

ЕМ-алгоритм, стохастическая модель сигнала

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

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

Оценим угловые координаты источников θ :

$$\begin{aligned} \theta^{(\tau)} &= \arg \min_{\theta} \mathcal{Q}_1(\theta | \theta^{(\tau-1)}) = \arg \min_{\theta} \left[-2 \text{Tr} (\mathbf{\Lambda}^{-1} \mathbf{A}(\theta) \mathbb{E}_{(X_m, S)|\mathcal{I}} [X S^*]) \right. \\ & \left. + \text{Tr} (\mathbf{\Lambda}^{-1} \mathbf{A}(\theta) \mathbb{E}_{(X_m, S)|\mathcal{I}} [S S^*] \mathbf{A}^*(\theta)) \right] = \\ & \arg \min_{\theta} \|\mathbf{\Lambda}^{-1/2} (\mathbb{E}_{(X_m, S)|\mathcal{I}} [X S^*] - \mathbf{A}(\theta) \mathbb{E}_{(X_m, S)|\mathcal{I}} [S S^*])\|_F^2. \end{aligned}$$

На практике, если \mathbf{A} соответствует антенной решетке типа ULA, удобно искать оптимальный $u = \sin(\theta)$, а затем находить θ как $\arcsin(u)$. Для удобства введем новые обозначения: $\mathbf{C} = \mathbb{E}_{(X_m, S)|\mathcal{I}} [X S^*]$, $\mathbf{D} = \mathbb{E}_{(X_m, S)|\mathcal{I}} [S S^*]$. Соответственно минимизации подлежит функция

$$\mathcal{Q}_1(u | u^{(\tau-1)}) = \|\mathbf{\Lambda}^{-1/2} (\mathbf{C} - \tilde{\mathbf{A}}(u) \mathbf{D})\|_F^2, \quad (1)$$

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:

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

где

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

где $\nabla_i \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})$ — i -я компонента градиента $\mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})$. Теперь эту суррогатную функцию надо минимизировать (или, эквивалентно, найти направление шага $d^{(\tau)} = u - u^{(\tau-1)}$):

$$\min_u \mathcal{G}(u | u^{(\tau-1)}). \quad (3)$$

Подставим $u = u^{(\tau-1)} + d$:

$$\mathcal{G}(u^{(\tau-1)} + d | u^{(\tau-1)}) = \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)}) + \nabla \mathcal{Q}_1(u^{(\tau-1)} | u^{(\tau-1)})^T d + \frac{1}{2} d^T \mathbf{H} d. \quad (4)$$

Для нахождения минимума берем градиент по d :

$$\nabla_d \mathcal{G}(u^{(\tau-1)} + d|u^{(\tau-1)}) = \nabla \mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)}) + \mathbf{H}d, \quad (5)$$

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

$$0 = \nabla \mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)}) + \mathbf{H}d, \quad (6)$$

решаем относительно $d^{(\tau)}$:

$$d^{(\tau)} = -\mathbf{H}^{-1} \nabla \mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)}). \quad (7)$$

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

$$\nabla \mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)})^T d^{(\tau)} = -\nabla \mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)})^T \mathbf{H}^{-1} \nabla \mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)}), \quad (8)$$

и поскольку \mathbf{H} – положительно определенная матрица, производная по направлению может принимать лишь неположительные значения, а значит направление $d^{(\tau)}$ соответствует невозрастанию функции. Далее используется backtracking line search для гарантии невозрастания \mathcal{Q}_1 :

$$u^{(\tau)} = u^{(\tau-1)} + \alpha_m d^{(\tau)}, \quad (9)$$

где α_m – первое $\alpha = \alpha_0 \beta^m, m \in \mathbb{N}$, такое что $\mathcal{Q}_1(u^{(\tau)}|u^{(\tau-1)}) \leq \mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)})$ для фиксированных $\alpha_0 > 0, \beta \in (0; 1)$.

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

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

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

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

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

$$\mathbf{O} = \mathbf{\Gamma}^{-1} - \mathbf{\Gamma}^{-1} \mathbb{E}_{(X_m, S)|\mathcal{I}}[SS^*] \mathbf{\Gamma}^{-1} \Rightarrow \mathbb{E}_{(X_m, S)|\mathcal{I}}[SS^*] = \mathbf{\Gamma} \Rightarrow \mathbf{\Gamma}^{(\tau)} = \mathbb{E}_{(X_m, S)|\mathcal{I}}[SS^*].$$

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

$$\mathbf{\Gamma}^{(\tau)} = \mathcal{D} \left[\mathbb{E}_{(X_m, S)|\mathcal{I}}[SS^*] \right]. \quad (10)$$

Обновляем оценку ковариации наблюдений с учетом полученных оценок параметров:

$$\hat{\mathbf{R}}^{(\tau)} = \mathbf{A}(\theta^{(\tau)}) \mathbf{\Gamma} \mathbf{A}^*(\theta^{(\tau)}) + \mathbf{\Lambda}. \quad (11)$$