ЕМ-алгоритм, стохастическая модель сигнала

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\Upsilon^{(\tau)} = \arg\max_{\Upsilon} \mathbb{E}_{(X_m,S)|\mathcal{I}}[\log P(X,S)] =$$
$$\arg\min_{\Upsilon} \mathrm{T}\Big[\log|\mathbf{\Lambda}| + \mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\mathbb{E}_{(X_m,S)|\mathcal{I}}[XX^*]\right) - 2\,\mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\mathrm{A}(\theta)\mathbb{E}_{(X_m,S)|\mathcal{I}}[XS^*]\right)$$
$$+\mathrm{Tr}\big(\mathbf{\Lambda}^{-1}\mathrm{A}(\theta)\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^*]\mathrm{A}^*(\theta)\big) + \log|\mathbf{\Gamma}| + \mathrm{Tr}\big(\mathbf{\Gamma}^{-1}\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^*]\big)\Big].$$

Оценим угловые координаты источников θ:

$$\theta^{(\tau)} = \arg\min_{\theta} \mathcal{Q}_1(\theta|\theta^{(\tau-1)}) = \arg\min_{\theta}\Big[-2\,\mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\mathrm{A}(\theta)\mathbb{E}_{(X_m,S)|\mathcal{I}}[XS^*]\right)$$
$$+\mathrm{Tr}\big(\mathbf{\Lambda}^{-1}\mathrm{A}(\theta)\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^*]\mathrm{A}^*(\theta)\big)\Big] =$$
$$\arg\min_{\theta} ||\mathbf{\Lambda}^{-1/2}(\mathbb{E}_{(X_m,S)|\mathcal{I}}[XS^*] - \mathrm{A}(\theta)\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^*])||_F^2.$$

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

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

где

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

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

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

где

$$\mathbf{H} = \begin{bmatrix} \max(|\boldsymbol{\nabla}_1\mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)})|,\varepsilon) & 0 & \dots & 0 \\ 0 & \max(|\boldsymbol{\nabla}_2\mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)})|,\varepsilon) & \dots & 0 \\ \dots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & \max(|\boldsymbol{\nabla}_{\mathrm{K}}\mathcal{Q}_1(u^{(\tau-1)}|u^{(\tau-1)})|,\varepsilon) \end{bmatrix}.$$

Теперь эту суррогатную функцию надо минимизировать (или, эквивалентно, найти направление шага $d^{(\tau)} = u - u^{(\tau-1)}$):

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

Подставим $u = u^{(\tau-1)} + d$:

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

Для нахождения минимума берем градиент по d:

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

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

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

решаем относительно $d^{(\tau)}$:

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

Далее используется backtracking line search для гарантии невозрастания $\mathcal{Q}_1$.
Оценим ковариацию сигналов $\mathbf{\Gamma}$:

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

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

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

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 некоррелированны, будем использовать лишь диагональное приближение матрицы, приравняв элементы вне главной диагонали к нулю:

$$\mathbf{\Gamma}^{(\tau)} = \mathcal{D}\Big[\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^*]\Big]. \tag{8}$$

Обновляем оценку ковариации наблюдений с учетом полученных оценок параметров:

$$\hat{\mathrm{R}}^{(\tau)} = \mathrm{A}(\theta^{(\tau)}) \mathbf{\Gamma} \mathrm{A}^*(\theta^{(\tau)}) + \mathbf{\Lambda}. \tag{9}$$