非均匀噪声场中的随机最大似然测向

Ming-yan Gong 和 Bin Lyu

arXiv:2302.04609v1 [eess.SP] 9 Feb 2023

摘要——本文采用并设计了期望–条件最大化（ECME）算法。该算法是 EM 算法的一种推广，用于解决未知非均匀噪声下相关随机信源的最大似然测向问题。与交替最大化不同，ECME 算法通过显式公式更新信源和噪声协方差矩阵的估计值，并能保证这两个估计值分别为半正定和正定。因此，ECME 算法计算效率高且运行稳定。仿真结果验证了该算法的有效性。

索引词——阵列处理、期望最大化、非均匀高斯噪声、随机信号模型。

I. 引言

IT 众所周知，在克拉美-罗下界（CRLB）和最大似然（ML）测向中，广泛使用两种信源信号模型，即确定性信号模型（信号为确定且未知）和随机信号模型（信号服从高斯分布）。例如，已有研究推导了基于这两种模型的各种 CRLB [1]–[7]。然而，对于这两种模型，最大似然测向通常都需要采用高维搜索算法，导致计算复杂度显著增加。

为降低计算复杂度，已有两种经典方法：交替最大化（AM）[8] 和期望最大化（EM） [9]–[12] 类算法。早期，这两种方法应用于均匀高斯噪声情形，从而减少了参数数量并简化了问题。然而，在许多情况下，均匀噪声模型并不现实，因此已有大量论文研究非均匀噪声 [4], [13]–[19]。在非均匀噪声下，协方差矩阵仍保持对角结构，但对角元素不再相同，这使到达方向（DOA）估计变得困难。为解决未知非均匀噪声下的测向问题，文献中提出了多种基于子空间技术的子空间分离方法 [13]–[18]。

为获得基于最大似然的解，AM和EM类算法也已应用于这一问题。然而，由于每次迭代时噪声具有非均匀性，AM类算法通常需要进行高维数值搜索 [4]‚[19], , 因此计算负担沉重。此外，对于高斯信号源， [19] 中提出的AM算法存在一个严重缺陷：无法保证信源和噪声协方差矩阵的估计值分别为半正定和正定。为此，我们在 [20], 中设计了几种计算效率较高的EM类算法，每次迭代仅需进行低维（一维或二维）数值搜索。然而，这些采用随机信号模型的EM类算法要求信源彼此不相关。这限制了随机最大似然测向在某些情况下的应用，例如多径环境。因此，亟需高效算法来解决这一问题。

本文采用并设计期望条件最大化（ECME）算法 [21], ——一种EM算法的推广形式，用于解决未知非均匀噪声下随机信源（可能彼此相关）的最大似然测向问题。与文献 [19], 中的交替最大化算法不同，ECME算法通过显式公式更新信源和噪声协方差矩阵的估计值，并能保证这两种估计值分别为半正定和正定。因此，ECME算法具有较高的计算效率且运行稳定。仿真结果验证了该算法的有效性。

II. 问题陈述

为简单起见，设一个由 W 个传感器组成的均匀间距线性阵列接收来自波长为 ι 的 V（$V < W$）个窄带信源入射的平面波。任意相邻传感器之间的距离为 $\iota/2$。我们将与第 v 个信源对应的方向记为 $\beta_v \in$ （0，π）弧度，并将接收信号写为

$$\mathbf{r}(t) = \sum_{v=1}^{V} \mathbf{a}(\beta_v) k_v(t) + \mathbf{j}(t) = \mathbf{A}(\boldsymbol{\beta})\mathbf{k}(t) + \mathbf{j}(t), \quad (1)$$

其中，a（β_v）$= [1\ a_v\ \cdots\ a_v^{W-1}]^T$，$a_v =$exp（$-\jmath\pi$cos（$\beta_v$）），$[\cdot]^T$ 表示转置，，（）表示第 v 个信源对应的信号，而 j（t）表示均值为零、协方差为 Q 的非均匀复高斯噪声，即 j（t）$\sim\mathcal{CN}$（0，Q）。这里，Q 为对角矩阵，表示为

$$\mathbf{Q} = \mathrm{diag}\{\boldsymbol{\delta}\}, \boldsymbol{\delta} = [\delta_1\ \cdots\ \delta_W]^T > \mathbf{0},$$

其中，Q 为正定矩阵，即 Q $\succ \mathbf{0}_W$（$\mathbf{0}_W$ 为零矩阵）。此外，若 $\delta_1 = \cdots = \delta_W = \delta > 0$，则 Q $=\delta \mathrm{I}_W$（I_W 为单位矩阵），此时噪声均匀。在式（1）中，A（$\boldsymbol{\beta}$）$=$[a（β_1）$\cdots$a（β_V）] 为阵列流形矩阵，$\boldsymbol{\beta} = [\beta_1\ \cdots\ \beta_V]^T \in \boldsymbol{\Gamma}$ 满足 $\boldsymbol{\Gamma} =$（0，π）^{V}，且 k（t）$= [k_1(t)\ \cdots\ k_V(t)]^T$。为方便记述，下文使用 A 代替 A（$\boldsymbol{\beta}$）。

我们考虑可能相关的高斯信号源，其满足 k(t) $\sim \mathcal{CN}$(0, O)，其中 O 是信源

协方差矩阵，且为半正定，即 $\mathbf{O} \succeq \mathbf{0}_V$。设信源与噪声不相关，则

$$\mathbf{r}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{G}), \mathbf{G} = \mathbf{AOA}^H + \mathbf{Q} \succ \mathbf{0}_W,$$

其中，$[\cdot]^H$ 表示共轭转置。在此基础上，L 个统计独立快拍的对数似然函数（LLF）可表示为

$$\begin{aligned} \mathcal{J}(\boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta}) = & \sum_{t=1}^{L} \log p\big(\mathbf{r}(t); \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta}\big) \\ = & f - L\big(\log|\mathbf{G}| + \operatorname{trace}\big[\mathbf{G}^{-1}\hat{\mathbf{R}}\big]\big), \end{aligned} \tag{2}$$

其中，$|\cdot|$、迹$[\cdot]$和$(\cdot)^{-1}$ 分别表示行列式、迹和求逆。在式（2）中，f为常数，$\mathbf{R} = (1/L)\sum_{t=1}^{L}\mathbf{r}(t)\mathbf{r}^H(t$表示快拍的协方差矩阵。此外，

$$\begin{aligned} \boldsymbol{\rho} = & \big([\mathbf{O}]_{1,1}, \ldots, [\mathbf{O}]_{V,V}, \operatorname{Re}\big\{[\mathbf{O}]_{1,2}\big\}, \operatorname{Im}\big\{[\mathbf{O}]_{1,2}\big\}, \ldots, \\ & \operatorname{Re}\big\{[\mathbf{O}]_{V-1,V}\big\}, \operatorname{Im}\big\{[\mathbf{O}]_{V-1,V}\big\}\big), \end{aligned}$$

其中，$[\mathbf{O}]_{p,\,q}$ 表示O的第(p,q)个元素，$\operatorname{Re}\{a\}$ 和$\operatorname{Im}\{a\}$分别表示 a的实部和虚部。因此，基于最大似然的到达方向估计问题为

$$\max_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}, \mathbf{O}\succeq\mathbf{0}_V, \boldsymbol{\delta}>\mathbf{0}} \mathcal{J}(\boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta}). \tag{3}$$

我们假设 $\mathcal{R}[\mathrm{A}] = V$，其中 $\mathcal{R}[\mathrm{A}]$ 表示A的秩，因此可以通过 [22]消去式（3）中的O

$$\hat{\mathbf{O}}(\boldsymbol{\beta}, \boldsymbol{\delta}) = \big(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\big)^{-1}\tilde{\mathbf{A}}^H\big(\widetilde{\mathbf{R}} - \mathbf{I}_W\big)\tilde{\mathbf{A}}\big(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\big)^{-1}, \tag{4}$$

其中，$\mathbf{Q}^{-1/2} = \operatorname{diag}\{1/\sqrt{\delta_1}, \ldots, 1/\sqrt{\delta_W}\}$,

$$\tilde{\mathbf{A}} = \mathbf{Q}^{-1/2}\mathbf{A}, \text{ and } \widetilde{\mathbf{R}} = \mathbf{Q}^{-1/2}\hat{\mathbf{R}}\mathbf{Q}^{-1/2}.$$

换言之，可以利用 $\boldsymbol{\beta}$和 $\boldsymbol{\delta}$的估计值来估计 $\boldsymbol{\rho}$。根据 (4)，G 可改写为

$$\mathbf{G} = \mathbf{Q}^{1/2}\big(\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}\widetilde{\mathbf{R}}\boldsymbol{\Pi}_{\tilde{\mathbf{A}}} + \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp}\big)\mathbf{Q}^{1/2},$$

其中，$\boldsymbol{\Pi}_{\tilde{\mathbf{A}}} = \tilde{\mathrm{A}}\big(\tilde{\mathrm{A}}^H\tilde{\mathrm{A}}\big)^{-1}\tilde{\mathrm{A}}^H$，且 $\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp} = \mathrm{I}_W - \boldsymbol{\Pi}_{\tilde{\mathrm{A}}}$。于是，问题 (3) 可化简为 [22]

$$\begin{aligned} \min_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}, \boldsymbol{\delta}>\mathbf{0}} \mathcal{H}(\boldsymbol{\beta}, \boldsymbol{\delta}) = & \log\big|\mathbf{Q}^{1/2}\big(\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}\widetilde{\mathbf{R}}\boldsymbol{\Pi}_{\tilde{\mathbf{A}}} + \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp}\big)\mathbf{Q}^{1/2}\big| \\ & + \operatorname{trace}\big[\big(\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}\widetilde{\mathbf{R}}\boldsymbol{\Pi}_{\tilde{\mathbf{A}}} + \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}^{\perp}\big)^{-1}\widetilde{\mathbf{R}}\big]. \end{aligned} \tag{5}$$

特别地，如果噪声为均匀高斯噪声，问题 (5) 可进一步化简为 [23]

$$\min_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}} \mathcal{G}(\boldsymbol{\beta}) = \big|\mathbf{A}\hat{\mathbf{O}}(\boldsymbol{\beta})\mathbf{A}^H + \hat{\delta}(\boldsymbol{\beta})\mathbf{I}_W\big|, \tag{6}$$

其中

$$\begin{aligned} \hat{\delta}(\boldsymbol{\beta}) &= \operatorname{trace}\big[\big(\mathbf{I}_W - \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\big)\hat{\mathbf{R}}\big]/(W-V), \\ \hat{\mathbf{O}}(\boldsymbol{\beta}) &= (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\big(\hat{\mathbf{R}} - \hat{\delta}(\boldsymbol{\beta})\mathbf{I}_W\big)\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}. \end{aligned}$$

遗憾的是，在非均匀高斯噪声下，很难将问题 (5) 化简为参数更少的问题。当然，由于搜索空间的维数为 $W+V$，且 $\mathcal{H}(\boldsymbol{\beta}, \boldsymbol{\delta})$ 的复杂度较高，应用梯度类算法搜索问题 (5) 的解在计算上开销很大。

事实上，当无法直接对所有参数进行最大化时，始终可以采用 AM算法。正如前文所述，[19] 中的作者针对问题 (3) 提出了一种 AM算法，每次迭代包含两个步骤。具体而言，第一步通过一种基于梯度的算法获得第 d 次迭代时 $\boldsymbol{\delta}$ 的估计值 $\boldsymbol{\delta}^{(d)}$。该算法称为“修正逆迭代算法”，并满足

$$\mathcal{J}(\boldsymbol{\beta}^{(d-1)}, \boldsymbol{\rho}^{(d-1)}, \boldsymbol{\delta}^{(d)}) \geq \mathcal{J}(\boldsymbol{\beta}^{(d-1)}, \boldsymbol{\rho}^{(d-1)}, \boldsymbol{\delta}^{(d-1)}), \tag{7}$$

其中，$[\cdot]^{(0)}$ 表示一个初始估计。然后，第二步通过以下方式同时求得 $\boldsymbol{\beta}^{(d)}$ 和 $\boldsymbol{\rho}^{(d)}$

$$(\boldsymbol{\beta}^{(d)}, \boldsymbol{\rho}^{(d)}) = \arg\max_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}, \mathbf{O}\succeq\mathbf{0}_V} \mathcal{J}(\boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta}^{(d)}), \tag{8}$$

以可分离的方式求解，即，

$$\boldsymbol{\beta}^{(d)} = \arg\min_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}} \mathcal{H}(\boldsymbol{\beta}, \boldsymbol{\delta}^{(d)}), \tag{9}$$

$$\mathbf{O}^{(d)} = \hat{\mathbf{O}}(\boldsymbol{\beta}^{(d)}, \boldsymbol{\delta}^{(d)}). \tag{10}$$

然而，AM算法有两个缺点：1）获取 $\boldsymbol{\delta}^{(d)}$ 和 $\boldsymbol{\beta}^{(d)}$ 的计算开销很大；2）无法保证$\mathrm{Q}^{(d)} \succ \mathbf{0}_W$（或 $\boldsymbol{\delta}^{(d)} > \mathbf{0}$）和$\mathrm{O}^{(d)} \succeq \mathbf{0}_V$ [24]，[25]。为高效获得式（3）中 $\boldsymbol{\beta}$ 的最大似然估计，我们将在下一节采用并设计ECME算法。

III. ECME算法

现有用于随机最大似然测向的EM类算法仅适用于不相关源 [9], [12]，[20]，，即O为对角矩阵。本节采用 ECME算法 [21]，（EM算法的一种推广）来求解与相关信源有关的问题（3）。

A. 步骤

式（1）中的信源可能相关，因此我们选择$\mathrm{K} = [\mathrm{k}(1) \ \cdots \mathrm{k}(L)]$ 和$\mathrm{J} = [\mathrm{j}(1) \ \cdots \mathrm{j}(L)]$ 作为增广数据。我们将增广数据对数似然函数表示为

$$\begin{aligned} \mathcal{M}(\mathbf{K}, \mathbf{J}, \boldsymbol{\rho}, \boldsymbol{\delta}) = & \sum_{t=1}^{L}\big[\log p(\mathbf{k}(t); \boldsymbol{\rho}) + \log p(\mathbf{j}(t); \boldsymbol{\delta})\big] \\ = & h - L\big(\log|\mathbf{O}| + \operatorname{trace}\big[\mathbf{O}^{-1}\hat{\mathbf{N}}_k\big]\big) + \\ & f - L\big(\log|\mathbf{Q}| + \operatorname{trace}\big[\mathbf{Q}^{-1}\hat{\mathbf{N}}_j\big]\big), \end{aligned} \tag{11}$$

其中 h 是常数，$\hat{\mathrm{N}}_k = (1/L)\sum_{t=1}^{L}\mathrm{k}(t)\mathrm{k}^H(t)$，且 $\hat{\mathrm{N}}_j = (1/L)\sum_{t=1}^{L}\mathrm{j}(t)\mathrm{j}^H(t)$。根据式（11），我们首先构建EM算法，并在下文推导第 d 次迭代中的期望步和最大化步骤。令 $\mathcal{E}\{\cdot\}$ 和 $\mathcal{D}\{\cdot\}$ 分别表示期望和协方差。

1）期望步：计算增广数据对数似然函数的条件期望，即

$$\begin{aligned} & \mathcal{M}\big(\boldsymbol{\rho}, \boldsymbol{\delta}; \boldsymbol{\Omega}^{(d-1)}\big) = \mathcal{E}\big\{\mathcal{M}(\mathbf{K}, \mathbf{J}, \boldsymbol{\rho}, \boldsymbol{\delta})\big|\mathbf{F}; \boldsymbol{\Omega}^{(d-1)}\big\} \\ = & \ h - L\big(\log|\mathbf{O}| + \operatorname{trace}\big[\mathbf{O}^{-1}\hat{\mathbf{N}}_k^{(d)}\big]\big) + \\ & \ f - L\big(\log|\mathbf{Q}| + \operatorname{trace}\big[\mathbf{Q}^{-1}\hat{\mathbf{N}}_j^{(d)}\big]\big) \end{aligned} \tag{12}$$

其中 $\Omega = (\boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta})$，且 $\mathrm{F} = [\mathrm{r}(1) \ \cdots \mathrm{r}(L)]$。此外，

$$\begin{aligned} \hat{\mathbf{N}}_k^{(d)} = & \ \mathcal{E}\big\{\hat{\mathbf{N}}_k\big|\mathbf{F}; \boldsymbol{\Omega}^{(d-1)}\big\} = [\mathbf{H}^{(d-1)}]^H\hat{\mathbf{R}}\mathbf{H}^{(d-1)} + \\ & \mathbf{O}^{(d-1)} - [\mathbf{H}^{(d-1)}]^H\mathbf{G}^{(d-1)}\mathbf{H}^{(d-1)} \succeq \mathbf{0}_V, \end{aligned} \tag{13}$$

$$\begin{aligned} \hat{\mathbf{N}}_j^{(d)} = & \ \mathcal{E}\big\{\hat{\mathbf{N}}_j\big|\mathbf{F}; \boldsymbol{\Omega}^{(d-1)}\big\} \\ = & \ \mathbf{Q}^{(d-1)}[\mathbf{G}^{(d-1)}]^{-1}\hat{\mathbf{R}}[\mathbf{G}^{(d-1)}]^{-1}\mathbf{Q}^{(d-1)} + \\ & \mathbf{Q}^{(d-1)} - \mathbf{Q}^{(d-1)}[\mathbf{G}^{(d-1)}]^{-1}\mathbf{Q}^{(d-1)} \succeq \mathbf{0}_W, \end{aligned} \tag{14}$$

其中 $\mathbf{H}^{(d-1)} = [\mathbf{G}^{(d-1)}]^{-1}\mathbf{A}^{(d-1)}\mathbf{O}^{(d-1)}$，$\mathbf{k}(t)$ 和 $\mathbf{j}(t)$ 的条件分布可在 [26], 中获得，且

$$\begin{aligned}
\mathcal{E}\big\{\mathbf{k}(t)\big|\mathbf{F};\boldsymbol{\Omega}^{(d-1)}\big\} &= [\mathbf{H}^{(d-1)}]^H\mathbf{r}(t),\\
\mathcal{D}\big\{\mathbf{k}(t)\big|\mathbf{F};\boldsymbol{\Omega}^{(d-1)}\big\} &= \mathbf{O}^{(d-1)} - \\
&\quad [\mathbf{H}^{(d-1)}]^H\mathbf{G}^{(d-1)}\mathbf{H}^{(d-1)} \succeq \mathbf{0}_V,\\
\mathcal{E}\big\{\mathbf{j}(t)\big|\mathbf{F};\boldsymbol{\Omega}^{(d-1)}\big\} &= \mathbf{Q}^{(d-1)}[\mathbf{G}^{(d-1)}]^{-1}\mathbf{r}(t),\\
\mathcal{D}\big\{\mathbf{j}(t)\big|\mathbf{F};\boldsymbol{\Omega}^{(d-1)}\big\} &= \mathbf{Q}^{(d-1)} - \\
&\quad \mathbf{Q}^{(d-1)}[\mathbf{G}^{(d-1)}]^{-1}\mathbf{Q}^{(d-1)} \succeq \mathbf{0}_W.
\end{aligned}$$

2) 最大化步骤：通过针对 $\boldsymbol{\rho}$ 和 $\boldsymbol{\delta}$ 最大化 $\mathcal{M}(\boldsymbol{\rho},\boldsymbol{\delta};\Omega^{(d-1)})$，得到 $\boldsymbol{\rho}^{(d)}$ 和 $\boldsymbol{\delta}^{(d)}$，从而得到两个并行子问题

$$\min_{\mathbf{O}\succeq\mathbf{0}_V} \log|\mathbf{O}| + \mathrm{trace}\big[\mathbf{O}^{-1}\hat{\mathbf{N}}_k^{(d)}\big], \tag{15}$$

$$\min_{\mathbf{Q}\succ\mathbf{0}_W} \log|\mathbf{Q}| + \mathrm{trace}\big[\mathbf{Q}^{-1}\hat{\mathbf{N}}_j^{(d)}\big]. \tag{16}$$

$\boldsymbol{\rho}^{(d)}$ 和 $\boldsymbol{\delta}^{(d)}$ 同时通过以下方式得到

$$\mathbf{O}^{(d)} = \hat{\mathbf{N}}_k^{(d)} \succeq \mathbf{0}_V, \tag{17}$$

$$\delta_w^{(d)} = \begin{cases} [\hat{\mathbf{N}}_j^{(d)}]_{w,w}, & [\hat{\mathbf{N}}_j^{(d)}]_{w,w} > 0,\\ \delta_w^{(d-1)}/2, & [\hat{\mathbf{N}}_j^{(d)}]_{w,w} = 0, \end{cases} \forall w. \tag{18}$$

由 (17) 和 (18) 可知，广义 EM 算法具有单调性 [27], ，即

$$\mathcal{J}\big(\boldsymbol{\beta}^{(d-1)},\boldsymbol{\rho}^{(d)},\boldsymbol{\delta}^{(d)}\big) \geq \mathcal{J}\big(\boldsymbol{\beta}^{(d-1)},\boldsymbol{\rho}^{(d-1)},\boldsymbol{\delta}^{(d-1)}\big). \tag{19}$$

显然，$\boldsymbol{\beta}^{(d)}$ 并非在 EM 算法的第 d 次迭代中得到。

3) 条件最大化步骤：为了得到 $\boldsymbol{\beta}^{(d)}$，我们现在在本次迭代中增加一个条件最大化步骤。考虑到单调性

$$\mathcal{J}\big(\boldsymbol{\beta}^{(d)},\boldsymbol{\rho}^{(d)},\boldsymbol{\delta}^{(d)}\big) \geq \mathcal{J}\big(\boldsymbol{\beta}^{(d-1)},\boldsymbol{\rho}^{(d)},\boldsymbol{\delta}^{(d)}\big), \boldsymbol{\beta}^{(d)} \in \boldsymbol{\Gamma}, \tag{20}$$

我们可以将此步骤设计为

$$\boldsymbol{\beta}^{(d)} = \arg\max_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}} \mathcal{J}\big(\boldsymbol{\beta},\boldsymbol{\rho}^{(d)},\boldsymbol{\delta}^{(d)}\big), \tag{21}$$

或者使用梯度型算法，根据（20）求得 $\boldsymbol{\beta}^{(d)}$，例如下一节中的算法 1。由于增加了与增广数据无关的步骤，上述 EM算法变为 ECME算法 [21]。

B. 稳定性和复杂度

算法的稳定运行要求 $\mathbf{Q}^{(d)} \succ \mathbf{0}_W$ （或 $\boldsymbol{\delta}^{(d)} > \mathbf{0}$）和 $\mathbf{O}^{(d)} \succeq \mathbf{0}_V$ for $d \geq 0$，因此我们给出以下注释。

注 1。在 ECME算法中，如果 $\mathbf{Q}^{(0)} \succ \mathbf{0}_W$ 和 $\mathbf{O}^{(0)} \succeq \mathbf{0}_V$，则 $\mathbf{Q}^{(d)} \succ \mathbf{0}_W$ （或 $\boldsymbol{\delta}^{(d)} > \mathbf{0}$）和 $\mathbf{O}^{(d)} \succeq \mathbf{0}_V$ for $d \geq 1$。

证明。我们采用数学归纳法。若 $\mathbf{Q}^{(u)} \succ \mathbf{0}_W$ （或 $\boldsymbol{\delta}^{(u)} > \mathbf{0}$）且 $\mathbf{O}^{(u)} \succeq \mathbf{0}_V$，则有 $\mathbf{G}^{(u)} = \mathbf{A}^{(u)}\mathbf{O}^{(u)}[\mathbf{A}^{(u)}]^H + \mathbf{Q}^{(u)} \succ \mathbf{0}_W$，由此得到式（14）中的 $\hat{\mathbf{N}}_j^{(u+1)} \succeq \mathbf{0}_W$，进而在式（18）中得到 $\delta_w^{(u+1)} > 0$、$\forall w$，即 $\mathbf{Q}^{(u+1)} \succ \mathbf{0}_W$。此外，式（13）中的 $\mathbf{O}^{(u+1)} = \hat{\mathbf{N}}_k^{(u+1)} \succeq \mathbf{0}_V$。证明完毕。 □

注 1 表明，在 ECME算法中，当 $\mathbf{Q}^{(0)} \succ \mathbf{0}_W$ 且 $\mathbf{O}^{(0)} \succeq \mathbf{0}_V$ 时，第 d 次迭代得到的 $\boldsymbol{\rho}^{(d)}$ 和 $\boldsymbol{\delta}^{(d)}$ 分别属于参数空间。因此，ECME算法运行稳定。

由于 $\boldsymbol{\rho}^{(d)}$ 和 $\boldsymbol{\delta}^{(d)}$ 可通过（17）和（18）中的显式公式获得，ECME算法的计算复杂度主要取决于在（20）中求得 $\boldsymbol{\beta}^{(d)}$。与 [19], 中的 AM算法相比，ECME算法因此具有较高的计算效率。

C. 极限点

根据 [21], [28], ，我们知道ECME算法满足一定的正则性条件，并且始终收敛于 $\mathcal{J}(\boldsymbol{\beta},\boldsymbol{\rho},\boldsymbol{\delta})$的某个驻点。不幸的是，$\mathcal{J}(\boldsymbol{\beta},\boldsymbol{\rho},\boldsymbol{\delta})$往往具有多个驻点，而ECME算法的极限点可能是一个不理想的驻点。为了解决这一问题，我们需要提供一个准确的初始点。按照 [19], 中的方法，我们可以假设噪声是均匀的，然后在一个粗略的V维网格上计算式（6）中的 $\mathcal{G}(\boldsymbol{\beta})$，以找到一个接近 $\mathcal{G}(\boldsymbol{\beta})$全局最小值的网格点，作为ECME算法的 $\boldsymbol{\beta}^{(0)}$。此外，鉴于 $\boldsymbol{\beta}$ [2]的随机最大似然估计具有更高的准确性，我们可以使用通过子空间 [29] 或基于稀疏表示 [30] 的算法获得的 $\boldsymbol{\beta}$估计，作为 $\boldsymbol{\beta}^{(0)}$。

在 $\boldsymbol{\rho}$的半正定区域边界上，即集合 $\eth = \{\boldsymbol{\rho} \mid \mathbf{O} \succeq \mathbf{0}_V, \mathcal{R}[\mathbf{O}] < V\}$，我们给出以下备注。令 $\mathcal{N}[\mathbf{O}]$ 表示O的零空间。

注2. 在ECME算法中，如果 $\mathbf{Q}^{(0)} \succ \mathbf{0}_W$ 且 $\mathbf{O}^{(0)} \succeq \mathbf{0}_V$，则 $\mathcal{N}[\mathbf{O}^{(d)}] = \mathcal{N}[\mathbf{O}^{(0)}]$ for$d \geq 1$。

证明。根据注 1，我们首先知道，由于 $\mathbf{Q}^{(0)} \succ \mathbf{0}_W$ 且 $\mathbf{O}^{(0)} \succeq \mathbf{0}_V$，因此 $\mathbf{Q}^{(d)} \succ \mathbf{0}_W$、$\mathbf{O}^{(d)} \succeq \mathbf{0}_V$ 以及 $\mathbf{G}^{(d)} \succ \mathbf{0}_W$ for $d \geq 0$。然后，采用数学归纳法进行证明。

如果 $\mathbf{O}^{(u)}\mathbf{v} = \mathbf{0}$，则根据（13）有 $\mathbf{O}^{(u+1)}\mathbf{v} = \hat{\mathbf{N}}_k^{(u+1)}\mathbf{v} = \mathbf{0}$，从而 $\mathcal{N}[\mathbf{O}^{(u)}] \subseteq \mathcal{N}[\mathbf{O}^{(u+1)}]$。此外，如果 $\mathbf{O}^{(u+1)}\mathbf{v} = \hat{\mathbf{N}}_k^{(u+1)}\mathbf{v} = \mathbf{0}$，则有 $\mathbf{v}^H\hat{\mathbf{N}}_k^{(u+1)}\mathbf{v} = 0$，并且在（13）中

$$\mathbf{v}^H[\mathbf{H}^{(u)}]^H\hat{\mathbf{R}}\mathbf{H}^{(u)}\mathbf{v} = 0 \Rightarrow [\mathbf{H}^{(u)}]^H\hat{\mathbf{R}}\mathbf{H}^{(u)}\mathbf{v} = \mathbf{0}, \tag{22}$$

$$\begin{aligned}
&\mathbf{v}^H\big(\mathbf{O}^{(u)} - [\mathbf{H}^{(u)}]^H\mathbf{G}^{(u)}\mathbf{H}^{(u)}\big)\mathbf{v} = 0\\
&\Rightarrow \big(\mathbf{O}^{(u)} - [\mathbf{H}^{(u)}]^H\mathbf{G}^{(u)}\mathbf{H}^{(u)}\big)\mathbf{v} = \mathbf{0}.
\end{aligned} \tag{23}$$

为继续推导，我们使用矩阵求逆公式 [22]

$$\mathbf{G}^{-1} = \mathbf{Q}^{-1/2}\big[\mathbf{I}_W - \tilde{\mathbf{A}}(\mathbf{O}\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} + \mathbf{I}_V)^{-1}\mathbf{O}\tilde{\mathbf{A}}^H\big]\mathbf{Q}^{-1/2} \tag{24}$$

并得到

$$\mathbf{O} - \mathbf{H}^H\mathbf{G}\mathbf{H} = (\mathbf{O}\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} + \mathbf{I}_V)^{-1}\mathbf{O}, \tag{25}$$

这表明 $\mathcal{N}[\mathbf{O}] = \mathcal{N}[\mathbf{O} - \mathbf{H}^H\mathbf{G}\mathbf{H}]$ 以及 $\mathcal{N}[\mathbf{O}^{(u)}] = \mathcal{N}[\mathbf{O}^{(u)} - [\mathbf{H}^{(u)}]^H\mathbf{G}^{(u)}\mathbf{H}^{(u)}]$。因此，将 $\mathbf{O}^{(u)}\mathbf{v} = \mathbf{0}$ 代入式（22）和（23），得到 $\mathcal{N}[\mathbf{O}^{(u)}] \supseteq \mathcal{N}[\mathbf{O}^{(u+1)}]$。最后，结合 $\mathcal{N}[\mathbf{O}^{(u)}] \subseteq \mathcal{N}[\mathbf{O}^{(u+1)}]$和 $\mathcal{N}[\mathbf{O}^{(u)}] \supseteq \mathcal{N}[\mathbf{O}^{(u+1)}]$，得到 $\mathcal{N}[\mathbf{O}^{(u)}] = \mathcal{N}[\mathbf{O}^{(u+1)}]$。证明完毕。 □

注2指出，如果ECME算法中的 $\boldsymbol{\rho}^{(0)}$ 位于边界上，即 $\boldsymbol{\rho}^{(0)} \in \eth$ 和 $\mathcal{N}[\mathbf{O}^{(0)}]$ 非空，那么 $\boldsymbol{\rho}$ 的极限点也位于边界上。因此，令$(\boldsymbol{\beta}^*,\boldsymbol{\rho}^*,\boldsymbol{\delta}^*)$表示问题(3)的解；若 $\boldsymbol{\rho}^* \in \eth$，我们可能需要

在执行ECME算法之前估计 $\mathcal{N}[\mathrm{O}^*]$。幸运的是，在实践中，$(\boldsymbol{\beta}^*, \boldsymbol{\rho}^*, \boldsymbol{\delta}^*)$始终是参数空间的内点（即 $\boldsymbol{\rho}^* \notin \eth$、$\mathrm{O}^* \succ \mathbf{0}_V$ 以及 $\mathcal{N}[\mathrm{O}^*]$ 为空），即使 $\boldsymbol{\rho}$ 的真实值位于边界上也是如此。因此，我们始终可以在ECME算法中采用$\mathrm{O}^{(0)} \succ \mathbf{0}_V$，例如图1中与相干源相关的仿真结果。

IV. 仿真结果

给出仿真结果以验证ECME算法的有效性，即ECME算法能够获得式（3）中 $\boldsymbol{\beta}$ 的最大似然估计。我们设定 $V = 2$、$\beta_1 = 50°$、$W = 6$、$\beta_2 = 100°$和 $\boldsymbol{\delta} = [1\ 2\ 3\ 4\ 2\ 10]^T$。采用算法1获得式（20）中的 $\boldsymbol{\beta}^{(d)}$，并将 $\|\boldsymbol{\beta}^{(u+1)} - \boldsymbol{\beta}^{(u)}\|_2 \leq 0.001°$用作停止准则。为获得 $\boldsymbol{\beta}$的最大似然估计，ECME算法采用一个准确的初始点。在图1和图2中，我们分别考虑具有以下特性的相干（或完全相关）源模型和部分相关源模型

$$\mathrm{O} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \text{ and } \mathrm{O} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}.$$

算法 1 基于最速下降法的DOA估计

$f(\boldsymbol{\beta}) = -\mathcal{J}\big(\boldsymbol{\beta}, \boldsymbol{\rho}^{(d)}, \boldsymbol{\delta}^{(d)}\big)/L$, 初始化 $\boldsymbol{\beta} = \boldsymbol{\beta}^{(d-1)} \in \boldsymbol{\Gamma}$。
当 $\|\nabla f(\boldsymbol{\beta})\|_2 > 0.001$ 时
$$t_v = \begin{cases} -(\pi - \beta_v)/f'_v(\boldsymbol{\beta}), & f'_v(\boldsymbol{\beta}) < 0, \\ \beta_v / f'_v(\boldsymbol{\beta}), & f'_v(\boldsymbol{\beta}) > 0, \\ \infty, & f'_v(\boldsymbol{\beta}) = 0, \end{cases} \quad \forall v.$$
$t = 0.1 \times$ 最小值$\{t_1, \ldots, t_V\}$.
当 $f\big(\boldsymbol{\beta} - t\nabla f(\boldsymbol{\beta})\big) > f(\boldsymbol{\beta}) - 0.3t\|\nabla f(\boldsymbol{\beta})\|_2^2$时执行
$t = 0.5t$。
结束循环
$\boldsymbol{\beta} = \boldsymbol{\beta} - t\nabla f(\boldsymbol{\beta}) \in \boldsymbol{\Gamma}$。
结束循环
$\boldsymbol{\beta}^{(d)} = \boldsymbol{\beta}$。

在图1中，我们比较了ECME算法的均方根误差（RMSE）性能与克拉美-罗下界 [4], [5]。此外，我们还仿真了 [20] 中针对不相关源的第二种空间交替广义EM（SAGE）算法，该SAGE算法采用与 [20]相同的仿真设置。每个均方根误差均基于2000次独立试验计算，且这两种算法采用相同的初始点。正如预期，我们可以看到，ECME算法获得的均方根误差小于SAGE算法。更重要的是，我们可以观察到，当快拍数 L 较大时，ECME算法达到 $\boldsymbol{\beta}$ 的克拉美-罗下界，这与众所周知的结论一致： $\boldsymbol{\beta}$ 的随机克拉美–罗下界可以由 $\boldsymbol{\beta}$[2]的随机最大似然估计量渐近地达到。因此，只要初始点足够准确，ECME算法就能够获得式（3）中的 $\boldsymbol{\beta}$ 随机最大似然估计。

在图2中，我们将ECME算法与两种基于子空间的算法进行比较。这两种算法采用了 [17]–[18] 中最先进的子空间分离方法，分别称为“方法 1+Root-MUSIC”和“方法 2+Root-MUSIC”。在每次试验中，三种算法处理相同的快拍。我们可以观察到，正如预期的那样，ECME

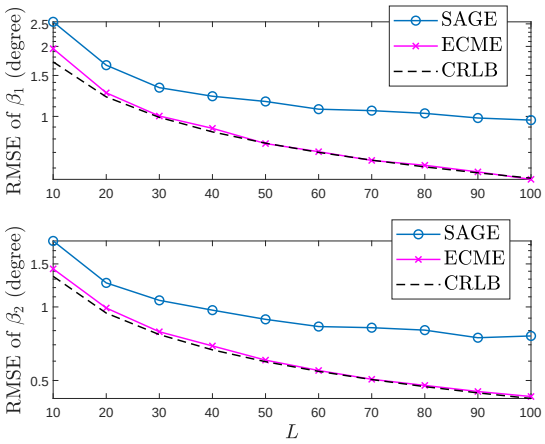

图1. ECME算法的均方根误差性能与克拉美-罗下界之间的关系。$\beta_1^{(0)} = 45°$、$\mathrm{O}^{(0)} = \mathrm{I}_V$、$\beta_2^{(0)} = 95°$和$\mathrm{Q}^{(0)} = \mathrm{I}_W$。

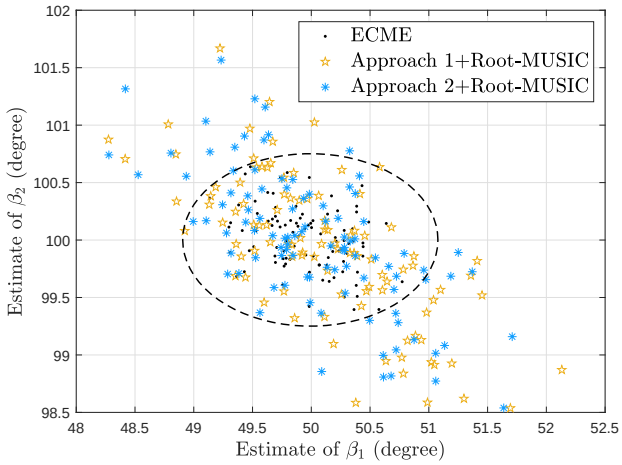

图2. 在100次独立试验中，ECME算法和两种基于子空间的算法得到的（β_1，β_2）估计值。$\beta_1^{(0)} = 45°$、$\beta_2^{(0)} = 95°$、$\mathrm{O}^{(0)} = \mathrm{I}_V$、$L = 100$和$\mathrm{Q}^{(0)} = \mathrm{I}_W$。

算法给出的(β_1, β_2)估计值更为接近，并以$(50°, 100°)$为中心，因为在到达方向估计中，最大似然技术在精度方面具有最高优势。

V. 结论

本文针对未知非均匀噪声下的随机最大似然测向（信源可能相关），采用并设计了ECME算法。理论分析表明，该算法计算效率高且运行稳定。仿真结果验证了该算法的有效性。

参考文献

[1] P. Stoica 和 A. Nehorai，《多重信号分类、最大似然与克拉美-罗界》，IEEE 声学、语音与信号处理汇刊，第37卷，第5期，第720–741页，1989年5月。[2] P. Stoica 和 A. Nehorai，《条件与非条件波达方向估计的性能研究》，IEEE 声学、语音与信号处理汇刊，第38卷，第10期，第1783–1795页，1990年10月。

[3] P. Stoica、E. G. Larsson 和 A. B. Gershman，《阵列处理的随机克拉美-罗界：教科书式推导》，IEEE 信号处理快报，第8卷，第5期，第148–150页，2001年5月。[4] M. Pesavento 和 A. B. Gershman，《存在未知非均匀噪声时的最大似然测向估计》，IEEE 信号处理汇刊，第49卷，第7期，第1310–1324页，2001年7月。[5] A. B. Gershman、M. Pesavento、P. Stoica 和 E. G. Larsson，《未知噪声场中的阵列处理随机克拉美-罗界》，载于 ICASSP 会议论文集，美国盐湖城，2001年5月。[6] J. Delmas 和 H. Abeida，《非圆信号的随机克拉美-罗界及其在到达方向估计中的应用》，IEEE 信号处理汇刊，第52卷，第11期，第3192–3199页，2004年11月。[7] H. Abeida 和 J. Delmas，《未知噪声场中非圆信号方向估计的高斯克拉美-罗界》，IEEE 信号处理汇刊，第53卷，第12期，第4610–4618页，2005年12月。[8] I. Ziskind 和 M. Wax，《通过交替投影对多个信源进行最大似然定位》，IEEE 声学、语音与信号处理汇刊，第36卷，第10期，第1553–1560页，1988年10月。[9] M. I. Miller 和 D. R. Fuhrmann，《最大似然窄带测向与 EM 算法》，IEEE 声学、语音与信号处理汇刊，第38卷，第9期，第1560–1577页，1990年9月。[10] P. Chung 和 J. F. Bohme，《DOA 估计中 EM 与 SAGE 算法的收敛性比较分析》，IEEE 信号处理汇刊，第49卷，第12期，第2940–2949页，2001年12月。[11] M. Gong 和 B. Lyu，《最大似然测向中的交替最大化与 EM 算法》，IEEE 车辆技术汇刊，第70卷，第10期，第9634–9645页，2021年10月。[12] M. Gong 和 B. Lyu，《存在未知均匀噪声时用于 DOA 估计的 EM 与 SAGE 算法》。 [在线]。可访问：https://arxiv.org/abs/2208 07510 。[13] A. M. Zoubir 和 S. Aouada，《非均匀噪声中的到达方向高分辨率估计》，载于 ICASSP 会议论文集，加拿大魁北克省蒙特利尔，2004年5月。[14] D. Madurasinghe，《一种适用于非均匀噪声的新型 DOA 估计器》，IEEE 信号处理快报，第12卷，第4期，第337–339页，2005年4月。[15] B. Liao、S. Chan、L. Huang 和 C. Guo，《非均匀噪声中子空间与 DOA 估计的迭代方法》，IEEE 信号处理汇刊，第64卷，第12期，第3008–3020页，2016年6月。[16] B. Liao、L. Huang、C. Guo 和 H. C. So，《未知非均匀噪声下利用传感器阵列进行到达方向估计的新方法》，IEEE 传感器期刊，第16卷，第24期，第8982–8989页，2016年12月。[17] M. Esfandiari、S. A. Vorobyov、S. Alibani 和 M. Karimi，《存在非均匀噪声时的非迭代子空间到达方向估计》，IEEE 信号处理快报，第26卷，第6期，第848–852页，2019年6月。[18] M. Esfandiari 和 S. A. Vorobyov，《存在非均匀噪声时的新型角度估计方法》，载于 ICASSP 会议论文集，新加坡，2022年4月。[19] C. E. Chen、F. Lorenzelli、R. E. Hudson 和 K. Yao，《存在未知非均匀噪声时的随机最大似然到达方向估计》，IEEE 信号处理汇刊，第56卷，第7期，第3038–3044页，2008年7月。[20] M. Gong 和 B. Lyu，《未知非均匀噪声下用于 DOA 估计的 EM 型算法》。 [在线]。可访问：https://arxiv.org/abs/2211 02458 。[21] C. Liu 和 D. B. Rubin，《ECME算法：对 EM 和 ECM 的简单扩展，可实现更快的单调收敛》，《生物计量学》，第81卷，第4期，第633–648页，1994年12月。[22] A. G. Jaffer，《随机信源的最大似然测向：一种可分离解》，载于 ICASSP 会议论文集，美国纽约，1988年4月。[23] P. Stoica 和 A. Nehorai，《阵列信号处理中的集中随机似然函数》，《电路、系统与信号处理》，第14卷，第5期，第669–674页，1995年9月。[24] Y. Bresler，《线性结构化协方差的最大似然估计及其在天线阵列处理中的应用》，载于第4届ASSP频谱估计与建模研讨会论文集，美国明尼苏达州明尼阿波利斯，1988年8月。[25] P. Stoica、B. Ottersten、M. Viberg 和 R. L. Moses，《随机相干信源的最大似然阵列处理》，IEEE 信号处理汇刊，第44卷，第1期，第96–105页，1996年1月。[26] I. B. Rhodes，《估计与滤波教程式导论》，IEEE 自动控制汇刊，第16卷，第6期，第688–706页，1971年12月。[27] A. P. Dempster、N. M. Laird 和 D. B. Rubin，《通过 EM 算法从不完整数据中进行最大似然估计》，皇家统计学会杂志，B辑（方法论），第 39 卷，38，1977 1，第 1 期，第 – 页。[28] C. F. Jeff Wu，《论 EM 算法的收敛性质》，统计学年刊，第 11 卷，第 1 期，第 95–103 页，1983 年 3 月。[29] P. Stoica 和 A. B. Gershman，《通过数据支持的网格搜索进行最大似然到达方向估计》，IEEE 信号处理快报，第 6 卷，第 10 期，第 273–275 页，1999 年 10 月。[30] D. Malioutov 、 M. Cetin 和 A. S. Willsky，《从稀疏信号重构视角看基于传感器阵列的源定位》，IEEE 信号处理汇刊，第 53 卷，第 8 期，第 3010–3022 页，2005 年 8 月。