

# ЕМ-алгоритм, стохастическая модель сигнала

## §1 Постановка проблемы

Предположим, имеется линейная антенная решетка, состоящая из  $L$  сенсоров, которая принимает сигналы, направленные из  $K$  источников, причем  $K < L$ . Этим источникам соответствуют угловые координаты (DoA)  $\theta$ , 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  $X$  состоит из наблюдаемой части  $X_o = \{X_{t,o_t}\}_{t=1}^T$  и ненаблюдаемой:  $X_m = \{X_{t,m_t}\}_{t=1}^T$ , причем  $o_t \cup m_t = \{1, \dots, T\}$ ,  $o_t \cap m_t = \emptyset, \forall t \in \{1, \dots, T\}$ . Предполагается, что  $\nexists o_t : |o_t| = 0$ , т.е. нет таких наблюдений, которые состоят лишь из ненаблюдаемой части.

Набор наблюдений  $X$  является результатом следующей модели наблюдений:

$$X = AS + N, \quad (1)$$

где

$$X = [X_1, \dots, X_T], S = [S_1, \dots, S_T], N = [N_1, \dots, N_T], \quad (2)$$

и  $X_t \in \mathbb{C}^{L \times T}, S_t \in \mathbb{C}^{K \times T}, N_t \in \mathbb{C}^{L \times T}$  — векторы-столбцы, соответствующие наблюдениям, источникам и шумам в момент времени  $t = 1, \dots, T$ ,  $A$  — матрица векторов направленности для равномерного линейного массива:

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

Сигналы, испускаемые источниками, шумы на сенсорах и наблюдения предполагаются стохастическими:  $S_t \sim \mathcal{CN}(\mathbf{O}_{K \times 1}, \mathbf{\Gamma})$ ,  $N_t \sim \mathcal{CN}(\mathbf{O}_{L \times 1}, \mathbf{\Lambda})$ ,  $X_t \sim \mathcal{CN}(\mathbf{O}_{L \times 1}, A\mathbf{\Gamma}A^* + \mathbf{\Lambda})$ ,  $t = 1, \dots, T$ . Матрицы  $\mathbf{\Gamma}$  и  $\mathbf{\Lambda}$  предполагаются диагональными, т.е. сигналы не коррелированы между собой, шумы также не коррелированы между собой, корреляция между сигналами и шумами также отсутствует. Составим ЕМ-алгоритм для двух случаев:

- Известный шум;
- Неизвестный шум.

## §2 Известный шум

Воспользуемся ЕМ-алгоритмом для того, чтобы определить значения параметров  $\Psi = (\theta, \Gamma)$ , пропущенные значения  $X_m = \{X_{t,m_t}\}_{t=1}^T$  и сигналы  $S$  рассматриваются как латентные переменные. Наблюдения  $X_t$ ,  $t = 1, \dots, T$  предполагаются независимыми и одинаково распределенными.

### Инициализация параметров

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

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

$$\nu \sim \mathcal{U}(-\pi; \pi); \quad (3)$$

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,  $a \bmod b = a - b \cdot \lfloor \frac{a}{b} \rfloor$ .

Начальную оценку ковариационной матрицы наблюдений  $\hat{R}^{(0)}$  получаем на основе доступных данных. Начальную ковариацию сигналов  $\Gamma^{(0)}$  получаем на основе метода наименьших квадратов с помощью  $\theta^{(0)}, \hat{R}^{(0)}$ .

### Е-шаг

Требуется найти математическое ожидание полного правдоподобия с учетом текущей оценки параметров и апостериорного совместного распределения пропущенных значений в наблюдениях  $X_m$  и сигналов  $S$

$$\mathbb{E}_{(X_m, S)|X_o=x_o, \Psi^{(\tau-1)}}[\log P(X, S)]. \quad (4)$$

Преобразуем выражение, обозначив через  $\mathcal{I}$  условную информацию  $X_o = x_o, \Psi^{(\tau-1)}$ , учитывая тот факт, что наблюдения являются независимыми:

$$\begin{aligned} & \mathbb{E}_{(X_m, S)|\mathcal{I}}[\log P(X, S)] = \\ & \mathbb{E}_{(X_m, S)|\mathcal{I}}[\log(P(X|S)P(S))] = \\ & \mathbb{E}_{(X_m, S)|\mathcal{I}}[\log P(X|S) + \log P(S)] = \\ & \mathbb{E}_{(X_m, S)|\mathcal{I}} \left[ \sum_{t=1}^T \log P(X_t|S_t) + \sum_{t=1}^T \log P(S_t) \right] = \\ & \left[ \sum_{t=1}^T \mathbb{E}_{(X_{t,m_t}, S_t)|\mathcal{I}_t} \log[P(X_t|S_t)] + \sum_{t=1}^T \mathbb{E}_{(X_{t,m_t}, S_t)|\mathcal{I}_t} [\log P(S_t)] \right] = \\ & - \left[ T \log |\Lambda| + \text{Tr}(\Lambda^{-1} \mathbb{E}_{(X_m, S)|\mathcal{I}}[XX^*]) - 2 \text{Tr}(\Lambda^{-1} A(\theta) \mathbb{E}_{(X_m, S)|\mathcal{I}}[XS^*]) \right. \\ & \left. + \text{Tr}(\Lambda^{-1} A(\theta) \mathbb{E}_{(X_m, S)|\mathcal{I}}[SS^*] A^*(\theta)) + T \log |\Gamma| + \text{Tr}(\Gamma^{-1} \mathbb{E}_{(X_m, S)|\mathcal{I}}[SS^*]) \right]. \end{aligned}$$

Для нахождения условных моментов, указанных в формуле выше, требуется найти апостериорное распределение скрытых переменных. Воспользуемся формулой произведения плотностей:

$$P(X_m, S|\mathcal{I}) = P(X_m|\mathcal{I}) \cdot P(S|X_m, \mathcal{I}). \quad (5)$$

Сначала найдем апостериорное распределение  $P(X_m|\mathcal{I})$ , причем ввиду того, что индексы, соответствующие пропущенным значениям в наблюдениях, могут отличаться в зависимости от  $t$ , будем находить апостериорное распределение для каждого  $X_{t,m_t}$ . 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:

$$\hat{\mathbf{R}} = \begin{pmatrix} \hat{\mathbf{R}}_{o_t, o_t} & \hat{\mathbf{R}}_{o_t, m_t} \\ \hat{\mathbf{R}}_{m_t, o_t} & \hat{\mathbf{R}}_{m_t, m_t} \end{pmatrix}, \quad (6)$$

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,  $P(X_{t,m_t}|X_{t,o_t} = x_{t,o_t}, \Psi^{(\tau-1)}), t = 1, \dots, T, |m_t| > 0$ . Обозначим через  $\mathcal{I}_t$  условную информацию  $X_{t,o_t} = x_{t,o_t}, \Psi^{(\tau-1)}$ . Параметры апостериорного распределения  $P(X_{t,m_t}|\mathcal{I}_t), t = 1, \dots, T$  на итерации  $\tau$  можно найти следующим образом:

$$\begin{cases} \mu_{X_{t,m_t}|\mathcal{I}_t} = \hat{\mathbf{R}}_{m_t, o_t} (\hat{\mathbf{R}}_{o_t, o_t})^{-1} \cdot x_{t,o_t}, \\ \Sigma_{X_{t,m_t}|\mathcal{I}_t} = \hat{\mathbf{R}}_{m_t, m_t} - \hat{\mathbf{R}}_{m_t, o_t} (\hat{\mathbf{R}}_{o_t, o_t})^{-1} \hat{\mathbf{R}}_{o_t, m_t}, \end{cases} \quad (7)$$

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  $X_t$  оценим условную ковариационную матрицу  $\tilde{\Sigma}_{X_t} = \mathbb{E}_{X_{m_t}|\mathcal{I}_t}[X_t X_t^*]$ :

$$\begin{aligned} \tilde{\Sigma}_{X_t} &= \begin{pmatrix} \mu_{X_{t,o_t}|\mathcal{I}_t} \cdot \mu_{X_{t,o_t}|\mathcal{I}_t}^* & \mu_{X_{t,o_t}|\mathcal{I}_t} \cdot \mu_{X_{t,m_t}|\mathcal{I}_t}^* \\ \mu_{X_{t,m_t}|\mathcal{I}_t} \cdot \mu_{X_{t,o_t}|\mathcal{I}_t}^* & \mu_{X_{t,m_t}|\mathcal{I}_t} \cdot \mu_{X_{t,m_t}|\mathcal{I}_t}^* + \Sigma_{X_{t,m_t}|\mathcal{I}_t} \end{pmatrix} = \\ &= \mathbb{E}[X_t|\mathcal{I}_t] \cdot \mathbb{E}[X_t|\mathcal{I}_t]^* + \begin{pmatrix} \mathbf{O}_{o_t, o_t} & \mathbf{O}_{o_t, m_t} \\ \mathbf{O}_{m_t, o_t} & \Sigma_{X_{t,m_t}|\mathcal{I}_t} \end{pmatrix}, \end{aligned}$$

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  $o_t, m_t$ .

Оценим условную ковариационную матрицу наблюдений  $\tilde{\Sigma}_X = \mathbb{E}_{(X_m, S)|\mathcal{I}}[XX^*]$ :

$$\tilde{\Sigma}_X = \frac{1}{T} \sum_{t=1}^T \tilde{\Sigma}_{X_t}. \quad (8)$$

Параметры апостериорного распределения  $P(S_t|\mathcal{I}_t, X_{t,m_t}), t = 1, \dots, T$  можно найти исходя из следующих формул:

$$\begin{cases} \mu_{S_t|\mathcal{I}_t, X_{t,m_t}} = \boldsymbol{\Gamma} \mathbf{A}^* (\tilde{\Sigma}_{X_t})^{-1} \mathbb{E}[X_t|\mathcal{I}_t], \\ \Sigma_{S_t|\mathcal{I}_t, X_{t,m_t}} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma} \mathbf{A}^* (\tilde{\Sigma}_{X_t})^{-1} \mathbf{A} \boldsymbol{\Gamma}, \end{cases} \quad (9)$$

где  $\mathbf{A} = \mathbf{A}(\theta^{(\tau-1)})$ ,  $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^{(\tau-1)}$ .

Оценим ковариационную матрицу сигналов с учетом текущей оценки параметров и доступных наблюдений  $\tilde{\Sigma}_S = \mathbb{E}_{(X_m, S)|\mathcal{I}}[SS^*]$ .

$$\tilde{\Sigma}_S = \frac{1}{T} \left[ \sum_{t=1}^T (\Sigma_{S_t|\mathcal{I}_t, X_{t,m_t}} + \mu_{S_t|\mathcal{I}_t, X_{t,m_t}} \cdot \mu_{S_t|\mathcal{I}_t, X_{t,m_t}}^H) \right]. \quad (10)$$

Остается оценить кросс-ковариацию  $\tilde{\Sigma}_{XS} = \mathbb{E}_{(X_m, S)|\mathcal{I}}[XS^*]$ :

$$\begin{aligned} \mathbb{E}_{(X_{t,m_t}, S_t)|\mathcal{I}_t}[X_t S_t^*] &= \mathbb{E}_{(X_{t,m_t}, S_t)|\mathcal{I}_t}[X_t X_t^*] (\hat{\mathbf{R}}^{(\tau-1)})^{-1} \mathbf{A}(\theta^{(\tau-1)}) \boldsymbol{\Gamma}^{(\tau-1)} \\ &= \tilde{\Sigma}_{X_t} (\hat{\mathbf{R}}^{(\tau-1)})^{-1} \mathbf{A}(\theta^{(\tau-1)}) \boldsymbol{\Gamma}^{(\tau-1)}. \end{aligned} \quad (11)$$

$$\begin{aligned}
\tilde{\Sigma}_{XS} &= \mathbb{E}_{(X_m, S)|\mathcal{I}}[XS^*] = \frac{1}{T} \sum_{t=1}^T \left( \tilde{\Sigma}_{X_t} (\hat{R}^{(\tau-1)})^{-1} A(\theta^{(\tau-1)}) \Gamma^{(\tau-1)} \right) = \\
&= \tilde{\Sigma}_X (\hat{R}^{(\tau-1)})^{-1} A(\theta^{(\tau-1)}) \Gamma^{(\tau-1)}
\end{aligned} \tag{12}$$

## М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\begin{aligned}\Psi^{(\tau)} &= \arg \max_{\Psi} \mathbb{E}_{(X_m, S) | \mathcal{I}} [\log P(X, S)] = \\ &\arg \min_{\Psi} \left[ T \log |\Lambda| + \operatorname{Tr} (\Lambda^{-1} \mathbb{E}_{(X_m, S) | \mathcal{I}} [XX^*]) - 2 \operatorname{Tr} (\Lambda^{-1} A(\theta) \mathbb{E}_{(X_m, S) | \mathcal{I}} [XS^*]) \right. \\ &\quad \left. + \operatorname{Tr} (\Lambda^{-1} A(\theta) \mathbb{E}_{(X_m, S) | \mathcal{I}} [SS^*] A^*(\theta)) + T \log |\Gamma| + \operatorname{Tr} (\Gamma^{-1} \mathbb{E}_{(X_m, S) | \mathcal{I}} [SS^*]) \right].\end{aligned}$$

Оценим угловые координаты источников  $\theta$ :

$$\begin{aligned}\theta^{(\tau)} &= \arg \min_{\theta} \mathcal{J}(\theta) = \arg \min_{\theta} \left[ -2 \operatorname{Tr} (\Lambda^{-1} A(\theta) \Sigma_{XS}) \right. \\ &\quad \left. + \operatorname{Tr} (\Lambda^{-1} A(\theta) \Sigma_S A^*(\theta)) \right] = \\ &\arg \min_{\theta} \|\Lambda^{-1/2} (\Sigma_{XS} - A(\theta) \Sigma_S)\|_F^2.\end{aligned}$$

Для решения этой оптимизационной задачи можно использовать сочетание метода последовательного квадратичного программирования и линейного поиска.

Оценим ковариацию сигналов  $\Gamma$ :

$$\Gamma^{(\tau)} = \arg \min_{\Gamma} \mathcal{K}(\Gamma | \Gamma^{(\tau-1)}) = T \left[ \log |\Gamma| + \operatorname{Tr} (\Gamma^{-1} \mathbb{E}_{(X_m, S) | \mathcal{I}} [SS^*]) \right].$$

Определим точку, где производная данной функции принимает значение 0, и, таким образом, находим минимум функции, при этом обозначим через  $M$  величину  $\mathbb{E}_{(X_m, S) | \mathcal{I}} [S_t S_t^*]$ :

$$\begin{aligned}\frac{\partial}{\partial \Gamma} \log(\operatorname{Det}(\Gamma)) &= \Gamma^{-1}, \\ \frac{\partial}{\partial \Gamma} \operatorname{Tr} (\Gamma^{-1} M) &= -\Gamma^{-1} M \Gamma^{-1}, \\ \frac{\partial \mathcal{K}(\Gamma)}{\partial \Gamma} &= \Gamma^{-1} - \Gamma^{-1} M \Gamma^{-1}.\end{aligned}$$

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

$$O = \Gamma^{-1} - \Gamma^{-1} M \Gamma^{-1} \Rightarrow M = T \Gamma \Rightarrow \Gamma^{(\tau)} = \tilde{\Sigma}_S.$$

Предполагая, что сигналы некоррелированы, будем использовать лишь диагональное приближение матрицы, приравняв элементы вне главной диагонали к нулю:

$$\Gamma^{(\tau)} = \mathcal{D} [\tilde{\Sigma}_S]. \tag{13}$$

Обновляем оценку ковариации наблюдений с учетом полученных оценок параметров:

$$\hat{R}^{(\tau)} = A(\theta^{(\tau)}) \Gamma A^*(\theta^{(\tau)}) + \Lambda. \tag{14}$$

## Список источников

1. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B (Methodol.) 1977, 39, 1–38.