ECM-алгоритм, стохастическая модель сигнала

Введем некоторые условные обозначения:

- θ — вектор направлений прибытия сигнала (DoA);
- τ — итерация ECM-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, представляет собой матрицу размера $G \times M$;
- 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, представляет собой матрицу размера $G \times L$;
- 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, представляет собой матрицу размера $G \times L$;
- 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;
- $\mathbf{O}_{D_1 \times D_2}$ — 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 $S_t \sim CN(\mathbf{O}_{M \times 1}, \mathbf{P}), t = \overline{1, G}$, $N_t \sim CN(\mathbf{O}_{L \times 1}, \mathbf{\Lambda})$, $\theta = [\theta_1, ..., \theta_M]^T$ — вектор направлений прибытия сигнала, $A(\theta)$ (далее – A) представляет собой матрицу управляющих векторов размера $L \times M$, $\mathbf{P}$ и $\mathbf{\Lambda}$ предполагаются диагональными. Рассматривается случай, когда массив антенн является линейным и равномерным.

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

- $L_{o,t}$ — число исправных сенсоров в момент времени t;
- $L_{m,t}$ — 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;
- $\mathbf{\Lambda}_{m,t}$ — ковариационная матрица шума на неисправных сенсорах в момент времени t;
- $\mathbf{\Lambda}_{o,t}$ — ковариационная матрица шума на исправных сенсорах в момент времени t.

Рассмотрим 2 случая:

- Известный шум
- Неизвестный шум

Известный шум

Воспользуемся ECM-алгоритмом для того, чтобы определить значения параметров $\Psi = (\theta, \mathbf{P})$, пропущенные значения $X_m = \{X_{m,t}\}_{t=1}^{G}$ рассматриваются как латентные переменные.

E-шаг

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

$$\mathbb{E}_{X_m|X_o=x_o,\Psi^{(\tau)}}[\log P(X_o, X_m)] \tag{2}$$

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

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

$$X_t = AS_t + N_t$$
$$X_t \sim CN(\mathbf{O}_{L\times 1}, A\mathbf{P}A^H + \mathbf{\Lambda})$$
$$X_{o,t} \sim CN(\mathbf{O}_{L\times 1}, A_{o,t}\mathbf{P}A_{o,t}^H + \mathbf{\Lambda_{o,t}})$$

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

$$P(X_o|\Psi) = \prod_{t=1}^{G} \frac{1}{\pi^{L_{o,t}}\,\mathrm{Det}(\mathbf{\Lambda}_{o,t})} e^{-(X_{o,t})^H(\mathbf{\Lambda}_{o,t})^{-1}(X_{o,t})}, \tag{5}$$

Параметры апостериорного распределения $P(X_{m,t}|X_{o,t} = x_{o,t}, \Psi)$ можно найти исходя из следующих формул:

$$\begin{cases} m_{x_{m,t}|x_{o,t}} = m_{x_{m,t}} + K_{x_{m,t},x_{o,t}} K^{-1}_{x_{o,t},x_{o,t}} \cdot (x_{o,t} - m_{x_{o,t}}) \\ K_{x_{m,t}|x_{o,t}} = K_{x_{m,t},x_{m,t}} - K_{x_{m,t},x_{o,t}} K^{-1}_{x_{o,t},x_{o,t}} K_{x_{o,t},x_{m,t}}. \end{cases} \tag{6}$$

В рамках данной задачи:

$$\begin{cases} K_{x_{o,t},x_{o,t}} = A_{o,t}\mathbf{P}A^H_{o,t} + \mathbf{\Lambda}_{o,t} \\ K_{x_{o,t},x_{m,t}} = A_{o,t}\mathbf{P}A^H_{m,t} \\ K_{x_{m,t},x_{o,t}} = A_{m,t}\mathbf{P}A^H_{o,t} \\ K_{x_{m,t},x_{m,t}} = A_{m,t}\mathbf{P}A^H_{m,t} + \mathbf{\Lambda}_{m,t} \\ m_{x_{o,t}} = \mathbf{O}_{L_{o,t}\times 1} \\ m_{x_{m,t}} = \mathbf{O}_{L_{m,t}\times 1} \end{cases} \tag{7}$$

$$\begin{cases} m_{x_{m,t}|x_{o,t}} = A_{m,t}\mathbf{P}A^H_{o,t}(A_{o,t}\mathbf{P}A^H_{o,t} + \mathbf{\Lambda}_{o,t})^{-1} \cdot x_{o,t} \\ K_{x_{m,t}|x_{o,t}} = A_{m,t}\mathbf{P}A^H_{m,t} + \mathbf{\Lambda}_{m,t} - A_{m,t}\mathbf{P}A^H_{o,t}(A_{o,t}\mathbf{P}A^H_{o,t} + \mathbf{\Lambda}_{o,t})^{-1}A_{o,t}\mathbf{P}A^H_{m,t}, \end{cases} \tag{8}$$

Оцениваем пропущенные значения для каждого наблюдения через условное математическое ожидание: $\hat{x}_{m,t} = m_{x_{m,t}|x_{o,t}}$. Вернемся к ранее рассмотренному условному математическому ожиданию:

$$\mathbb{E}_{X_m|X_o=x_o,\Psi^{(\tau)}}[\log P(X_o, X_m)].$$

Его следует максимизировать, мы можем перейти от логарифма произведения к сумме логарифмов. Определим, как будет выглядеть это УМО для произвольно выбранного элемента выборки X_t:

$$\mathbb{E}_{X_m|X_o=x_o,\Psi^{(\tau)}}[\log P(X_o, X_m)] =$$
$$\sum_{t=1}^{G} \mathbb{E}_{X_{m,t}|X_{o,t}=x_{o,t},\Psi^{(\tau)}}[\log P(X_{o,t}, X_{m,t})]$$

$$\mathbb{E}_{X_{m,t}|X_{o,t}=x_{o,t},\Psi^{(\tau-1)}}[\log P(X_{o,t}, X_{m,t})] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\Big[\log P(X_{o,t}, X_{m,t}\Big|X_{m,t}|X_{o,t} = x_{o,t})\Big] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\left[\log\left(\frac{1}{\pi^L\,\mathrm{Det}(\Sigma)} e^{-(X_t-\mu)^H\Sigma^{-1}(X_t-\mu)}\right)\Big|X_{m,t}|X_{o,t} = x_{o,t}\right] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\Big[-L\log(\pi) - \log(\mathrm{Det}(\Sigma)) - (X_t-\mu)^H\Sigma^{-1}(X_t-\mu)\Big|X_{m,t}|X_{o,t} = x_{o,t}\Big] =$$
$$-L\log(\pi) - \log\big(\mathrm{Det}(\Sigma_{X_t|(X_{m,t}|X_{o,t}=x_{o,t})})\big) +$$
$$+\mathbb{E}_{\Psi^{(\tau-1)}}\Big[-(X_t-\mu)^H\Sigma^{-1}(X_t-\mu)\Big|X_{m,t}|X_{o,t} = x_{o,t}\Big] =$$
$$-L\log(\pi) - \log\big(\mathrm{Det}(\Sigma_{X_t|(X_{m,t}|X_{o,t}=x_{o,t})})\big) +$$
$$+\mathbb{E}_{\Psi^{(\tau-1)}}\left[-\begin{bmatrix} X_{m,t} - \mu_{X_{m,t}} \\ X_{o,t} - \mu_{X_{o,t}} \end{bmatrix}^H \Sigma^{-1} \begin{bmatrix} X_{m,t} - \mu_{X_{m,t}} \\ X_{o,t} - \mu_{X_{o,t}} \end{bmatrix} \Big|X_{m,t}|X_{o,t} = x_{o,t}\right] =$$
$$-L\log(\pi) - \log\big(\mathrm{Det}(\Sigma_{X_t|(X_{m,t}|X_{o,t}=x_{o,t})})\big) +$$
$$+\mathbb{E}_{\Psi^{(\tau-1)}}\left[-\begin{bmatrix} X_{m,t} \\ X_{o,t} \end{bmatrix}^H \begin{bmatrix} \Sigma_{(m,t)(m,t)} & \Sigma_{(m,t)(o,t)} \\ \Sigma_{(o,t)(m,t)} & \Sigma_{(o,t)(o,t)} \end{bmatrix}^{-1} \begin{bmatrix} X_{m,t} \\ X_{o,t} \end{bmatrix} \Big|X_{m,t}|X_{o,t} = x_{o,t}\right] =$$

$$-L\log(\pi) - \log\big(\mathrm{Det}(\Sigma_{X_t|(X_{m,t}|X_{o,t}=x_{o,t})})\big) -$$
$$-\mathbb{E}_{\Psi^{(\tau-1)}}\left[X^H_{m,t}\Sigma^{-1}_{(m,t)(m,t)}X_{m,t} + 2X^H_{o,t}\Sigma^{-1}_{(m,t)(o,t)}X_{m,t} + X^H_{o,t}\Sigma^{-1}_{(o,t)(o,t)}X_{o,t}\Big|X_{m,t}|X_{o,t}=x_{o,t}\right] =$$
$$-L\log(\pi) - \log\big(\mathrm{Det}(\Sigma_{X_t|(X_{m,t}|X_{o,t}=x_{o,t})})\big) -$$
$$-x^H_{o,t}\Sigma^{-1}_{(o,t)(o,t)}x_{o,t} - 2x^H_{o,t}\Sigma^{-1}_{(m,t)(o,t)}\hat{x}^{(\tau)}_{m,t} - \mathbb{E}_{\Psi^{(\tau-1)}}\left[X^H_{m,t}\Sigma^{-1}_{(m,t)(m,t)}X_{m,t}\Big|X_{m,t}|X_{o,t}=x_{o,t}\right]$$

Заметим, что:

$$[\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}X_{m,t} \sim N\Big([\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}\hat{x}^{(\tau)}_{m,t}, [\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}\Sigma^{(\tau)}_{x_{m,t}|x_{o,t}}[\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}\Big)$$

Выполняем последующие преобразования, используя тот факт, что: $\mathbb{E}(WW^H) = \mathsf{cov}(W,W) + \mathbb{E}(W)\mathbb{E}(W)^H$.

$$\mathbb{E}_{\Psi^{(\tau-1)}}\left[X^H_{m,t}\Sigma^{-1}_{(m,t)(m,t)}X_{m,t}\Big|X_{m,t}|X_{o,t}=x_{o,t}\right] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)},X_{m,t}|X_{o,t}=x_{o,t}}\left[[[\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}\Sigma^{(\tau)}_{x_{m,t}|x_{o,t}}]^H[[\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}\Sigma^{(\tau)}_{x_{m,t}|x_{o,t}}]\right] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)},X_{m,t}|X_{o,t}=x_{o,t}}[W^H_tW_t] =$$
$$\mathrm{Tr}\left(\mathbb{E}_{\Psi^{(\tau-1)},X_{m,t}|X_{o,t}=x_{o,t}}[W_tW^H_t]\right) =$$
$$\mathrm{Tr}\left([\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}\Sigma^{(\tau)}_{x_{m,t}|x_{o,t}}[\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}\right) +$$
$$\mathrm{Tr}\left([[\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}\hat{x}^{(\tau)}_{m,t}][[\Sigma^{-1}_{(m,t)(m,t)}]^{\frac{1}{2}}\hat{x}^{(\tau)}_{m,t}]^H\right) =$$
$$\mathrm{Tr}\left([\Sigma^{-1}_{(m,t)(m,t)}]\Sigma^{(\tau)}_{x_{m,t}|x_{o,t}}\right) +$$
$$\mathrm{Tr}\left((\hat{x}^{(\tau)}_{m,t})^H[\Sigma^{-1}_{(m,t)(m,t)}]\hat{x}^{(\tau)}_{m,t}\right)$$

M-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\Psi^{(\tau)} = \underset{\Psi}{\mathrm{argmax}}\,\mathbb{E}_{X_m|X_o=x_o,\Psi^{(\tau-1)}}[\log P(X_o,X_m)] =$$
$$\underset{\Psi}{\mathrm{argmax}}\sum_{t=1}^{G}\left\{\mathbb{E}_{\Psi^{(\tau-1)}}\left[-\log\big(\mathrm{Det}(\Sigma_{X_t|(X_{m,t}|X_{o,t}=x_{o,t})})\big) - \begin{bmatrix}\hat{x}_{m,t}\\ x_{o,t}\end{bmatrix}^H \Sigma^{-1}_{X_t|(X_{m,t}|X_{o,t}=x_{o,t})}\begin{bmatrix}\hat{x}_{m,t}\\ x_{o,t}\end{bmatrix}\right]\right\}$$

Первый CM-шаг

Оценим углы прибытия сигналов θ, но оставляем оценку ковариации сигналов $\mathbf{P}$ фиксированной: $\mathbf{P} = \mathbf{P}^{(\tau-1)}$.

$$\theta^{(\tau)} = \underset{\theta}{\mathrm{argmax}}\sum_{t=1}^{G}\left\{\mathbb{E}_{\Psi^{(\tau-1)}}\left[-\log\big(\mathrm{Det}(\Sigma_{X_t|(X_{m,t}|X_{o,t}=x_{o,t})})\big) - \begin{bmatrix}\hat{x}_{m,t}\\ x_{o,t}\end{bmatrix}^H \Sigma^{-1}_{X_t|(X_{m,t}|X_{o,t}=x_{o,t})}\begin{bmatrix}\hat{x}_{m,t}\\ x_{o,t}\end{bmatrix}\right]\right\}$$

Второй CM-шаг

Оценим ковариацию сигналов $\mathbf{P}$, но оставляем оценку углов прибытия сигналов θ фиксированной: $\theta = \theta^{(\tau)}$.

Неизвестный шум

Е-шаг

М-шаг

Первый СМ-шаг

Второй СМ-шаг

Список источников