

ЕСМ, Детерминированная модель сигналов

§1 Постановка проблемы

Предположим, у нас имеется линейная антенная решетка, состоящая из L сенсоров. Решетка принимает волны, направленные из M различных источников. Этим источникам соответствует вектор углов прибытия (DoA) θ , практически не изменяющийся с течением времени. По итогам измерений, было получено G снимков полученного сигнала, причем ввиду технических неполадок, связанных с сенсорами, большая часть таких снимков содержит помимо надежных данных ненадежные, которые в рамках данной задачи рассматриваются как пропуски. Пусть X — полный набор сигналов, полученных датчиками в моменты времени $t = \overline{1, G}$, X_t соответствует сигналу в момент времени t , через x и x_t будем обозначать реализации полного набора сигналов и сигнала в отдельный момент времени t соответственно. Ввиду наличия пропусков в данных, будем считать, что X состоит из наблюдаемой части $X_o = \{X_{o_t}\}_{t=1}^G$ и ненаблюдаемой: $X_m = \{X_{m_t}\}_{t=1}^G$. Полученный сигнал является результатом следующей модели наблюдений:

$$X = AS + N, \quad (1)$$

где $N = \{N_t\}_{t=1}^G$ соответствует набору шумов, связанных с датчиками в моменты времени $t = \overline{1, G}$, $S = \{S_t\}_{t=1}^G$ соответствует набору сигналов, испускаемых источниками в моменты времени $t = \overline{1, G}$, A — матрица управляющих векторов для равномерного линейного массива:

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

Сигналы, испускаемые источниками, рассматриваются как детерминированные; в то же время шумы на сенсорах, предполагаются стохастическими: $N_t \sim CN(\mathbf{0}_{L \times 1}, \mathbf{\Lambda})$. Матрица $\mathbf{\Lambda}$ предполагается диагональной, т.е. шумы рассматриваются как некоррелированные. Для простоты дальнейших рассуждений введем также следующие величины:

- 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 сенсoram в момент времени t ;
- A_{m_t} — матрица, образованная теми строками матрицы A , которые соответствуют неисправным сенсoram 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 .

Составим ЕСМ-алгоритм (Expectation Conditional Maximization алгоритм) для двух случаев:

- Известный шум;
- Неизвестный шум.

§2 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. Наблюдения $X_t, t = 1, \bar{L}$ предполагаются независимыми и одинаково распределенными.

Инициализация параметров

$$\theta^{(0)} \sim \mathcal{U}([- \pi, \pi]^M) \quad (2)$$

$$S^{(0)} = A^+ X_{full}, \quad (3)$$

где $X_{full} = \{x_t : x_t = x_{o_t}, \forall t \in \{1, \dots, G\}\}$.

Е-шаг

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

$$\mathbb{E}_{X_m|X_o, \Psi^{(\tau-1)}} [\log P(X_o, X_m)] \quad (4)$$

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

$$P(X_m|X_o = x_o, \Psi) = \frac{P(X_o, X_m|\Psi)}{P(X_o|\Psi)} = \frac{P(X|\Psi)}{P(X_o|\Psi)} \quad (5)$$

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

$$P(X|\Psi) = \prod_{t=1}^G \frac{1}{\pi^L \text{Det}(\Lambda)} e^{-(X_t - AS_t)^H (\Lambda)^{-1} (X_t - AS_t)}, \quad (6)$$

$$P(X_o|\Psi) = \prod_{t=1}^G \frac{1}{\pi^{L_{o_t}} \text{Det}(\Lambda_{o_t})} e^{-(X_{o_t} - A_{o_t}S_t)^H (\Lambda_{o_t})^{-1} (X_{o_t} - A_{o_t}S_t)}, \quad (7)$$

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

$$\begin{cases} \mu_{x_{m_t}|x_{o_t}} = \mu_{x_{m_t}} + \Sigma_{x_{m_t}, x_{o_t}} \Sigma_{x_{o_t}, x_{o_t}}^{-1} \cdot (x_{o_t} - \mu_{x_{o_t}}) \\ \Sigma_{x_{m_t}|x_{o_t}} = \Sigma_{x_{m_t}, x_{m_t}} - \Sigma_{x_{m_t}, x_{o_t}} \Sigma_{x_{o_t}, x_{o_t}}^{-1} \Sigma_{x_{o_t}, x_{m_t}} \end{cases} \quad (8)$$

В рамках данной задачи:

$$\begin{cases} \Sigma_{x_{o_t}, x_{o_t}} = \Lambda_{o_t} \\ \Sigma_{x_{o_t}, x_{m_t}} = \hat{\Sigma}_{x_{o_t}, x_{m_t}} \\ \Sigma_{x_{m_t}, x_{o_t}} = \hat{\Sigma}_{x_{o_t}, x_{m_t}}^H \\ \Sigma_{x_{m_t}, x_{m_t}} = \Lambda_{m_t} \\ \mu_{x_{o_t}} = A_{o_t} S_t \\ \mu_{x_{m_t}} = A_{m_t} S_t \end{cases} \quad (9)$$

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

$$\begin{cases} \mu_{x_{m_t}|x_{o_t}} = A_{m_t}S_t + \hat{\Sigma}_{x_{o_t}, x_{m_t}}^H (\mathbf{\Lambda}_{o_t})^{-1} \cdot (x_o - A_{o_t}S_t) \\ \Sigma_{x_{m_t}|x_{o_t}} = \mathbf{\Lambda}_{m_t} - \hat{\Sigma}_{x_{o_t}, x_{m_t}}^H (\mathbf{\Lambda}_{o_t})^{-1} \hat{\Sigma}_{x_{o_t}, x_{m_t}} \end{cases} \quad (10)$$

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) | 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) | 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) | 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[- \left[\frac{X_{m_t} - \mu_{X_{m_t}}}{X_{o_t} - \mu_{X_{o_t}}} \right]^H \mathbf{\Lambda}^{-1} \left[\frac{X_{m_t} - \mu_{X_{m_t}}}{X_{o_t} - \mu_{X_{o_t}}} \right] | 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[- \left[\frac{X_{m_t} - A_{m_t}S_t}{X_{o_t} - A_{o_t}S_t} \right]^H \begin{bmatrix} \mathbf{\Lambda}_{(m_t, m_t)} & \mathbf{O} \\ \mathbf{O} & \mathbf{\Lambda}_{(o_t, o_t)} \end{bmatrix}^{-1} \left[\frac{X_{m_t} - A_{m_t}S_t}{X_{o_t} - A_{o_t}S_t} \right] | 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) | X_{m_t}|X_{o_t}=x_{o_t} \right] = \\ & \quad -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) - \\ & \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) | X_{m_t}|X_{o_t}=x_{o_t} \right] = \end{aligned}$$

Заметим, что:

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

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

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

Таким образом, оптимизируемая функция (с учетом исключения слагаемых, независимых от параметров) примет следующий вид:

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

Рассмотрим второе слагаемое: внутри следа находится скаляр, соответственно, преобразуем выражение:

$$\begin{aligned}
& \sum_{t=1}^G \left[- (x_{o_t} - \mu_{x_{o_t}})^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} (x_{o_t} - \mu_{x_{o_t}}) - \text{Tr} \left((\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] (\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}}) \right) \right] = \\
& \sum_{t=1}^G \left[- (x_{o_t} - \mu_{x_{o_t}})^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} (x_{o_t} - \mu_{x_{o_t}}) - (\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] (\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}}) \right] = \\
& = \sum_{t=1}^G \left(- (X_t - AS_t)^H \mathbf{\Lambda}^{-1} (X_t - AS_t) \right) = \\
& \quad - \|\mathbf{\Lambda}^{-1/2} (X - AS)\|_F^2
\end{aligned}$$

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\begin{aligned} \Psi^{(\tau)} &= \arg \max_{\Psi} \mathbb{E}_{X_m | X_o = x_o, \Psi^{(\tau-1)}} [\log P(X_o, X_m)] = \\ &= \arg \max_{\Psi} \sum_{t=1}^G \left[-L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) - (x_{o_t} - \mu_{X_{o_t}})^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} (x_{o_t} - \mu_{X_{o_t}}) - \right. \\ &\quad \left. \text{Tr} \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \mathbf{\Lambda}_{m_t} \right) - \text{Tr} \left((\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] (\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}}) \right) \right] = \\ &= \arg \min_{\Psi} \|\mathbf{\Lambda}^{-1/2} (X - AS)\|_F^2 \end{aligned}$$

Первый СМ-шаг

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

$$\begin{aligned} \theta^{(\tau)} &= \arg \max_{\theta} Q(\theta | \theta^{(\tau-1)}) = \\ &= \arg \min_{\theta} \|\mathbf{\Lambda}^{-1/2} (X - AS)\|_F^2 \end{aligned}$$

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 $\mathbf{S}^{(\tau)} = \{S_t^{(\tau)}\}_{t=1}^G$:

$$S^\tau = \arg \min_S (X - AS)^H Q^{-1} (X - AS) \quad (11)$$

$$\begin{cases} S_1^{(\tau)} = (A(\theta^{(\tau)})^H Q^{-1} A(\theta^{(\tau)}))^{-1} A(\theta^{(\tau)})^H Q^{-1} \tilde{x}_1^{(\tau)} \\ S_2^{(\tau)} = (A(\theta^{(\tau)})^H Q^{-1} A(\theta^{(\tau)}))^{-1} A(\theta^{(\tau)})^H Q^{-1} \tilde{x}_2^{(\tau)} \\ \vdots \\ S_G^{(\tau)} = (A(\theta^{(\tau)})^H Q^{-1} A(\theta^{(\tau)}))^{-1} A(\theta^{(\tau)})^H Q^{-1} \tilde{x}_G^{(\tau)} \end{cases} \quad (12)$$

Шаги повторяются либо до достижения максимального числа итераций либо до сходимости оценок параметров, полученных на соседних двух итерациях.

§3 Неизвестный шум

Воспользуемся ЕСМ-алгоритмом для того, чтобы определить значения параметров $\Psi = (\theta, S, \Lambda)$, пропущенные значения $X_m = \{X_{m_t}\}_{t=1}^G$ рассматриваются как латентные переменные. Наблюдения X_t , $t = \overline{1, L}$ предполагаются независимыми и одинаково распределенными.

Инициализация параметров

$$\theta^{(0)} \sim \mathcal{U}([- \pi, \pi]^M) \quad (13)$$

$$S^{(0)} = A^+ X_{full}, \quad (14)$$

где $X_{full} = \{x_t : x_t = x_{o_t}, \forall t \in \{1, \dots, G\}\}$. Шумы на сенсорах задаются с учетом дисперсии остатков

$$\hat{q}_{jj}^{(0)} = \frac{1}{G} \sum_{t=1}^G |X_{j,t} - (AS)_{j,t}|^2 \quad (15)$$

Е-шаг

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

$$\mathbb{E}_{X_m|X_o, \Psi^{(\tau-1)}} [\log P(X_o, X_m)] \quad (16)$$

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

$$P(X_m|X_o = x_o, \Psi) = \frac{P(X_o, X_m|\Psi)}{P(X_o|\Psi)} = \frac{P(X|\Psi)}{P(X_o|\Psi)} \quad (17)$$

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

$$P(X|\Psi) = \prod_{t=1}^G \frac{1}{\pi^L \text{Det}(\Lambda)} e^{-(X_t - AS_t)^H (\Lambda)^{-1} (X_t - AS_t)}, \quad (18)$$

$$P(X_o|\Psi) = \prod_{t=1}^G \frac{1}{\pi^{L_{o_t}} \text{Det}(\Lambda_{o_t})} e^{-(X_{o_t} - A_{o_t}S_t)^H (\Lambda_{o_t})^{-1} (X_{o_t} - A_{o_t}S_t)}, \quad (19)$$

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

$$\begin{cases} \mu_{x_{m_t}|x_{o_t}} = \mu_{x_{m_t}} + \Sigma_{x_{m_t}, x_{o_t}} \Sigma_{x_{o_t}, x_{o_t}}^{-1} \cdot (x_{o_t} - \mu_{x_{o_t}}) \\ \Sigma_{x_{m_t}|x_{o_t}} = \Sigma_{x_{m_t}, x_{m_t}} - \Sigma_{x_{m_t}, x_{o_t}} \Sigma_{x_{o_t}, x_{o_t}}^{-1} \Sigma_{x_{o_t}, x_{m_t}} \end{cases} \quad (20)$$

В рамках данной задачи:

$$\begin{cases} \Sigma_{x_{o_t}, x_{o_t}} = \Lambda_{o_t} \\ \Sigma_{x_{o_t}, x_{m_t}} = \hat{\Sigma}_{x_{o_t}, x_{m_t}} \\ \Sigma_{x_{m_t}, x_{o_t}} = \hat{\Sigma}_{x_{o_t}, x_{m_t}}^H \\ \Sigma_{x_{m_t}, x_{m_t}} = \Lambda_{m_t} \\ \mu_{x_{o_t}} = A_{o_t} S_t \\ \mu_{x_{m_t}} = A_{m_t} S_t \end{cases} \quad (21)$$

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

$$\begin{cases} \mu_{x_{m_t}|x_{o_t}} = A_{m_t}S_t + \hat{\Sigma}_{x_{o_t}, x_{m_t}}^H (\mathbf{\Lambda}_{o_t})^{-1} \cdot (x_o - A_{o_t}S_t) \\ \Sigma_{x_{m_t}|x_{o_t}} = \mathbf{\Lambda}_{m_t} - \hat{\Sigma}_{x_{o_t}, x_{m_t}}^H (\mathbf{\Lambda}_{o_t})^{-1} \hat{\Sigma}_{x_{o_t}, x_{m_t}} \end{cases} \quad (22)$$

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) | 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) | 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) | 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[- \left[\frac{X_{m_t} - \mu_{X_{m_t}}}{X_{o_t} - \mu_{X_{o_t}}} \right]^H \mathbf{\Lambda}^{-1} \left[\frac{X_{m_t} - \mu_{X_{m_t}}}{X_{o_t} - \mu_{X_{o_t}}} \right] | 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[- \left[\frac{X_{m_t} - A_{m_t}S_t}{X_{o_t} - A_{o_t}S_t} \right]^H \begin{bmatrix} \mathbf{\Lambda}_{(m_t, m_t)} & \mathbf{O} \\ \mathbf{O} & \mathbf{\Lambda}_{(o_t, o_t)} \end{bmatrix}^{-1} \left[\frac{X_{m_t} - A_{m_t}S_t}{X_{o_t} - A_{o_t}S_t} \right] | 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) | X_{m_t}|X_{o_t}=x_{o_t} \right] = \\ & \quad -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) - \\ & \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) | X_{m_t}|X_{o_t}=x_{o_t} \right] = \end{aligned}$$

Заметим, что:

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

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

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

Таким образом, оптимизируемая функция (с учетом исключения слагаемых, независимых от параметров) примет следующий вид:

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

Рассмотрим второе слагаемое: внутри следа находится скаляр, соответственно, преобразуем выражение:

$$\begin{aligned}
& \sum_{t=1}^G \left[- (x_{o_t} - \mu_{x_{o_t}})^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} (x_{o_t} - \mu_{x_{o_t}}) - \text{Tr} \left((\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] (\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}}) \right) \right] = \\
& \sum_{t=1}^G \left[- (x_{o_t} - \mu_{x_{o_t}})^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} (x_{o_t} - \mu_{x_{o_t}}) - (\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] (\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}}) \right] = \\
& = \sum_{t=1}^G \left(- (X_t - AS_t)^H \mathbf{\Lambda}^{-1} (X_t - AS_t) \right) = \\
& \quad - \|\mathbf{\Lambda}^{-1/2} (X - AS)\|_F^2
\end{aligned}$$

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\begin{aligned} \Psi^{(\tau)} &= \arg \max_{\Psi} \mathbb{E}_{X_m | X_o = x_o, \Psi^{(\tau-1)}} [\log P(X_o, X_m)] = \\ &= \arg \max_{\Psi} \sum_{t=1}^G \left[-L \log(\pi) - \log(\text{Det}(\mathbf{\Lambda})) - (x_{o_t} - \mu_{X_{o_t}})^H (\mathbf{\Lambda}^{-1})_{(o_t, o_t)} (x_{o_t} - \mu_{X_{o_t}}) - \right. \\ &\quad \left. \text{Tr} \left([(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] \mathbf{\Lambda}_{m_t} \right) - \text{Tr} \left((\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H [(\mathbf{\Lambda}^{-1})_{(m_t, m_t)}] (\hat{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}}) \right) \right] = \\ &= \arg \min_{\Psi} \|\mathbf{\Lambda}^{-1/2} (X - AS)\|_F^2 \end{aligned}$$

Первый СМ-шаг

Оценим углы прибытия сигналов θ , но оставляем оценку сигналов S и оценку ковариации шума $\mathbf{\Lambda}$ фиксированными: $S = S^{(\tau-1)}$, $\mathbf{\Lambda} = \mathbf{\Lambda}^{(\tau-1)}$.

$$\begin{aligned} \theta^{(\tau)} &= \arg \max_{\theta} Q(\theta | \theta^{(\tau-1)}) = \\ &= \arg \min_{\theta} \|[\mathbf{\Lambda}^{(\tau-1)}]^{-1/2} (X^{(\tau)} - A^{(\tau-1)} S^{(\tau-1)})\|_F^2 \end{aligned}$$

Второй СМ-шаг

Оценим сигналы S , но оставляем оценку углов прибытия сигналов θ и оценку ковариации шума $\mathbf{\Lambda}$ фиксированными: $\theta = \theta^{(\tau)}$, $\mathbf{\Lambda} = \mathbf{\Lambda}^{(\tau-1)}$. Обозначим через $A^{(\tau)}$ оценку матрицы управляющих векторов $A(\theta^{(\tau)})$ после получения $\theta^{(\tau)}$. Можем теперь численно решить следующую систему уравнений относительно $\mathbf{S}^{(\tau)} = \{S_t^{(\tau)}\}_{t=1}^G$:

$$S^\tau = \arg \min_S (X^{(\tau)} - A^{(\tau)} S^{(\tau-1)})^H [\mathbf{\Lambda}^{(\tau-1)}]^{-1} (X^{(\tau)} - A^{(\tau)} S^{(\tau-1)}) \quad (23)$$

$$\begin{cases} S_1^{(\tau)} = (A^{(\tau)})^H [\mathbf{\Lambda}^{(\tau-1)}]^{-1} A^{(\tau)-1} (A^{(\tau)})^H [\mathbf{\Lambda}^{(\tau-1)}]^{-1} \tilde{x}_1^{(\tau)} \\ S_2^{(\tau)} = (A^{(\tau)})^H [\mathbf{\Lambda}^{(\tau-1)}]^{-1} A^{(\tau)-1} (A^{(\tau)})^H [\mathbf{\Lambda}^{(\tau-1)}]^{-1} \tilde{x}_2^{(\tau)} \\ \vdots \\ S_G^{(\tau)} = (A^{(\tau)})^H [\mathbf{\Lambda}^{(\tau-1)}]^{-1} A^{(\tau)-1} (A^{(\tau)})^H [\mathbf{\Lambda}^{(\tau-1)}]^{-1} \tilde{x}_G^{(\tau)} \end{cases} \quad (24)$$

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

Третий СМ-шаг

Оценим ковариацию шума $\mathbf{\Lambda}$, но оставляем оценку углов прибытия сигналов θ и оценку сигналов фиксированными: $\theta = \theta^{(\tau)}$, $S = S^{(\tau)}$. Один из вариантов оценить диагональные элементы ковариационной матрицы шума:

$$\hat{q}_{jj}^{(\tau)} = \frac{1}{G} \sum_{t=1}^G |X_{j,t}^{(\tau)} - (A^{(\tau)} S^{(\tau)})_{j,t}|^2 \quad (25)$$

Если же напрямую оптимизировать целевую функцию $Q(\theta|\theta^{(\tau-1)})$, можем получить иную аналитическую оценку ковариационной матрицы шума: Пусть $\tilde{x}_t^{(\tau)}$ — вектор x_t , в котором пропущенные значения x_{m_t} оценены с помощью $\hat{x}_{m_t}^{(\tau)}$.

$$\begin{aligned}
\nabla_{\Sigma^{-1}} Q(\theta \mid \theta^{(\tau-1)}) &= \sum_{t=1}^G \nabla_{\Sigma^{-1}} Q_t(\theta \mid \theta^{(\tau-1)}) \\
&= \sum_{t=1}^G \frac{1}{2} \nabla_{\Sigma^{-1}} \log |\Sigma^{-1}| - \frac{1}{2} \nabla_{\Sigma^{-1}} \text{Tr}((\Sigma^{-1})_{(m_t, m_t)} \Sigma_{x_{m_t}|x_{o_t}}^{(\tau)}) \\
&\quad - \sum_{t=1}^G \frac{1}{2} \nabla_{\Sigma^{-1}} (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)})^H \Sigma^{-1} (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)}) \\
&= \frac{1}{2} \sum_{t=1}^G \nabla_{\Sigma^{-1}} \log |\Sigma^{-1}| - \nabla_{\Sigma^{-1}} \text{Tr}((\Sigma^{-1})_{(m_t, m_t)} \Sigma_{x_{m_t}|x_{o_t}}^{(\tau)}) \\
&\quad - \frac{1}{2} \sum_{t=1}^G \nabla_{\Sigma^{-1}} \text{Tr}((\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)})^H \Sigma^{-1} (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)})) \\
\nabla_{\Sigma^{-1}} Q(\theta \mid \theta^{(\tau-1)}) &= \frac{1}{2} \left[\sum_{t=1}^G \Sigma - \tilde{\Sigma}_t^{(\tau)} \right] - \frac{1}{2} \sum_{t=1}^G (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)}) (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)})^H \\
&= \frac{1}{2} \left[G \Sigma - \sum_{t=1}^G \tilde{\Sigma}_t^{(\tau)} \right] - \frac{1}{2} \sum_{t=1}^G (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)}) (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)})^H
\end{aligned}$$

где $\Sigma_t^{(\tau)}$ — матрица размера $L \times L$, в которой все элементы являются нулями, за исключением тех, что стоят на пересечении строк с номерами $j_1 \in m_t$ и столбцов с номерами $j_2 \in m_t$: они заменены величиной $\Sigma_{x_{m_t}|x_{o_t}}^{(\tau)}$. Приравняем производную к нулю, чтобы получить значение Σ , соответствующее максимуму.

$$\begin{aligned}
O &= \frac{1}{2} \left[G \Sigma^{(\tau)} - \sum_{t=1}^G \tilde{\Sigma}_t^{(\tau)} \right] - \frac{1}{2} \sum_{t=1}^G (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)}) (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)})^H \\
&= \left[G \Sigma^{(\tau)} - \sum_{t=1}^G \tilde{\Sigma}_t^{(\tau)} \right] - \sum_{t=1}^G \tilde{x}_t^{(\tau)} (\tilde{x}_t^{(\tau)})^H \\
\iff G \Sigma^{(\tau)} &= \sum_{t=1}^G \tilde{\Sigma}_t^{(\tau)} + \sum_{t=1}^G (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)}) (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)})^H \\
\iff \Sigma^{(\tau)} &= \frac{1}{G} \sum_{t=1}^G \left[\tilde{\Sigma}_t^{(\tau)} + (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)}) (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)})^H \right]
\end{aligned}$$

Учтем тот факт, что матрица шума оценивается как диагональная:

$$\iff \mathbf{\Lambda}^{(\tau)} = \Sigma^{(\tau)} = \text{diag} \left(\frac{1}{G} \sum_{t=1}^G \left[\tilde{\Sigma}_t^{(\tau)} + (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)}) (\tilde{x}_t^{(\tau)} - A^{(\tau)} S^{(\tau)})^H \right] \right)$$

Шаги повторяются либо до достижения максимального числа итераций либо до сходимости оценок параметров, полученных на соседних двух итерациях.

Список источников