EM-алгоритм для оценки направления прибытия сигнала

Введем некоторые условные обозначения:

- θ — вектор направлений прибытия сигнала (DOA);
- τ — итерация EM-алгоритма, начальная оценка параметров θ;
- t — момент времени (а заодно и номер кадра (snapshot));
- L — число датчиков;
- M — число источников (источники разделяют общую длину центральной волны χ);
- G — число независимых кадров/снимков (snapshot), сделанных в разные моменты времени;
- S — набор сигналов (случайная величина), испускаемых источниками в моменты времени $t = \overline{1, G}$, S_t соответствует сигналу в момент времени t;
- s — набор сигналов (реализация), испускаемых источниками в моменты времени $t = \overline{1, G}$, s_t соответствует сигналу в момент времени t;
- N — набор шумов (случайная величина), связанных с датчиками в моменты времени $t = \overline{1, G}$, N_t соответствует шуму в момент времени t;
- n — набор шумов (реализация), связанных с датчиками в моменты времени $t = \overline{1, G}$, n_t соответствует шуму в момент времени t;
- X — набор сигналов (случайная величина), полученных датчиками в моменты времени $t = \overline{1, G}$, X_t соответствует сигналу в момент времени t;
- x — набор сигналов (реализация), полученных датчиками в моменты времени $t = \overline{1, G}$, x_t соответствует сигналу в момент времени t;
- X_o — наблюдаемая часть (случайная величина) X, $X_{o,t}$ соответствует сигналу в момент времени t;
- x_o — наблюдаемая часть (реализация) X, $x_{o,t}$ соответствует сигналу в момент времени t;
- X_m — ненаблюдаемая часть (случайная величина) X, $X_{m,t}$ соответствует сигналу в момент времени t;
- x_m — ненаблюдаемая часть (реализация) X, $x_{m,t}$ соответствует сигналу в момент времени t;
- Ψ — оцениваемые параметры (θ, S);
- $O_{D_1 \times D_2}$ — нулевая матрица размера $D_1 \times D_2$;
- Полученный сигнал (итоговый сигнал, получаемый массивом датчиков):

$$X_t = A(\theta) S_t + N_t, \tag{1}$$

где $S_t \sim CN(\mathbf{O}_{M\times 1}, \mathbf{P}), t = \overline{1, G}$, $N_t \sim CN(\mathbf{O}_{L\times 1}, \mathbf{Q}), t = \overline{1, G}$, S_t имеет размер $M \times 1$, N_t имеет размер $L \times 1$, $\theta = [\theta_1, ..., \theta_M]^T$ — вектор направлений прибытия сигнала, $A(\theta)$ (далее

– A) представляет собой матрицу управляющих векторов размера $L \times M$, $\mathbf{P}$ и $\mathbf{Q}$ предполагаются диагольными. Рассматривается случай, когда массив антенн является линейным и равномерным.

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

- $L_{1,t}$ — число исправных сенсоров в момент времени t;
- $L_{2,t}$ — число неисправных сенсоров в момент времени t;
- $A_{o,t}$ — матрица, образованная теми строками матрицы A, которые соответствуют работающим сенсорам в момент времени t;
- $A_{m,t}$ — матрица, образованная теми строками матрицы A, которые соответствуют неисправным сенсорам в момент времени t;
- $Q_{m,t}$ — ковариационная матрица шума на неисправных сенсорах в момент времени t;
- $Q_{o,t}$ — ковариационная матрица шума на исправных сенсорах в момент времени t.

Воспользуемся EM-алгоритмом для того, чтобы определить значения параметров θ, значения сигналов $S_t, t = \overline{1,G}$ рассматриваются как латентные переменные. Пусть X, S и N набор итоговых сигналов полученных L датчиками за моменты времени $t = \overline{1,G}$ и набор выпущенных M источниками сигналов и набор шумов за моменты времени $t = \overline{1,G}$, соответственно. X, S и N представляют из себя матрицы размеров $G \times L$, $G \times M$ и $G \times L$ соответственно.

E-шаг

Требуется найти условное математическое ожидание с учетом апостериорного распределения ненаблюдаемых/пропущенных принятых сигналов и текущей оценки параметров

$$E_{X_m|X_o,\Psi^{(\tau)}}[\log P(X_o, X_m)] \tag{2}$$

Сначала найдем апостериорное распределение $P(X_m|X_o = x_o, \Psi)$, воспользуемся формулой Байеса:

$$P(X_m|X_o = x_o, \Psi) = \frac{P(X_o, X_m|\Psi)}{P(X_o|\Psi)} = \frac{P(X|\Psi)}{P(X_o|\Psi)} \tag{3}$$

$$X_t = AS_t + N_t$$
$$X_t \sim CN(\mathbf{O}_{L\times 1}, A\mathbf{P}A^H + \mathbf{Q})$$

$$P(X|\Psi) = \prod_{t=1}^{G} \frac{1}{\pi^L \operatorname{Det}(A\mathbf{P}A^H + \mathbf{Q})} e^{-X_t^H (A\mathbf{P}A^H + \mathbf{Q})^{-1} X_t}, \tag{4}$$

$$P(X_o|\Psi) = \prod_{t=1}^{G} \frac{1}{\pi^L \operatorname{Det}(A_{o,t}\mathbf{P}A_{o,t}^H + \mathbf{Q}_{o,t})} e^{-X_{o,t}^H (A_{o,t}\mathbf{P}A_{o,t}^H + \mathbf{Q}_{o,t})^{-1} X_{o,t}}, \tag{5}$$

Параметры апостериорного распределения $P(X_m|X_o = x_o, \Psi)$ можно найти исходя из следующих формул:

$$\begin{cases} m_{x_m|x_o} = m_{x_m} + K_{x_m,x_o} K_{x_o,x_o}^{-1} \cdot (x_o - m_{x_o}) \\ K_{x_m|x_o} = K_{x_m,x_m} - K_{x_m,x_o} K_{x_o,x_o}^{-1} K_{x_o,x_m}, \end{cases} \tag{6}$$

В рамках данной задачи: $K_{x_o,x_o} = A_{o,t}\mathbf{P}A_{o,t}^H + \mathbf{Q}_{o,t}$, $K_{x_o,x_m} = A_{o,t}\mathbf{P}A_{m,t}^H$, $K_{x_m,x_o} = A_{m,t}\mathbf{P}A_{o,t}^H$, $K_{x_m,x_m} = A_{m,t}\mathbf{P}A_{m,t}^H + \mathbf{Q}_{m,t}$

$$\begin{cases} m_{x_m|x_o} = A_{m,t}\mathbf{P}A_{o,t}^H (A_{o,t}\mathbf{P}A_{o,t}^H + \mathbf{Q}_{o,t})^{-1} \cdot x_o \\ K_{x_m|x_o} = (A_{m,t}\mathbf{P}A_{m,t}^H + \mathbf{Q}_{m,t}) - A_{m,t}\mathbf{P}A_{o,t}^H (A_{o,t}\mathbf{P}A_{o,t}^H + \mathbf{Q}_{o,t})^{-1} A_{o,t}\mathbf{P}A_{m,t}^H, \end{cases} \tag{7}$$

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\begin{gathered} \Psi^{(\tau+1)} = \underset{\Psi}{\operatorname{argmax}}\, E_{X_m|X_o=x_o,\Psi^{(\tau)}}[\log P(X_o, X_m)] = \\ \underset{\Psi}{\operatorname{argmax}}\, E_{X_m|X_o=x_o,\Psi^{(\tau)}} \left[\sum_{t=1}^{G} \log P(X_{o,t}, X_{m,t})\right] = \\ \underset{\Psi}{\operatorname{argmax}}\, E_{X_m|X_o=x_o,\Psi^{(\tau)}} \left[\sum_{t=1}^{G} \log\left(\frac{1}{\pi^L \operatorname{Det}(\Sigma)} e^{-(X_t-\mu_{X_t})^H \Sigma^{-1}(X_t-\mu_{X_t})}\right)\right] \end{gathered} \tag{8}$$

$$\begin{gathered} \log\left(\frac{1}{\pi^L \operatorname{Det}(\Sigma)} e^{-(X_t-\mu_{X_t})^H \Sigma^{-1}(X_t-\mu_{X_t})}\right) = \\ -L\log(\pi) - \log(\operatorname{Det}(\Sigma)) - (X_t-\mu_{X_t})^H \Sigma^{-1}(X_t-\mu_{X_t}) = \end{gathered} \tag{9}$$

$$\begin{aligned} & E_{X_m|X_o=x_o,\Psi^{(\tau)}}\left[\begin{bmatrix} X_{m,t} \\ X_{o,t} \end{bmatrix}^H \begin{bmatrix} (\Sigma^{-1})_{x_m,x_m} & (\Sigma^{-1})_{x_m,x_o} \\ (\Sigma^{-1})_{x_o,x_m} & (\Sigma^{-1})_{x_o,x_o} \end{bmatrix} \begin{bmatrix} X_{m,t} \\ X_{o,t} \end{bmatrix} \right] \\ = \quad & E_{X_m|X_o=x_o,\Psi^{(\tau)}}\left[(X_{m,t})^H (\Sigma^{-1})_{x_m,x_m} (X_{m,t})\right] \\ & + 2E_{X_m|X_o=x_o,\Psi^{(\tau)}}\left[(X_{o,t})^H (\Sigma^{-1})_{x_o,x_m} (X_{m,t})\right] \\ & + \ E_{X_m|X_o=x_o,\Psi^{(\tau)}}\left[(X_{o,t})^H (\Sigma^{-1})_{x_o,x_o} (X_{o,t})\right] \\ = \quad & E_{X_m|X_o=x_o,\Psi^{(\tau)}}\left[(X_{m,t})^H (\Sigma^{-1})_{x_m,x_m} (X_{o,t})\right] \\ & + 2(x_{o,t})^H (\Sigma^{-1})_{x_o,x_m} E_{X_m|X_o=x_o,\Psi^{(\tau)}}\left[X_{m,t}\right] \\ & + \ x_{o,t}^H (\Sigma^{-1})_{x_o,x_o} x_{o,t} \end{aligned}$$