

ЕСМ, Детерминированная модель сигналов

Введем некоторые условные обозначения:

- θ — вектор направлений прибытия сигнала (DoA);
- τ — итерация ЕСМ-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 детерминированных сигналов, испускаемых источниками в моменты времени $t = \overline{1, G}$, S_t соответствует сигналу в момент времени t , представляет собой матрицу размера $G \times M$;
- 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):

$$X_t = A(\theta)S_t + N_t, \quad (1)$$

где S_t — детерминированные сигналы, $t = \overline{1, G}$, $N_t \sim CN(\mathbf{O}_{L \times 1}, \mathbf{\Lambda})$, $\theta = [\theta_1, \dots, \theta_M]^T$ — вектор направлений прибытия сигнала, $A(\theta)$ (далее — A) представляет собой матрицу управляющих векторов размера $L \times M$, $\mathbf{\Lambda}$ предполагается диагональной. 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 ;
- Λ_{m_t} — ковариационная матрица шума на неисправных сенсорах в момент времени t ;
- Λ_{o_t} — ковариационная матрица шума на исправных сенсорах в момент времени t .

Рассмотрим 2 случая:

- Известный шум
- Неизвестный шум

§1 Известный шум

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

Е-шаг

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

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

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

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

$$\begin{aligned} X_t &= AS_t + N_t \\ X_t &\sim CN(AS_t, \Lambda) \\ X_{o_t} &\sim CN(A_{o_t}S_t, \Lambda_{o_t}) \end{aligned}$$

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

$$P(X_o|\Psi) = \prod_{t=1}^G \frac{1}{\pi^{L_{o_t}} \text{Det}(\Lambda_{o_t})} e^{-(X_{o_t} - A_{o_t}S_t)^H (\Lambda_{o_t})^{-1} (X_{o_t} - A_{o_t}S_t)}, \quad (5)$$

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

$$\begin{cases} \mu_{x_{m_t}|x_{o_t}} = \mu_{x_{m_t}} + \Sigma_{x_{m_t}, x_{o_t}} \Sigma_{x_{o_t}, x_{o_t}}^{-1} \cdot (x_{o_t} - \mu_{x_{o_t}}) \\ \Sigma_{x_{m_t}|x_{o_t}} = \Sigma_{x_{m_t}, x_{m_t}} - \Sigma_{x_{m_t}, x_{o_t}} \Sigma_{x_{o_t}, x_{o_t}}^{-1} \Sigma_{x_{o_t}, x_{m_t}} \end{cases} \quad (6)$$

В рамках данной задачи:

$$\begin{cases} \Sigma_{x_{o_t}, x_{o_t}} = \mathbf{\Lambda}_{o_t} \\ \Sigma_{x_{o_t}, x_{m_t}} = \hat{\Sigma}_{x_{o_t}, x_{m_t}} \\ \Sigma_{x_{m_t}, x_{o_t}} = \hat{\Sigma}_{x_{o_t}, x_{m_t}}^H \\ \Sigma_{x_{m_t}, x_{m_t}} = \mathbf{\Lambda}_{m_t} \\ \mu_{x_{o_t}} = A_{o_t} S_t \\ \mu_{x_{m_t}} = A_{m_t} S_t \end{cases} \quad (7)$$

где $\hat{\Sigma}_{x_{o_t}, x_{m_t}}$ – выборочная оценка ковариации на основе полных наблюдений, т.е. таких, что $x_t = x_{o_t}$

$$\begin{cases} \mu_{x_{m_t}|x_{o_t}} = A_{m_t} S_t + \hat{\Sigma}_{x_{o_t}, x_{m_t}}^H (\mathbf{\Lambda}_{o_t})^{-1} \cdot (x_{o_t} - A_{o_t} S_t) \\ \Sigma_{x_{m_t}|x_{o_t}} = \mathbf{\Lambda}_{m_t} - \hat{\Sigma}_{x_{o_t}, x_{m_t}}^H (\mathbf{\Lambda}_{o_t})^{-1} \hat{\Sigma}_{x_{o_t}, x_{m_t}} \end{cases} \quad (8)$$

Оцениваем пропущенные значения для каждого наблюдения через условное математическое ожидание: $\hat{x}_{m_t} = \mu_{x_{m_t}|x_{o_t}}$. Вернемся к ранее рассмотренному условному математическому ожиданию:

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

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 :

$$\begin{aligned} & \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)} \left[\log P(X_{o_t}, X_{m_t} | X_{m_t} | X_{o_t} = x_{o_t}) \right] = \\ & \mathbb{E}_{\Psi(\tau-1)} \left[\log \left(\frac{1}{\pi^L \text{Det}(\Sigma)} e^{-(X_t - \mu)^H \Sigma^{-1} (X_t - \mu)} \right) \middle| X_{m_t} | X_{o_t} = x_{o_t} \right] = \\ & \mathbb{E}_{\Psi(\tau-1)} \left[-L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) - (X_t - \mu)^H \mathbf{\Lambda}^{-1} (X_t - \mu) \middle| X_{m_t} | X_{o_t} = x_{o_t} \right] = \\ & \quad -L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) + \\ & \quad + \mathbb{E}_{\Psi(\tau-1)} \left[- (X_t - \mu)^H \mathbf{\Lambda}^{-1} (X_t - \mu) \middle| X_{m_t} | X_{o_t} = x_{o_t} \right] = \\ & \quad -L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) + \\ & \quad + \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 \mathbf{\Lambda}^{-1} \begin{bmatrix} X_{m_t} - \mu_{X_{m_t}} \\ X_{o_t} - \mu_{X_{o_t}} \end{bmatrix} \middle| X_{m_t} | X_{o_t} = x_{o_t} \right] = \\ & \quad -L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) + \\ & \quad + \mathbb{E}_{\Psi(\tau-1)} \left[- \begin{bmatrix} X_{m_t} - A_{m_t} S_t \\ X_{o_t} - A_{o_t} S_t \end{bmatrix}^H \begin{bmatrix} \mathbf{\Lambda}_{(m_t, m_t)} & \mathbf{O} \\ \mathbf{O} & \mathbf{\Lambda}_{(o_t, o_t)} \end{bmatrix}^{-1} \begin{bmatrix} X_{m_t} - A_{m_t} S_t \\ X_{o_t} - A_{o_t} S_t \end{bmatrix} \middle| X_{m_t} | X_{o_t} = x_{o_t} \right] = \\ & \quad -L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) - \\ & \quad \mathbb{E}_{\Psi(\tau-1)} \left[(X_{m_t} - A_{m_t} S_t)^H (\mathbf{\Lambda}^{-1})_{(m_t, m_t)} (X_{m_t} - A_{m_t} S_t) + \right. \\ & \quad \left. (X_{o_t} - A_{o_t} S_t)^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} (X_{o_t} - A_{o_t} S_t) \middle| X_{m_t} | X_{o_t} = x_{o_t} \right] = \end{aligned}$$

$$-L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) - (X_{o_t} - A_{o_t} S_t)^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} (X_{o_t} - A_{o_t} S_t) -$$

$$\mathbb{E}_{\Psi^{(\tau-1)}} \left[(X_{m_t} - A_{m_t} S_t)^H (\mathbf{\Lambda}^{-1})_{(m_t, m_t)} (X_{m_t} - A_{m_t} S_t) \middle| X_{m_t} | X_{o_t} = x_{o_t} \right] =$$

Заметим, что:

$$[(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} X_{m_t} \sim N \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} \hat{x}_{m_t}^{(\tau)}, [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} \mathbf{\Lambda}_{m_t} [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} \right)$$

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

$$\begin{aligned} & \mathbb{E}_{\Psi^{(\tau-1)}} \left[X_{m_t}^H (\mathbf{\Lambda}^{-1})_{(m_t, m_t)} X_{m_t} \middle| X_{m_t} | X_{o_t} = x_{o_t} \right] = \\ & \mathbb{E}_{\Psi^{(\tau-1)}, X_{m_t} | X_{o_t} = x_{o_t}} \left[\left[[(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} \mathbf{\Lambda}_{m_t} \right]^H \left[[(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} \mathbf{\Lambda}_{m_t} \right] \right] = \\ & \mathbb{E}_{\Psi^{(\tau-1)}, X_{m_t} | X_{o_t} = x_{o_t}} [W_t^H W_t] = \\ & \text{Tr} \left(\mathbb{E}_{\Psi^{(\tau-1)}, X_{m_t} | X_{o_t} = x_{o_t}} [W_t W_t^H] \right) = \\ & \text{Tr} \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} \mathbf{\Lambda}_{m_t} [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} \right) + \text{Tr} \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} \hat{x}_{m_t}^{(\tau)} \left[[(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}]^{\frac{1}{2}} \hat{x}_{m_t}^{(\tau)} \right]^H \right) = \\ & \text{Tr} \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \mathbf{\Lambda}_{m_t} \right) + \text{Tr} \left((\hat{x}_{m_t}^{(\tau)})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \hat{x}_{m_t}^{(\tau)} \right) \\ & \mathbb{E}_{X_{m_t} | X_{o_t} = x_{o_t}, \Psi^{(\tau-1)}} [\log P(X_{o_t}, X_{m_t})] = \\ & -L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) - x_{o_t}^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} x_{o_t} - \\ & \text{Tr} \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \mathbf{\Lambda}_{m_t} \right) - \text{Tr} \left((\hat{x}_{m_t}^{(\tau)})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \hat{x}_{m_t}^{(\tau)} \right) \end{aligned}$$

Таким образом, оптимизируемая функция примет следующий вид:

$$\begin{aligned} & \mathbb{E}_{X_m | X_o = x_o, \Psi^{(\tau)}} [\log P(X_o, X_m)] = \\ & \sum_{t=1}^G \mathbb{E}_{X_{m_t} | X_{o_t} = x_{o_t}, \Psi^{(\tau)}} [\log P(X_{o_t}, X_{m_t})] = \\ & \sum_{t=1}^G \left[-L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) - x_{o_t}^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} x_{o_t} - \right. \\ & \left. \text{Tr} \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \mathbf{\Lambda}_{m_t} \right) - \text{Tr} \left((\hat{x}_{m_t}^{(\tau)})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \hat{x}_{m_t}^{(\tau)} \right) \right] \end{aligned}$$

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\begin{aligned} & \Psi^{(\tau)} = \underset{\Psi}{\text{argmax}} \mathbb{E}_{X_m | X_o = x_o, \Psi^{(\tau-1)}} [\log P(X_o, X_m)] = \\ & \underset{\Psi}{\text{argmax}} \sum_{t=1}^G \left[-L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) - x_{o_t}^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} x_{o_t} - \right. \\ & \left. \text{Tr} \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \mathbf{\Lambda}_{m_t} \right) - \text{Tr} \left((\hat{x}_{m_t}^{(\tau)})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \hat{x}_{m_t}^{(\tau)} \right) \right] \end{aligned}$$

Первый СМ-шаг

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

$$\begin{aligned}\theta^{(\tau)} &= \operatorname{argmax}_{\theta} Q(\theta | \theta^{(\tau-1)}) = \\ \operatorname{argmax}_{\Psi} \sum_{t=1}^G &\left[-L \log(\pi) - \log(\operatorname{Det}(\mathbf{\Lambda})) - x_{o_t}^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} x_{o_t} - \right. \\ &\left. \operatorname{Tr} \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \mathbf{\Lambda}_{m_t} \right) - \operatorname{Tr} \left((\hat{x}_{m_t}^{(\tau)})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \hat{x}_{m_t}^{(\tau)} \right) \right]\end{aligned}$$

Внутри данной функции, вектор θ возникает только как переменная, влияющая на μ . Если будет найдено оптимальное значение $\mu^{(\tau)}$ вектора μ , можно будет найти θ , численно решив $\sum_{t=1}^G A(\theta^{(\tau)}) S_t = G \cdot \mu^{(\tau)}$ относительно θ .

Пусть $\tilde{x}_t^{(\tau)}$ — вектор x_t , в котором пропущенные значения x_{m_t} оценены с помощью $\hat{x}_{m_t}^{(\tau)}$.

$$\begin{aligned}\nabla_{\mu} Q(\theta | \theta^{(\tau-1)}) &= \sum_{t=1}^G \nabla_{\mu} Q_t(\theta | \theta^{(\tau-1)}) \\ &= \sum_{t=1}^G \nabla_{\mu} \left(-\frac{1}{2} (\tilde{x}_t^{(\tau)} - \mu)^H \mathbf{\Lambda}^{-1} (\tilde{x}_t^{(\tau)} - \mu) \right) \\ &= -\frac{1}{2} \sum_{t=1}^G \nabla_{\mu} (\tilde{x}_t^{(\tau)} - \mu)^H \mathbf{\Lambda}^{-1} (\tilde{x}_t^{(\tau)} - \mu) \\ &= -\frac{1}{2} \sum_{t=1}^G 2 \mathbf{\Lambda}^{-1} (\tilde{x}_t^{(\tau)} - \mu) \nabla_{\mu} (\tilde{x}_t^{(\tau)} - \mu) \\ &= -\frac{1}{2} \sum_{t=1}^G 2 \mathbf{\Lambda}^{-1} (\tilde{x}_t^{(\tau)} - \mu) (O - I_p) \\ &= \mathbf{\Lambda}^{-1} \sum_{t=1}^G (\tilde{x}_t^{(\tau)} - \mu)\end{aligned}$$

Приравняем производную к нулю, чтобы получить значение μ , соответствующее максимуму.

$$\begin{aligned}\mathbf{\Lambda}^{-1} \sum_{t=1}^G (\tilde{x}_t^{(\tau)} - \mu^{(\tau)}) &= \mathbf{O} \\ \iff \sum_{t=1}^G (\tilde{x}_t^{(\tau)} - \mu^{(\tau)}) &= \mathbf{O} \\ \iff \sum_{t=1}^G \tilde{x}_t^{(\tau)} &= G \mu^{(\tau)} \\ \iff \mu^{(\tau)} &= \frac{1}{G} \sum_{t=1}^G \tilde{x}_t^{(\tau)}\end{aligned}$$

Можем теперь численно решить следующее уравнение относительно $\theta^{(\tau)}$:

$$\sum_{t=1}^G A(\theta^{(\tau)}) S_t = G \cdot \mu^{(\tau)}. \quad (9)$$

Для упрощений вычислений, преобразуем равенство:

$$A(\theta^{(\tau)}) \sum_{t=1}^G S_t = G \cdot \mu^{(\tau)}. \quad (10)$$

Второй СМ-шаг

Оценим сигналы S , но оставляем оценку углов прибытия сигналов θ фиксированной: $\theta = \theta^{(\tau)}$. Можем теперь численно решить следующую систему уравнений относительно $\mathbf{S}^{(\tau)} = \{S_t^{(\tau)}\}_{t=1}^G$:

$$\begin{cases} A(\theta^{(\tau)})S_1^{(\tau)} = \tilde{x}_1^{(\tau)} \\ A(\theta^{(\tau)})S_2^{(\tau)} = \tilde{x}_2^{(\tau)} \\ \vdots \\ A(\theta^{(\tau)})S_G^{(\tau)} = \tilde{x}_G^{(\tau)} \end{cases} \quad (11)$$

§2 Неизвестный шум

Е-шаг

М-шаг

Первый СМ-шаг

Второй СМ-шаг

Список источников