

Задача оценивания угловых координат неподвижных объектов по наблюдениям с пропусками.

М. В. Лебедев^{*_a}, К. В. Семенихин^{*_b}, Г. Ф. Суржиков^{*_c}

**Московский авиационный институт (национальный исследовательский университет),
Москва, Россия*

e-mail: ^amax_max82@inbox.ru, ^csiemenku@gmail.com, ^cgfsurzhikov@gmail.com

Поступила в редакцию XX.XX.2026

Аннотация—В работе построен ЕМ-алгоритм для оценивания угловых координат неподвижных объектов при наличии пропусков в наблюдениях.

КЛЮЧЕВЫЕ СЛОВА: оценивание угловых координат, ЕМ-алгоритм, наблюдения с пропусками, линейная равномерная антенна решетка, оценка максимального правдоподобия.

1. ВВЕДЕНИЕ

В парадигме статистической обработки сигналов особое место занимает задача оценивания угловых координат (DoA) источников сигналов [1, 2]. Существуют две основные категории методов для оценивания угловых координат: подпространственные (к примеру, CAPON, MUSIC, ESPRIT), основанные на разбиении ковариационной матрицы наблюдений на подпространство, соответствующее сигналу, и подпространство, соответствующее шуму, и методы, основанные на поиске ММП-оценки. Последние оказываются особенно полезными в постановках задач, предполагающих относительно низкое отношение сигнал-шум, корреляцию между сигналами или малое число наблюдений.

Одним из классических подходов к поиску оценки максимального правдоподобия в моделях с неполными наблюдениями является ЕМ-алгоритм [3]. В контексте задач оценивания угловых координат источников скрытые переменные позволяют переформулировать задачу таким образом, что применение ЕМ-алгоритма становится естественным и вычислительно удобным. В практических приложениях зачастую используются модификации этого алгоритма, позволяющие упростить оптимизацию на M-шаге (ЕСМ-алгоритм, [4]), либо обеспечивающие более быструю сходимость (SAGE-алгоритм, [5]), и подобные методы уже применяются для оценивания угловых координат источников [7–9]. Данные алгоритмы используют предположение об известной ковариационной матрице шума. В последние годы начали появляться статьи [10], посвященные использованию ЕМ-алгоритма в условиях неопределенности скалярной ковариационной матрицы шума. В подобных алгоритмах, как правило, на M-шаге происходит вычисление сразу нескольких групп параметров, к примеру, могут быть вычислены угловые координаты и ковариационная матрица сигналов.

В реальных условиях эксплуатации сенсоры могут быть подвержены стохастическим аппаратным сбоям, вследствие чего их показания в определенные моменты времени могут не быть надежными.

В работе формулируется задача поиска оптимальных оценок угловых координат источников сигналов в условиях, когда часть сенсоров подвержены аппаратным сбоям или случайным помехам, и, ввиду этого, часть наблюдений содержит случайные пропуски. Используется стохастическая модель сигнала, ковариационная матрица шума известная и диагональная. Сигнал

и шум предполагаются независимыми и имеющими комплексное нормальное распределение. В качестве скрытых переменных рассматриваются исходные сигналы и пропуски в наблюдениях. Ввиду чувствительности EM-алгоритма к инициализации, используется мультистарт, для первой итерации используется оценка, полученная алгоритмом MUSIC, на последующих - случайные смещения указанной оценки. В условиях, когда число сенсоров оказывается меньше числа наблюдений, причем порядка трети сенсоров ненадежны, оценки, полученные с помощью алгоритма MUSIC оказываются существенно смещенными, в то же время предложенный в работе EM-алгоритм позволяет получить оценки угловых координат, для которых характерно значительно меньшее смещение относительно истинных значений параметров.

2. ПОСТАНОВКА ЗАДАЧИ

Предположим, имеется линейная равномерная антенная решетка, состоящая из L сенсоров, которая принимает узкополосные сигналы, направленные из K источников, причем $K < L$. Этим источникам соответствуют угловые координаты (DoA) $\theta = [\theta_1, \dots, \theta_K]^T$, практически не изменяющиеся во времени. По итогам измерений было получено T снимков полученного сигнала, причем ввиду стохастических технических сбоев, связанных с сенсорами, большая часть таких снимков содержит помимо надежных данных ненадежные, которые в рамках данной задачи рассматриваются как пропуски. Пусть X — набор наблюдений, полученных сенсорами в моменты времени $t = 1, \dots, T$, X_t соответствует наблюдению в момент времени t , через x и x_t будем обозначать реализации полного набора наблюдений и наблюдения в отдельный момент времени t соответственно. Ввиду наличия пропусков в данных, будем считать, что набор наблюдений X состоит из доступной части $X_o = \{X_{t,o_t}\}_{t=1}^T$ и недоступной: $X_m = \{X_{t,m_t}\}_{t=1}^T$, причем $o_t \cup m_t = \{1, \dots, L\}$, $o_t \cap m_t = \emptyset, \forall t \in \{1, \dots, T\}$. Предполагается, что $\#o_t : o_t = \emptyset$, т.е. нет таких наблюдений, которые состоят лишь из недоступной части.

Матрица X является результатом следующей модели наблюдений:

$$X = \mathbf{A}(\theta)S + N, \quad (1)$$

где

$$X = [X_1, \dots, X_T], S = [S_1, \dots, S_T], N = [N_1, \dots, N_T],$$

и $X_t \in \mathbb{C}^L, S_t \in \mathbb{C}^K, N_t \in \mathbb{C}^L$ — векторы-столбцы, соответствующие наблюдениям, сигналам и шумам в момент времени $t = 1, \dots, T$, $\mathbf{A}(\theta)$ — матрица направляющих векторов для линейной равномерной антенной решетки размера $L \times K$:

$$\mathbf{A}(\theta) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-2i\pi \frac{d}{\lambda} \sin(\theta_1)} & e^{-2i\pi \frac{d}{\lambda} \sin(\theta_2)} & \dots & e^{-2i\pi \frac{d}{\lambda} \sin(\theta_K)} \\ \dots & \dots & \ddots & \vdots \\ e^{-2i\pi(L-1)\frac{d}{\lambda} \sin(\theta_1)} & e^{-2i\pi(L-1)\frac{d}{\lambda} \sin(\theta_2)} & \dots & e^{-2i\pi(L-1)\frac{d}{\lambda} \sin(\theta_K)} \end{bmatrix},$$

где i — мнимая единица, d — расстояние между сенсорами, λ — длина волны, θ_k — угловая координата k -го источника, $k = 1, \dots, K$.

Сигналы, испускаемые источниками, шумы на сенсорах и наблюдения предполагаются стохастическими: $S_t \sim \mathcal{CN}(\mathbf{O}_{K \times 1}, \boldsymbol{\Gamma})$, $N_t \sim \mathcal{CN}(\mathbf{O}_{L \times 1}, \boldsymbol{\Lambda})$, $X_t \sim \mathcal{CN}(\mathbf{O}_{L \times 1}, \mathbf{A}(\theta)\boldsymbol{\Gamma}\mathbf{A}(\theta)^* + \boldsymbol{\Lambda})$, $t = 1, \dots, T$. Матрица $\boldsymbol{\Lambda}$ предполагается диагональной, т.е. шумы не коррелированы между собой, также в любой момент времени $S_t \perp N_t$. Требуется найти оптимальные оценки параметров $\Upsilon = (\theta, \boldsymbol{\Gamma})$.

3. ОПИСАНИЕ АЛГОРИТМА

Воспользуемся ЕМ-алгоритмом для того, чтобы оценить значения параметров $\Upsilon, \Xi = (X_m, S)$ – латентные переменные (недоступные значения и сигналы). Пусть $\Xi_t = (X_{t,m_t}, S_t)$ – набор латентных переменных, соответствующих наблюдению X_t , причем для отдельных наблюдений $X_{t,m_t} = \emptyset$. Алгоритм состоит из двух шагов:

- **E-шаг:** найти математическое ожидание полного правдоподобия с учетом апостериорного распределения латентных переменных:

$$\mathbb{E}_{q(\Xi)}[\log P(X, S)],$$

где

$$q(\Xi) = P(X_m, S | X_o = x_o, \Upsilon^{(\tau-1)}),$$

причем τ – номер итерации ЕМ-алгоритма.

- **M-шаг:**

$$\Upsilon^{(\tau)} = \arg \max_{\Upsilon \in \mathbb{R}^{K^2+K}} \mathbb{E}_{q(\Xi)}[\log P(X, S)].$$

3.1. E-шаг

Из построения модели наблюдений следует независимость и одинаковая распределенность наблюдений $X_t, t = 1, \dots, T$. Требуется найти математическое ожидание полного правдоподобия с учетом текущей оценки параметров и апостериорного совместного распределения пропущенных значений в наблюдениях X_m и сигналов S

$$\mathbb{E}_{q(\Xi)}[\log P(X, S)] = \mathbb{E}_{(X_m, S) | X_o = x_o, \Upsilon^{(\tau-1)}}[\log P(X, S)].$$

Преобразуем выражение, учитывая тот факт, что наблюдения являются независимыми и одинаково распределенными, и обозначив через \mathcal{I} доступную информацию $X_o = x_o, \Upsilon^{(\tau-1)}$ на итерации τ и через $q(\Xi_t)$ распределение $P(X_{t,m_t}, S_t | X_{t,o_t} = x_{t,o_t}, \Upsilon^{(\tau-1)})$:

$$\begin{aligned} \mathbb{E}_{q(\Xi)}[\log P(X, S)] &= \\ \mathbb{E}_{q(\Xi)}[\log P(X | S) + \log P(S)] &= \\ \left[\sum_{t=1}^T \mathbb{E}_{q(\Xi_t)} \log [P(X_t | S_t)] + \sum_{t=1}^T \mathbb{E}_{q(\Xi_t)} [\log P(S_t)] \right] &= \\ -T \left[\log |\Lambda| + \text{Tr} (\Lambda^{-1} \mathbb{E}_{q(\Xi)}[XX^*]) - 2 \operatorname{Re} \text{Tr} (\mathbf{A}^*(\theta) \Lambda^{-1} \mathbb{E}_{q(\Xi)}[XS^*]) \right. & \\ \left. + \text{Tr} (\Lambda^{-1} \mathbf{A}(\theta) \mathbb{E}_{q(\Xi)}[SS^*] \mathbf{A}^*(\theta)) + \log |\Gamma| + \text{Tr} (\Gamma^{-1} \mathbb{E}_{q(\Xi)}[SS^*]) \right]. \end{aligned}$$

Распределение $P(X_{t,m_t}, S_t | X_{t,o_t})$ задается однозначно следующими условиями:

1. S_t и N_t независимы $\forall t \in \{1, \dots, T\}$;
2. $\mathbf{A}(\theta)$ – матрица линейного преобразования;
3. S_t, N_t – комплексные гауссовые векторы, которые имеют нулевую псевдоковариацию.

На практике, не требуется напрямую находить ковариации для распределений $q(\Xi_t), t = 1, \dots, T$, достаточно лишь найти условные вторые моменты $\mathbb{E}_{q(\Xi)}[XX^*], \mathbb{E}_{q(\Xi)}[XS^*], \mathbb{E}_{q(\Xi)}[SS^*]$.

Ввиду того, что наблюдения $X_t, t = 1, \dots, T$ являются независимыми и одинаково распределенными, мы можем найти их путем усреднения соответствующих условных вторых моментов по всем имеющимся наблюдениям:

$$\mathbb{E}_{q(\Xi)}[XX^*] = \frac{1}{T} \sum_{t=1}^T \mathbb{E}[X_t X_t^* | \mathcal{I}_t],$$

$$\mathbb{E}_{q(\Xi)}[SS^*] = \frac{1}{T} \sum_{t=1}^T \mathbb{E}[S_t S_t^* | \mathcal{I}_t],$$

$$\mathbb{E}_{q(\Xi)}[XS^*] = \frac{1}{T} \sum_{t=1}^T \mathbb{E}[X_t S_t^* | \mathcal{I}_t].$$

Рассмотрим случай, когда наблюдение X_t является полностью доступным, т.е. $m_t = \emptyset$. Тогда, апостериорное распределение $\mathsf{P}(X_{t,m_t}, S_t | \mathcal{I}_t)$ сводится к распределению $\mathsf{P}(S_t | \mathcal{I}_t)$, а условные моменты $\mathbb{E}[X_t X_t^* | \mathcal{I}_t]$, $\mathbb{E}[X_t S_t^* | \mathcal{I}_t]$, $\mathbb{E}[S_t S_t^* | \mathcal{I}_t]$ вычисляются по следующей схеме. Находим параметры апостериорного распределения $\mathsf{P}(S_t | \mathcal{I}_t), t = 1, \dots, T$:

$$\begin{cases} \mathbb{E}[S_t | \mathcal{I}_t] = \boldsymbol{\Gamma}^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} x_t, \\ \text{Cov}(S_t | \mathcal{I}_t) = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} \mathbf{A} \boldsymbol{\Gamma}. \end{cases}$$

И вычисляем условные моменты следующим образом:

$$\mathbb{E}[X_t X_t^* | \mathcal{I}_t] = \mathbb{E}[X_t | \mathcal{I}_t] \cdot \mathbb{E}[X_t^* | \mathcal{I}_t] = x_t x_t^*,$$

$$\mathbb{E}[X_t S_t^* | \mathcal{I}_t] = \mathbb{E}[X_t | \mathcal{I}_t] \cdot \mathbb{E}[S_t^* | \mathcal{I}_t] = x_t \cdot \mathbb{E}[S_t^* | \mathcal{I}_t] = x_t \cdot x_t^* \cdot \hat{\mathbf{R}}^{-1} \mathbf{A} \boldsymbol{\Gamma},$$

$$\mathbb{E}[S_t S_t^* | \mathcal{I}_t] = \mathbb{E}[S_t | \mathcal{I}_t] \cdot \mathbb{E}[S_t^* | \mathcal{I}_t] + \text{Cov}(S_t | \mathcal{I}_t).$$

Теперь рассмотрим случай, когда наблюдение X_t содержит пропуски, т.е. $m_t \neq \emptyset$. Сначала найдем апостериорное распределение $\mathsf{P}(X_m | \mathcal{I})$. Для достижения этой цели, для каждой пары $\{(o_t, m_t) : m_t \neq \emptyset\}$ создадим разбиение оценки ковариационной матрицы наблюдений $\hat{\mathbf{R}}$ на блоки, индуцированное этим разбиением множества индексов, оно имеет следующий вид:

$$\hat{\mathbf{R}} = \begin{pmatrix} \hat{\mathbf{R}}_{o_t, o_t} & \hat{\mathbf{R}}_{o_t, m_t} \\ \hat{\mathbf{R}}_{m_t, o_t} & \hat{\mathbf{R}}_{m_t, m_t} \end{pmatrix},$$

где каждый блок определяется как

$$\hat{\mathbf{R}}_{a,b} = (\hat{\mathbf{R}}_{ij})_{i \in a, j \in b}.$$

Для каждого наблюдения, содержащего пропуски, требуется найти апостериорное распределение пропущенных значений, $\mathsf{P}(X_{t,m_t} | X_{t,o_t} = x_{t,o_t}, \Upsilon^{(\tau-1)}), t = 1, \dots, T, m_t \neq \emptyset$. Обозначим через \mathcal{I}_t доступную информацию $X_{t,o_t} = x_{t,o_t}, \Upsilon^{(\tau-1)}$ на итерации τ . Параметры апостериорного распределения $\mathsf{P}(X_{t,m_t} | \mathcal{I}_t), t = 1, \dots, T$ на итерации τ можно найти следующим образом:

$$\begin{cases} \mathbb{E}[X_{t,m_t} | \mathcal{I}_t] = \hat{\mathbf{R}}_{m_t, o_t} (\hat{\mathbf{R}}_{o_t, o_t})^{-1} \cdot x_{t,o_t}, \\ \text{Cov}(X_{t,m_t} | \mathcal{I}_t) = \hat{\mathbf{R}}_{m_t, m_t} - \hat{\mathbf{R}}_{m_t, o_t} (\hat{\mathbf{R}}_{o_t, o_t})^{-1} \hat{\mathbf{R}}_{o_t, m_t}, \end{cases}$$

где $\hat{\mathbf{R}}_{o_t, o_t} = \hat{\mathbf{R}}_{o_t, o_t}^{(\tau-1)}$, $\hat{\mathbf{R}}_{o_t, m_t} = \hat{\mathbf{R}}_{o_t, m_t}^{(\tau-1)}$, $\hat{\mathbf{R}}_{m_t, o_t} = \hat{\mathbf{R}}_{m_t, o_t}^{(\tau-1)}$, $\hat{\mathbf{R}}_{m_t, m_t} = \hat{\mathbf{R}}_{m_t, m_t}^{(\tau-1)}$.

Для каждого наблюдения X_t , содержащего пропуски, определим условные первый и второй начальный момент $\mathbb{E}[X_t X_t^* | \mathcal{I}_t]$:

$$\begin{aligned}\mathbb{E}[X_t | \mathcal{I}_t] &= \mathbb{E}\left[\begin{pmatrix} X_{t,o_t} \\ X_{t,m_t} \end{pmatrix} | \mathcal{I}_t\right] = \begin{pmatrix} x_{t,o_t} \\ \mathbb{E}[X_{t,m_t} | \mathcal{I}_t] \end{pmatrix}, \\ \mathbb{E}[X_t X_t^* | \mathcal{I}_t] &= \mathbb{E}[X_t | \mathcal{I}_t] \cdot \mathbb{E}[X_t^* | \mathcal{I}_t] + \text{Cov}(X_t | \mathcal{I}_t) \\ &= \mathbb{E}[X_t | \mathcal{I}_t] \cdot \mathbb{E}[X_t^* | \mathcal{I}_t] + \begin{pmatrix} \mathbf{O}_{o_t, o_t} & \mathbf{O}_{o_t, m_t} \\ \mathbf{O}_{m_t, o_t} & \text{Cov}(X_{t,m_t} | \mathcal{I}_t) \end{pmatrix},\end{aligned}$$

где \mathbf{O}_{o_t, o_t} , \mathbf{O}_{o_t, m_t} , \mathbf{O}_{m_t, o_t} – нулевые блочные матрицы. Разбиение указанной матрицы на четыре блока, три из которых состоят из нулей, индуктировано разбиением множества индексов $\{1, \dots, L\}$ на множества o_t, m_t .

Параметры апостериорного распределения $\mathsf{P}(S_t | \mathcal{I}_t), t = 1, \dots, T$ можно найти следующим образом:

$$\begin{cases} \mathbb{E}[S_t | \mathcal{I}_t] = \mathbf{A}^* \hat{\mathbf{R}}^{-1} \mathbb{E}[X_t | \mathcal{I}_t], \\ \text{Cov}(S_t | \mathcal{I}_t) = \mathbf{A} - \mathbf{A}^* \hat{\mathbf{R}}^{-1} \mathbf{A} \Gamma + \mathbf{A}^* \hat{\mathbf{R}}^{-1} \text{Cov}(X_t | \mathcal{I}_t) \hat{\mathbf{R}}^{-1} \mathbf{A} \Gamma, \end{cases}$$

где $\mathbf{A} = \mathbf{A}(\theta^{(\tau-1)})$, $\Gamma = \Gamma^{(\tau-1)}$, $\hat{\mathbf{R}} = \hat{\mathbf{R}}^{(\tau-1)}$; эти же обозначения используются и далее, если не оговорено иное. Для вывода этих формул используется телескопическое свойство условного математического ожидания и закон полной дисперсии (подробности приведены в приложении Б).

Оценим $\mathbb{E}[S_t S_t^* | \mathcal{I}_t]$:

$$\mathbb{E}[S_t S_t^* | \mathcal{I}_t] = \mathbb{E}[S_t | \mathcal{I}_t] \cdot \mathbb{E}[S_t^* | \mathcal{I}_t] + \text{Cov}(S_t | \mathcal{I}_t),$$

Оценим $\mathbb{E}[X_t S_t^* | \mathcal{I}_t]$:

$$\mathbb{E}[X_t S_t^* | \mathcal{I}_t] = \mathbb{E}[X_t X_t^* | \mathcal{I}_t] \hat{\mathbf{R}}^{-1} \mathbf{A} \Gamma.$$

3.2. M-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\begin{aligned}\Upsilon^{(\tau)} &= \arg \max_{\Upsilon \in \mathbb{R}^{K^2+K}} \mathbb{E}_{q(\Xi)} [\log \mathsf{P}(X, S)] = \\ \arg \min_{\Upsilon \in \mathbb{R}^{K^2+K}} \text{T} \Big[\log |\Lambda| &+ \text{Tr} (\Lambda^{-1} \mathbb{E}_{q(\Xi)} [XX^*]) - 2 \operatorname{Re} \text{Tr} (\mathbf{A}^*(\theta) \Lambda^{-1} \mathbb{E}_{q(\Xi)} [XS^*]) \\ &+ \text{Tr} (\Lambda^{-1} \mathbf{A}(\theta) \mathbb{E}_{q(\Xi)} [SS^*] \mathbf{A}^*(\theta)) + \log |\Gamma| + \text{Tr} (\Gamma^{-1} \mathbb{E}_{q(\Xi)} [SS^*]) \Big].\end{aligned}$$

Оценим угловые координаты источников θ :

$$\begin{aligned}\theta^{(\tau)} &= \arg \min_{\theta \in \mathbb{R}^K} \mathcal{Q}_1(\theta | \theta^{(\tau-1)}) = \arg \min_{\theta \in \mathbb{R}^K} \Big[-2 \operatorname{Re} \text{Tr} (\mathbf{A}^*(\theta) \Lambda^{-1} \mathbb{E}_{q(\Xi)} [XS^*]) \\ &\quad + \text{Tr} (\Lambda^{-1} \mathbf{A}(\theta) \mathbb{E}_{q(\Xi)} [SS^*] \mathbf{A}^*(\theta)) \Big].\end{aligned}$$

Эту задачу можно решить численно, подробности приведены в приложении В.

Оценим ковариацию сигналов Γ :

$$\Gamma^{(\tau)} = \arg \min_{\Gamma \in \mathbb{R}^{K^2}} \mathcal{Q}_2(\Gamma | \Gamma^{(\tau-1)}) = \arg \min_{\Gamma \in \mathbb{R}^{K^2}} \text{T} \Big[\log |\Gamma| + \text{Tr} (\Gamma^{-1} \mathbb{E}_{q(\Xi)} [SS^*]) \Big] = \mathbb{E}_{q(\Xi)} [SS^*].$$

Обновляем оценку ковариации наблюдений с учетом полученных оценок параметров:

$$\hat{\mathbf{R}}^{(\tau)} = \mathbf{A}(\theta^{(\tau)}) \Gamma^{(\tau)} \mathbf{A}^*(\theta^{(\tau)}) + \Lambda.$$

4. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО ЭКСПЕРИМЕНТА

4.1. Пример 1

Число сенсоров	1000
Частота	1000 Гц
Приведенное давление шума p_n	0.0016 Па/ $\sqrt{\text{Гц}}$
Приведенное давление сигнала p_s	0.01 Па/ $\sqrt{\text{Гц}}$
Длина волны	1.5 м
Расстояние до источника r	100 м, 1000 м
Расстояние между сенсорами	0.75 м

Таблица 1. Исходные данные

Дисперсия сигнала:

$$d_s = p_s^2 / (r/r_0)^2, \quad r_0 = 1 \text{ м},$$

Дисперсия шума:

$$d_n = p_n^2.$$

4.2. Первый набор начальных условий

 $L = 25, K = 1, T = 12, 5$ сенсоров неисправны, 50% пропусков для каждого из них,

$$\Gamma = 0.5 \cdot E_K, \Lambda = 8.1 \cdot E_L, \theta = [0.7] \approx [40.107^\circ].$$

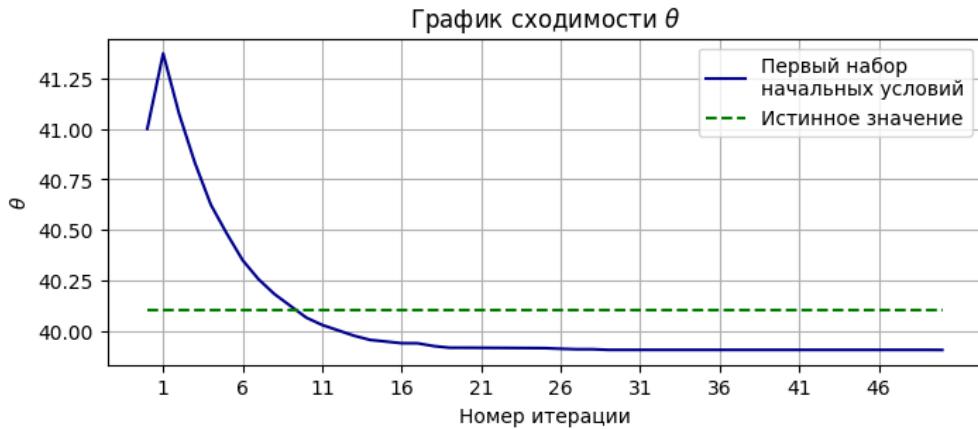


Рис. 1. График сходимости угловой координаты для первого набора начальных условий

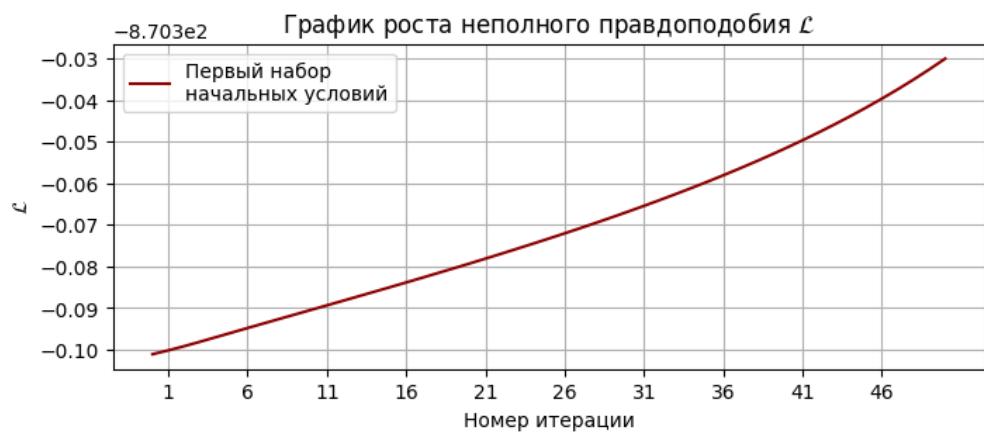


Рис. 2. График роста $\log P(X_o | \Upsilon)$ для первого набора начальных условий

4.3. Второй набор начальных условий

$L = 25, K = 1, T = 11, 8$ сенсоров неисправны, 50% пропусков для каждого из них,
 $\Gamma = 0.5 \cdot E_K, \Lambda = 6.1 \cdot E_L, \theta = [0.7] \approx [40.107^\circ]$.

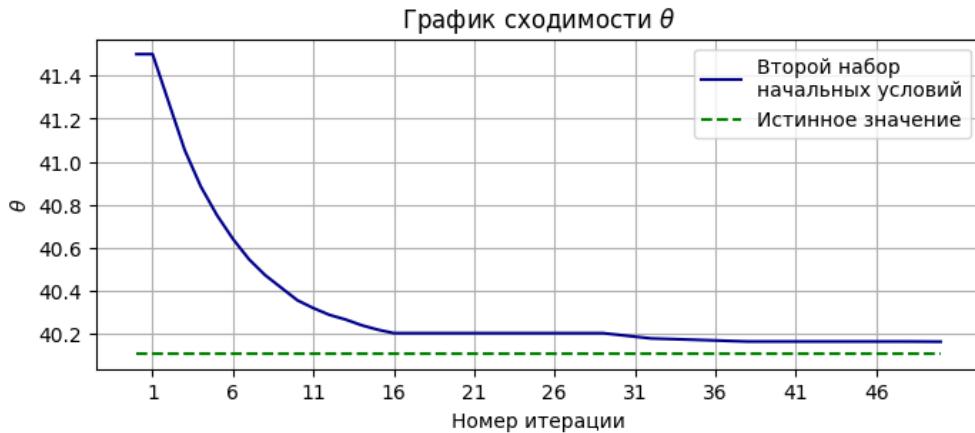


Рис. 3. График сходимости угловой координаты для второго набора начальных условий

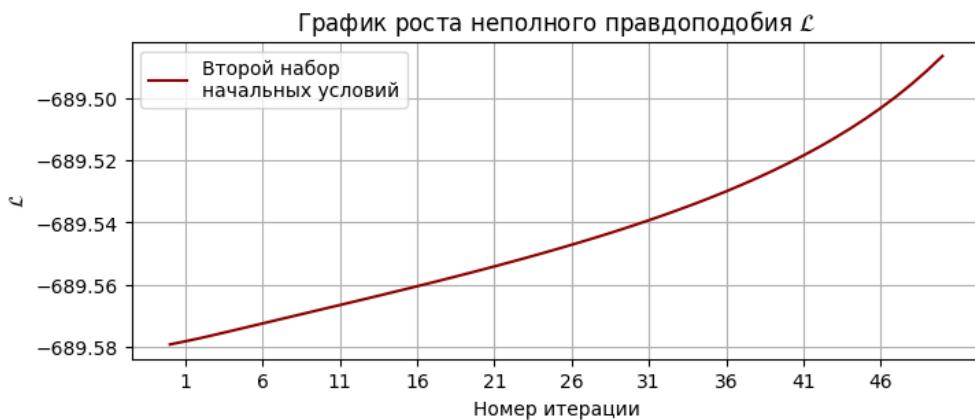


Рис. 4. График роста $\log P(X_o | \Upsilon)$ для второго набора начальных условий

5. ЗАКЛЮЧЕНИЕ

ПРИЛОЖЕНИЕ А. ИНИЦИАЛИЗАЦИЯ

Оценим ковариационную матрицу $\hat{\mathbf{R}}^{(0)}$ следующим образом: если $\sum_t \mathbf{I}_{x_t=x_{t,o_t}} \geq L$, оцениваем $\hat{\mathbf{R}}^{(0)}$ по выборке, образованной $x_t : x_t = x_{t,o_t}$. В противном случае используется следующий подход: пусть x^l – строка матрицы X , соответствующая сенсору l , $l = 1, \dots, L$, x_o^l – выборка, образованная доступными наблюдениями в строке l , x_m^l – выборка, образованная недоступными наблюдениями в строке l , x_t^l – компонента наблюдения x_t , соответствующая сенсору l . Для всех недоступных значений построим следующую оценку: если $x_t^l \in x_m^l$, $l = 1, \dots, L$, $t = 1, \dots, T$, то $\hat{x}_t^l = \overline{x_o^l}$. Оцениваем ковариацию по такому оцененному набору \hat{x} :

$$\hat{\mathbf{R}}^{(0)} = \text{Cov}(\hat{x}).$$

Оценим вектор угловых координат источников $\theta^{(0)}$ следующим образом:

1. Выберем число ν , которое будет соответствовать первому компоненту вектора $\theta^{(0)}$:

$$\nu \sim \mathcal{U}(-\pi; \pi);$$

2. Оценим компоненты вектора $\theta^{(0)}$ так: $\theta_i^{(0)} = (\nu + (i - 1) \cdot \frac{2\pi}{K}) \bmod 2\pi$, $i = 1, \dots, K$. При этом, $a \bmod b = a - b \cdot \lfloor \frac{a}{b} \rfloor$.

Альтернативный подход: используем $\hat{\mathbf{R}}^{(0)}$ для оценки $\theta^{(0)}$ с помощью алгоритма MUSIC, используем эту оценку на первой итерации мультистарта EM, на следующих используем оценки в окрестности указанной.

Начальную ковариацию сигналов $\Gamma^{(0)}$ получаем на основе метода наименьших квадратов с помощью $\theta^{(0)}$, $\hat{\mathbf{R}}^{(0)}$. Матрица $\mathbf{A}(\theta^{(0)})$ – матрица векторов направленности и представима в следующем виде:

$$\mathbf{A}(\theta^{(0)}) = [\mathbf{A}(\theta_1^{(0)}), \dots, \mathbf{A}(\theta_K^{(0)})],$$

где $\mathbf{A}(\theta_k^{(0)})$ – вектор направленности для источника k . Перед применением МНК каждый такой вектор отнормируем:

$$\check{\mathbf{A}}(\theta_k^{(0)}) = \frac{\mathbf{A}(\theta_k^{(0)})}{\|\mathbf{A}(\theta_k^{(0)})\|}, k = 1, \dots, K,$$

сформируем из них нормированную матрицу векторов направленности:

$$\check{\mathbf{A}}(\theta^{(0)}) = [\check{\mathbf{A}}(\theta_1^{(0)}), \dots, \check{\mathbf{A}}(\theta_K^{(0)})],$$

применяем МНК для получения нормированной ковариации мощностей:

$$\check{\Gamma}^{(0)} = \mathcal{D}[\check{\mathbf{A}}(\theta^{(0)})^+ (\hat{\mathbf{R}}^{(0)} - \Lambda) (\check{\mathbf{A}}(\theta^{(0)})^+)^*],$$

где $\mathcal{D}[\cdot]$ – диагональное приближение матрицы, $(\cdot)^+$ – псевдообратная матрица. Такой подход не гарантирует, что все элементы матрицы на главной диагонали будут неотрицательными, поэтому используем следующую корректировку: $\check{\Gamma}_{ll}^{(0)} = \max(\check{\Gamma}_{ll}^{(0)}, \varepsilon)$, $l = 1, \dots, L$. Переходим к ненормированной мощности:

$$\Gamma_{ll}^{(0)} = \frac{\check{\Gamma}_{ll}^{(0)}}{\|a(\theta_k^{(0)})\|_l}, l = 1, \dots, L.$$

Требуется гарантировать, что при выбранной инициализации на Е-шаге не будет получена условная ковариация сигналов, не являющаяся положительной определенной. Вычисляем оценку условной ковариации сигналов, используя начальную оценку, и проверяем, выполняется ли свойство неотрицательной определенности:

$$\boldsymbol{\Gamma}^{(0)} - (\boldsymbol{\Gamma}^{(0)})^* \mathbf{A}(\theta^{(0)})^* (\hat{\mathbf{R}}^{(0)})^{-1} \mathbf{A}(\theta^{(0)}) \boldsymbol{\Gamma}^{(0)} \succeq 0,$$

пока это свойство не выполняется, домножаем матрицу на 0.5: $\boldsymbol{\Gamma}^{(0)} = 0.5 \cdot \boldsymbol{\Gamma}^{(0)}$ и повторно проверяем выполнимость вышеуказанного свойства.

ПРИЛОЖЕНИЕ Б. О КОРРЕКТНОСТИ Е-ШАГА

В рамках Е-шага необходимо вычислить $\mathbb{E}[S_t | X_{t,o_t} = x_{t,o_t}]$, $\text{Cov}(S_t | X_{t,o_t} = x_{t,o_t})$. Вычисление $\mathbb{E}[S_t | X_{t,o_t} = x_{t,o_t}]$ реализуется так:

$$\begin{aligned}\mathbb{E}[S_t | X_{t,o_t} = x_{t,o_t}] &= \mathbb{E}\mathbb{E}[S_t | X_t] | X_{t,o_t} = x_{t,o_t}] \\ &= \mathbb{E}[\boldsymbol{\Gamma}^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} X_t | X_{t,o_t} = x_{t,o_t}] = \boldsymbol{\Gamma}^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} \mathbb{E}[X_t | X_{t,o_t} = x_{t,o_t}].\end{aligned}$$

Для нахождения $\text{Cov}(S_t | X_{t,o_t} = x_{t,o_t})$ может быть использовано равенство, называемое законом полной дисперсии (Law of total variance, [12]):

$$\text{Cov}(W) = \mathbb{E}[\text{Cov}(W | Y)] + \text{Cov}(\mathbb{E}[W | Y]).$$

Пусть $\mathcal{F}_Z = \sigma(Z)$ – сигма-алгебра, порожденная Z . Тогда все числовые характеристики W , включая условные моменты и ковариации относительно Y , можно обусловить на \mathcal{F}_Z . И, соответственно, из закона полной ковариации следует:

$$\text{Cov}(W | \mathcal{F}_Z) = \mathbb{E}[\text{Cov}(W | Y, Z) | \mathcal{F}_Z] + \text{Cov}(\mathbb{E}[W | Y, Z] | \mathcal{F}_Z).$$

Вектора W, Y, Z предполагаются комплексными гауссовскими (случай круговой симметрии), зависимость между Z и Y линейная, $\sigma(Z) \subseteq \sigma(Y)$, предполагается, что $Z = CY$, где C – булев селектор. Для линейной гауссовой модели значение Z не уменьшает условную ковариацию $W | Y$ и не влияет на условное математическое ожидание $\mathbb{E}[W | Y, Z]$, поскольку Z не содержит никакой новой информации, которой не было бы в Y :

$$\begin{aligned}\text{Cov}(W | Y, Z) &= \text{Cov}(W | Y), \\ \mathbb{E}[W | Y, Z] &= \mathbb{E}[W | Y].\end{aligned}$$

Получаем:

$$\text{Cov}(W | \mathcal{F}_Z) = \mathbb{E}[\text{Cov}(W | Y) | \mathcal{F}_Z] + \text{Cov}(\mathbb{E}[W | Y] | \mathcal{F}_Z).$$

Если $Z = z$, имеем:

$$\text{Cov}(W | Z = z) = \mathbb{E}[\text{Cov}(W | Y) | Z = z] + \text{Cov}(\mathbb{E}[W | Y] | Z = z).$$

В рамках исходной задачи по оцениванию DoA и ковариации сигналов, W соответствует величине S_t , Y соответствует величине X_t , Z соответствует величине X_{t,o_t} , z соответствует величине $x_{t,o_t}, t = 1, \dots, T$.

Получаем:

$$\text{Cov}(S_t | X_{t,o_t} = x_{t,o_t}) = \mathbb{E}[\text{Cov}(S_t | X_t) | X_{t,o_t} = x_{t,o_t}] + \text{Cov}(\mathbb{E}[S_t | X_t] | X_{t,o_t} = x_{t,o_t}).$$

Преобразуем первое слагаемое, учитывая тот факт, что ковариация не зависит от реализации X_{t,o_t} .

$$\mathbb{E}[\text{Cov}(S_t | X_t) | X_{t,o_t} = x_{t,o_t}] = \text{Cov}(S_t | X_t) = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} \mathbf{A} \boldsymbol{\Gamma}.$$

Теперь преобразуем второе слагаемое:

$$\begin{aligned}\text{Cov}(\mathbb{E}[S_t | X_t] | X_{t,o_t} = x_{t,o_t}) &= \text{Cov}(\boldsymbol{\Gamma}^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} X_t | X_{t,o_t} = x_{t,o_t}) \\ &= \boldsymbol{\Gamma}^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} \text{Cov}(X_t | X_{t,o_t} = x_{t,o_t}) \hat{\mathbf{R}}^{-1} \mathbf{A} \boldsymbol{\Gamma}.\end{aligned}$$

Таким образом:

$$\text{Cov}(S_t | X_{t,o_t} = x_{t,o_t}) = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} \mathbf{A} \boldsymbol{\Gamma} + \boldsymbol{\Gamma}^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} \text{Cov}(X_t | X_{t,o_t} = x_{t,o_t}) \hat{\mathbf{R}}^{-1} \mathbf{A} \boldsymbol{\Gamma}.$$

Для полных наблюдений $X_{t,o_t} = X_t$, и, соответственно:

$$\text{Cov}(\mathbb{E}[S_t | X_t] | X_{t,o_t} = x_{t,o_t}) = \text{Cov}(\mathbb{E}[S_t | X_t] | X_t = x_t) = 0.$$

О выводе смешанных вторых моментов $\mathbb{E}[X_t S_t^* | X_t = x_t]$ и $\mathbb{E}[X_t S_t^* | X_{t,o_t} = x_{t,o_t}]$:

$$\mathbb{E}[X_t S_t^* | X_t] = X_t \mathbb{E}[S_t^* | X_t] = X_t (\Gamma^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} X_t)^* = X_t X_t^* \hat{\mathbf{R}}^{-1} \mathbf{A} \Gamma,$$

$$\begin{aligned} \mathbb{E}[X_t S_t^* | X_{t,o_t} = x_{t,o_t}] &= \mathbb{E}[X_t \mathbb{E}[S_t^* | X_t] | X_{t,o_t} = x_{t,o_t}] = \mathbb{E}[X_t (\Gamma^* \mathbf{A}^* \hat{\mathbf{R}}^{-1} X_t)^* | X_{t,o_t} = x_{t,o_t}] \\ &= \mathbb{E}[X_t X_t^* \hat{\mathbf{R}}^{-1} \mathbf{A} \Gamma | X_{t,o_t} = x_{t,o_t}] = \mathbb{E}[X_t X_t^* | X_{t,o_t} = x_{t,o_t}] \hat{\mathbf{R}}^{-1} \mathbf{A} \Gamma \end{aligned}$$

В приведенных выше переходах используется закон полного условного математического ожидания и уже полученное выражение для $\mathbb{E}[S_t^* | X_t]$.

ПРИЛОЖЕНИЕ В. О РЕАЛИЗАЦИИ М-ШАГА

Если антенная решетка является равномерной и линейной, удобно искать оптимальный $u = \sin(\theta)$, а затем находить θ как $\arcsin(u)$. Соответственно минимизации подлежит функция

$$\mathcal{Q}_1(u | u^{(\tau-1)}) = -2 \operatorname{Re} \operatorname{Tr} (\tilde{\mathbf{A}}^*(u) \mathbf{\Lambda}^{-1} \mathbb{E}_{q(\Xi)}[XS^*]) + \operatorname{Tr} (\mathbf{\Lambda}^{-1} \tilde{\mathbf{A}}(u) \mathbb{E}_{q(\Xi)}[SS^*] \tilde{\mathbf{A}}^*(u)),$$

где

$$\tilde{\mathbf{A}}(u) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-2j\pi \frac{d}{\lambda} u_1} & e^{-2j\pi \frac{d}{\lambda} u_2} & \dots & e^{-2j\pi \frac{d}{\lambda} u_K} \\ \dots & \dots & \ddots & \vdots \\ e^{-2j\pi(L-1) \frac{d}{\lambda} u_1} & e^{-2j\pi(L-1) \frac{d}{\lambda} u_2} & \dots & e^{-2j\pi(L-1) \frac{d}{\lambda} u_K} \end{bmatrix}.$$

Для ускорения оптимизации можно оптимизировать не $\mathcal{Q}_1(u | u^{(\tau-1)})$, а суррогатную функцию $\mathcal{G}(u | u^{(\tau-1)})$, построенную так:

$$\mathcal{G}(u | u^{(\tau-1)}) = \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)}) + \nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})^T (u - u^{(\tau-1)}) + \frac{1}{2} (u - u^{(\tau-1)})^T \mathbf{H} (u - u^{(\tau-1)}),$$

где

$$\mathbf{H} = \begin{bmatrix} \max(|\nabla_1 \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})|, \varepsilon) & 0 & \dots & 0 \\ 0 & \max(|\nabla_2 \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})|, \varepsilon) & \dots & 0 \\ \dots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & \max(|\nabla_K \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})|, \varepsilon) \end{bmatrix},$$

причем $\nabla_i \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})$ – i -я компонента градиента $\mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})$. Диагональная аппроксимация \mathbf{H} предотвращает слишком маленькие или слишком большие шаги, сохраняя численную стабильность. Теперь эту суррогатную функцию надо минимизировать (или, эквивалентно, найти направление шага $\rho^{(\tau)} = u - u^{(\tau-1)}$):

$$\min_{u \in \mathbb{R}^K} \mathcal{G}(u | u^{(\tau-1)}).$$

Подставим $u = u^{(\tau-1)} + \rho$:

$$\mathcal{G}(u^{(\tau-1)} + \rho | u^{(\tau-1)}) = \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)}) + \nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})^T \rho + \frac{1}{2} \rho^T \mathbf{H} \rho.$$

Для нахождения минимума берем градиент по ρ :

$$\nabla_\rho \mathcal{G}(u^{(\tau-1)} + \rho | u^{(\tau-1)}) = \nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)}) + \mathbf{H} \rho,$$

при $\nabla_\rho \mathcal{G} = 0$ будет выполняться:

$$0 = \nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)}) + \mathbf{H} \rho,$$

решаем относительно $\rho^{(\tau)}$:

$$\rho^{(\tau)} = -\mathbf{H}^{-1} \nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)}).$$

Важно заметить, что производная по направлению $\nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})^T \rho^{(\tau)}$ принимает вид:

$$\nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})^T \rho^{(\tau)} = -\nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})^T \mathbf{H}^{-1} \nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)}),$$

и поскольку \mathbf{H} – положительно определенная матрица, производная по направлению может принимать лишь неположительные значения, а значит направление $\rho^{(\tau)}$ соответствует невозрастанию функции [15]. Далее используется backtracking line search в сочетании с проекцией на допустимое множество:

$$\tilde{u}^{(\tau)} = u^{(\tau-1)} + \alpha_m \rho^{(\tau)}, \quad u^{(\tau)} = \Pi_{\mathcal{U}}(\tilde{u}^{(\tau)}),$$

где $\Pi_{\mathcal{U}}$ – евклидова проекция на множество

$$\mathcal{U} = [-1, 1]^L = \{u \in \mathbb{R}^L \mid \|u\|_\infty \leq 1\},$$

задаваемая покомпонентным ограничением:

$$(\Pi_{\mathcal{U}}(x))_i = \min\{1, \max\{-1, x_i\}\}.$$

Параметр шага α_m выбирается как первое $\alpha = \alpha_0 \beta^m$, $m \in \mathbb{N}$, такое что

$$\mathcal{Q}_1(u^{(\tau)} \mid u^{(\tau-1)}) \leq \mathcal{Q}_1(u^{(\tau-1)} \mid u^{(\tau-1)}),$$

при фиксированных $\alpha_0 > 0$ и $\beta \in (0, 1)$.

В рамках программного комплекса реализуется мультистарт для поиска оптимального u на М-шаге. Указанный выше оптимизационный алгоритм запускается не только из точки $u^{(\tau-1)}$, но и из случайно выбранных допустимых точек, в том числе находящихся в окрестности $u^{(\tau-1)}$. Для выбранной начальной точки с индексом j , обновление оценки синусов угловых координат $\tilde{u}^{(j)}$ может быть принято, только если:

1. $\mathcal{Q}_1(\tilde{u}^{(j)} \mid u^{(\tau-1)}) > \mathcal{Q}_1(u^{(\tau-1)} \mid u^{(\tau-1)}),$
2. $\mathcal{Q}_1(\tilde{u}^{(j)} \mid u^{(\tau-1)}) > \mathcal{Q}_1(\min(\tilde{u}^{(i)} : 1 \leq i \leq j-1) \mid u^{(\tau-1)}),$ при $j > 1$.

Выбранный подход корректен ввиду следующих соображений:

1. Функция $\mathcal{Q}_1(u \mid u^{(\tau-1)})$ дифференцируема, поскольку $\tilde{\mathbf{A}}(u)$ состоит из экспонент вида $e^{j\pi m u_k}$, $k = 1, \dots, K$, $m \in \mathbb{R}$, т.е. гладких функций, умножение подобных матриц на фиксированные матрицы, взятие следа, взятие действительной части не нарушают гладкость.
2. Функция $\mathcal{Q}_1(u \mid u^{(\tau-1)})$ ограничена снизу, поскольку вектор u принадлежит компакту \mathcal{U} , матрицы, определяющие вид функции $\mathcal{Q}_1(u \mid u^{(\tau-1)})$ и отличные от $\mathbf{A}(u)$, являются фиксированными, след от произведения таких матриц, принимает лишь конечные значения.
3. Градиент $\nabla \mathcal{Q}_1(u \mid u^{(\tau-1)})$ не принимает нулевые значения для любых $u \in \mathcal{U}$, поскольку $\mathcal{Q}_1(u \mid u^{(\tau-1)})$ – гладкая функция, не сводящаяся к константе.

Оценим ковариацию сигналов $\mathbf{\Gamma}$:

$$\mathbf{\Gamma}^{(\tau)} = \arg \min_{\mathbf{\Gamma}} \mathcal{Q}_2(\mathbf{\Gamma} \mid \mathbf{\Gamma}^{(\tau-1)}) = \arg \min_{\mathbf{\Gamma}} \text{T} \left[\log |\mathbf{\Gamma}| + \text{Tr}(\mathbf{\Gamma}^{-1} \mathbb{E}_{q(\Xi)}[SS^*]) \right].$$

Определим точку, где производная данной функции принимает значение 0, и, таким образом, находим минимум функции:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{\Gamma}} \log(\text{Det}(\mathbf{\Gamma})) &= \mathbf{\Gamma}^{-1}, \\ \frac{\partial}{\partial \mathbf{\Gamma}} \text{Tr}(\mathbf{\Gamma}^{-1} \mathbb{E}_{q(\Xi)}[SS^*]) &= -\mathbf{\Gamma}^{-1} \mathbb{E}_{q(\Xi)}[SS^*] \mathbf{\Gamma}^{-1}, \\ \frac{\partial \mathcal{Q}_2(\mathbf{\Gamma})}{\partial \mathbf{\Gamma}} &= \mathbf{\Gamma}^{-1} - \mathbf{\Gamma}^{-1} \mathbb{E}_{q(\Xi)}[SS^*] \mathbf{\Gamma}^{-1}. \end{aligned}$$

Приравняем производную к нулю (функция по $\mathbf{\Gamma}$ выпукла):

$$\mathbf{O} = \mathbf{\Gamma}^{-1} - \mathbf{\Gamma}^{-1} \mathbb{E}_{q(\Xi)}[SS^*] \mathbf{\Gamma}^{-1} \Rightarrow \mathbf{\Gamma}^{(\tau)} = \mathbb{E}_{q(\Xi)}[SS^*].$$

Ввиду того, что М-шаг не максимизирует, а лишь улучшает УМО полного правдоподобия, представленный алгоритм является не точным ЕМ, а обобщенным (Generalized) ЕМ.

СПИСОК ЛИТЕРАТУРЫ

1. Krim H., Viberg M. Two decades of array signal processing research: the parametric approach. *IEEE Signal Processing Magazine*, 1996, Vol. 13, No. 4, pp. 67–94.
2. Godara L. C. Application of antenna arrays to mobile communications. Part II: Beam-forming and direction-of-arrival considerations. *Proceedings of the IEEE*, 1997, Vol. 85, No. 8, pp. 1195-1245.
3. Dempster A. P., Laird N. M., Rubin D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1977, Vol. 39, No. 1, pp. 1-38.
4. Meng X.-L., Rubin D. B. Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika*, 1993, Vol. 80, No. 2, pp. 267-278.
5. Fessler J. A., Hero A. O. Space-alternating generalized expectation-maximization algorithm. *IEEE Transactions on Signal Processing*, 1994, Vol. 42, No. 10, pp. 2664-2677.
6. Wu C. F. J. On the convergence properties of the EM algorithm. *The Annals of Statistics*, 1983, Vol. 11, No. 1, pp. 95–103.
7. Miller M. I., Fuhrmann D. R. Maximum-likelihood narrow-band direction finding and the EM algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1990, Vol. 38, No. 9, pp. 1560-1577.
8. Feder M., Weinstein E. Parameter estimation of superimposed signals using the EM algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1988, Vol. 36, No. 4, pp. 477-489.
9. Gong M.-Y., Lyu B. Alternating maximization and the EM algorithm in maximum-likelihood direction finding. *IEEE Trans. Veh. Technol.*, 2021, Vol. 70, No. 10, pp. 9634-9645.
10. Gong M.-Y., Lyu B. EM and SAGE Algorithms for DOA Estimation in the Presence of Unknown Uniform Noise. *Sensors*, 2023, Vol. 23, No. 10.
11. Louis T. A. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1982, Vol. 44, No. 2, pp. 226–233.
12. Ross S. M. *A first course in probability*. 8th ed. Upper Saddle River, N. J.: Prentice Hall, 2010. 640p., p. 348.
13. Little R. J. A., Rubin D. B. *Statistical analysis with missing data* 3rd ed. Hoboken, N. J.: Wiley, 2019, 462p.
14. Schreier P. J., Scharf L. L. *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge: Cambridge University Press, 2010. 309p., pp. 41-42.
15. Nocedal J., Wright S.J. *Numerical Optimization*. 2nd ed. New York: Springer, 2006. 664p., pp. 10-37.