal{P}[Y^{(k)}] + \mathcal{P}^\perp[\Phi_{xx}]$
 - $G^{(k)}$ の固有値分解を計算する: $G^{(k)} = U^{(k)} \Lambda^{(k)} U^{(k)H}$ 。但し $U^{(k)}$ はユニタリ。
 - $\hat{\Theta}^{(k+1)} = U^{(k)} (\Lambda^{(k)} - \mu I) U^{(k)H}$ 。但し $(\cdot)_+$ は行列の負の要素を 0 で置換する演算。
 - $t^{(k+1)} = \frac{1 + \sqrt{1 + 4t^{(k)2}}}{2}$

4 実環境雑音を用いた実験による性能評価

直径 5 cm の正方形 4 素子アレイを用いて新宿駅構内で録音した雑音 [6] と、鏡像法 [7] により生成した目的信号成分を加算して観測信号を得た。目的音声は ATR 日本語データベース [8] から取った。鏡像法における幾何学的な条件は下記の通りである。

- 部屋の寸法: 9 m × 7 m × 3 m
- アレイ位置: 部屋の中央に水平に配置
- 音源数: 2
- 音源位置: アレイと同一水平面内、1 m の距離、方位角 60°、330°

残響時間は 130 ms であった。1 番目のマイクロフォンにおける観測信号の SN 比は両音源とも -5 dB とした。信号長は 10 秒、サンプリング周波数は 16 kHz であった。

観測共分散行列から MUSIC スペクトルを計算する従来法と、推定した信号共分散行列から MUSIC スペクトルを計算する提案法を比較した。STFT のフレーム長は 512 点、フレームシフトは 64 点、窓はハミング窓とした。 Φ_{xx} は $x(\tau, \omega)x^H(\tau, \omega)$ を全フレームにわたり時間平均して求めた。提案法では、式 (5) の部分空間 $\mathcal{V}(\omega)$ を用い、 $\Gamma(\omega)$ は球面等方雑音場 [3] に対するコヒーレンス行列とした。また、アルゴリズムの反復の停止条件は $\frac{\|\hat{\Theta}^{(t)} - \hat{\Theta}^{(t-1)}\|_F}{\|\hat{\Theta}^{(t)}\|_F} < 10^{-4}$ とした。

提案法による方向性信号の共分散行列の推定精度を確認するため、Fig. 1 に真の方向性信号の共分散行列に対する相対誤差 $\frac{\|\hat{\Theta} - \Theta\|_F}{\|\Theta\|_F}$ を周波数の関数としてプロットする。破線が観測された共分散行列の相対誤差、実線が推定共分散行列の相対誤差である。提案法による共分散行列の雑音除去により相対誤差が著しく減少しており、この技術の有効性が確認できた。

Fig. 2 に両手法で計算した平均 MUSIC スペクトルを示す。これは各帯域で得られる MUSIC スペクトルを幾何平均して得られたものである。但し、SN 比の著しく低い 3 番目の帯域 (63 Hz) までは無視し、4 番目以降の帯域を用いて平均を計算した。図中の直線は音源の真の方位角を示す。従来法 (破線) は 60° の音源に対応するピークが鈍く、ピーク位置が大きすぎておらず、偽のピークも存在する。それに対し、提案法 (実線) は比較的正確な位置に鋭いピークがあり、偽のピークも抑制されていることが分かる。

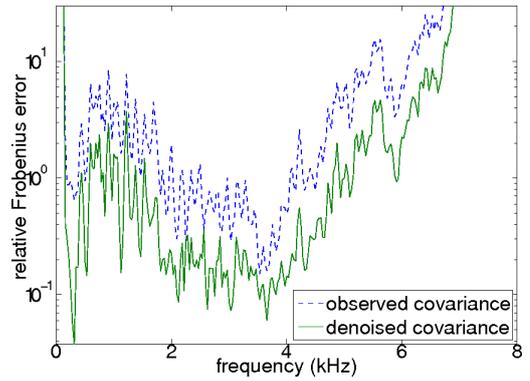


Fig. 1 真の方向性信号の共分散行列に対する相対誤差。破線: 観測共分散行列、実線: 推定共分散行列。

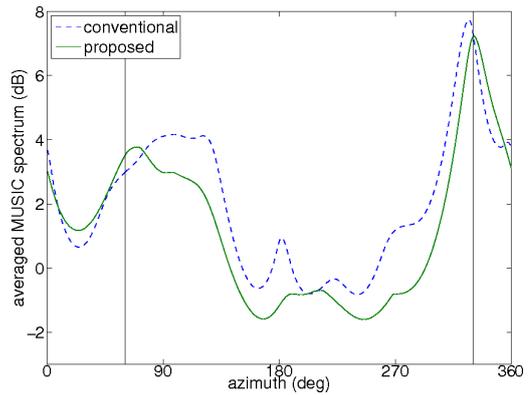


Fig. 2 平均 MUSIC スペクトルの例。破線: 従来法、実線: 提案法、直線: 真の音源の方位角。

5 結論

本稿では、観測された共分散行列から方向性信号の共分散行列を復元し、それに対して MUSIC を適用することにより、拡散性雑音の存在下でも正確に複数音源の方位角を推定する手法を示した。提案法では、方向性信号の共分散行列の低ランク性と、拡散性雑音の共分散行列の部分空間モデルに基づき、トレースノルム最小化により方向性信号の共分散行列を推定する。実環境雑音を用いた実験により、提案法を用いて従来の MUSIC よりも高精度に複数音源の方位角を推定できることが示された。

謝辞

本研究の一部は日本學術振興会科研費特別研究員奨励費 (22-6927) の助成を受けて行われた。

参考文献

- [1] Zelinski, *Proc. ICASSP*, pp. 2578–2581, 1988.
- [2] 伊藤ら, *信学技報*, EA2010-12, pp. 31–36, 2010.
- [3] Cook *et al.*, *JASA*, vol. 27, no. 6, pp. 1072–1077, 1955.
- [4] Ji *et al.*, *Proc. 26th Annual Int. Conf. Machine Learning*, pp. 457–464, 2009.
- [5] Toh *et al.*, <http://www.math.nus.edu.sg/~matys>
- [6] Ito *et al.*, *IEEE Trans. ASLP*. (Accepted).
- [7] Allen *et al.*, *JASA*, vol. 65, no. 4, pp. 943–950, 1979.
- [8] Kurematsu *et al.*, *Speech Commun.*, vol. 9, no. 4, pp. 357–363, 1990.