ECM, 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 во 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 наблюдений (сигналов, полученных сенсорами в моменты времени $t = \overline{1, G}$), X_t соответствует наблюдению в момент времени t, через x и x_t будем обозначать реализации полного набора наблюдений и наблюдения 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: $X_m = \{X_{m_t}\}_{t=1}^G$, причем $X_t = X_{o_t} \cup X_{m_t}, \forall t \in \{1, ..., G\}$. Предполагается, что нет таких наблюдений, которые состоят лишь из ненаблюдаемой части. Набор наблюдений X является результатом следующей модели наблюдений:

$$X = AS + N, \tag{1}$$

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: $N_t \sim CN(\mathbf{O}_{L\times 1}, \mathbf{\Lambda})$. 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 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.

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

- Известный шум;
- Неизвестный шум.

§2 Известный шум

Воспользуемся ECM-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, $t = \overline{1, L}$ предполагаются независимыми и одинаково распределенными.

Инициализация параметров

Оценим вектор углов $\theta^{(0)}$ следующим образом:

1. Выберем число ν, которое будет соответствовать первому компоненту вектора $\theta^{(0)}$:

$$\nu \sim \mathcal{U}([-\pi; \pi]); \tag{2}$$

2. Оценим компоненты вектора $\theta^{(0)}$ так: $\theta_i^{(0)} = (\nu + (i-1) \cdot \frac{2\pi}{M}) \bmod 2\pi, i = \overline{1, M}$. При этом, $a \bmod b = a - b \cdot \lfloor \frac{a}{b} \rfloor$.

Теперь следует составить первоначальную оценку детерминированных сигналов. Предположим, что детерминированные сигналы мы можем представить в форме:

$$S(t) = Be^{j\omega \cdot t + \phi}, \tag{3}$$

где B – амплитуда, ω – частота, ϕ – фаза. Предполагаем, что амплитуда, частота и фаза у сигналов от различных источников могут отличаться, генерируем эти параметры из равномерного распределения. На основе их строим оценку сигналов.

E-шаг

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

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

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

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

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

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

$$P(X_o|\Psi) = \prod_{t=1}^{G} \frac{1}{\pi^{L_{o_t}} \operatorname{Det}(\mathbf{\Lambda}_{o_t})} e^{-(X_{o_t} - A_{o_t}S_t)^H (\mathbf{\Lambda}_{o_t})^{-1}(X_{o_t} - A_{o_t}S_t)}, \tag{7}$$

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

$$\begin{cases} \mu^{(\tau)}_{X_{m_t}|X_{o_t}=x_{o,t}} = \mu^{(\tau)}_{X_{m_t}} + \Sigma^{(\tau)}_{X_{m_t},X_{o_t}} \left(\Sigma^{(\tau)}_{X_{o_t},X_{o_t}}\right)^{-1} \cdot (x_{o_t} - \mu^{(\tau)}_{X_{o_t}}) \\ \Sigma^{(\tau)}_{X_{m_t}|X_{o_t}=x_{o,t}} = \Sigma^{(\tau)}_{X_{m_t},X_{m_t}} - \Sigma^{(\tau)}_{X_{m_t},X_{o_t}} \left(\Sigma^{(\tau)}_{X_{o_t},X_{o_t}}\right)^{-1} \Sigma^{(\tau)}_{X_{o_t},X_{m_t}} \end{cases} \tag{8}$$

В рамках данной задачи:

$$\begin{cases} \Sigma^{(\tau)}_{X_{o_t},X_{o_t}} = \mathbf{\Lambda}_{o_t} \\ \Sigma^{(\tau)}_{X_{o_t},X_{m_t}} = \hat{\Sigma}^{(\tau)}_{X_{o_t},X_{m_t}} \\ \Sigma^{(\tau)}_{X_{m_t},X_{o_t}} = \hat{\Sigma}^{(\tau)}_{X_{m_t},X_{o_t}} \\ \Sigma^{(\tau)}_{X_{m_t},X_{m_t}} = \mathbf{\Lambda}_{m_t} \\ \mu^{(\tau)}_{X_{o_t}} = A^{(\tau)}_{o_t} S^{(\tau)}_t \\ \mu^{(\tau)}_{X_{m_t}} = A^{(\tau)}_{m_t} S^{(\tau)}_t \end{cases} \tag{9}$$

где $\hat{\Sigma}^{(\tau)}_{X_{o_t},X_{m_t}}$ — выборочная оценка кросс-ковариации на основе полных наблюдений, т.е. таких, что $x_t = x_{o_t}$, $\hat{\Sigma}^{(\tau)}_{X_{m_t},X_{o_t}} = \left(\hat{\Sigma}^{(\tau)}_{X_{o_t},X_{m_t}}\right)^H$;

$$\begin{cases} \mu^{(\tau)}_{X_{m_t}|X_{o_t}=x_{o_t}} = A^{(\tau)}_{m_t} S^{(\tau)}_t + \hat{\Sigma}^{(\tau)}_{X_{m_t},X_{o_t}} (\mathbf{\Lambda}_{o_t})^{-1} \cdot (x_o - A^{(\tau)}_{o_t} S^{(\tau)}_t) \\ \Sigma^{(\tau)}_{X_{m_t}|X_{o_t}=x_{o_t}} = \mathbf{\Lambda}_{m_t} - \hat{\Sigma}^{(\tau)}_{X_{m_t},X_{o_t}} (\mathbf{\Lambda}_{o_t})^{-1} \hat{\Sigma}^{(\tau)}_{X_{o_t},X_{m_t}} \end{cases} \tag{10}$$

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

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

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:

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

$$\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}[\log P(X_{o_t}, X_{m_t})] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\left[\log P\Big(X_{o_t}, X_{m_t}\Big| X_{o_t} = x_{o_t}\Big)\right] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\left[\log\left(\frac{1}{\pi^L \operatorname{Det}(\mathbf{\Lambda})} e^{-(X_t-\mu)^H \mathbf{\Lambda}^{-1}(X_t-\mu)}\right) \Big| X_{o_t} = x_{o_t}\right] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\left[-L\log(\pi) - \log(\operatorname{Det}(\mathbf{\Lambda})) - (X_t-\mu)^H \mathbf{\Lambda}^{-1}(X_t-\mu) \Big| X_{o_t} = x_{o_t}\right] =$$
$$-L\log(\pi) - \log(\operatorname{Det}(\mathbf{\Lambda})) +$$
$$+\mathbb{E}_{\Psi^{(\tau-1)}}\left[-(X_t-\mu)^H \mathbf{\Lambda}^{-1}(X_t-\mu) \Big| X_{o_t} = x_{o_t}\right] =$$
$$-L\log(\pi) - \log(\operatorname{Det}(\mathbf{\Lambda})) +$$
$$+\mathbb{E}_{\Psi^{(\tau-1)}}\left[-\begin{bmatrix} X_{m_t} - \mu_{X_{m_t}} \\ X_{o_t} - \mu_{X_{o_t}} \end{bmatrix}^H \mathbf{\Lambda}^{-1} \begin{bmatrix} X_{m_t} - \mu_{X_{m_t}} \\ X_{o_t} - \mu_{X_{o_t}} \end{bmatrix} \Big| X_{o_t} = x_{o_t}\right] =$$

$$-L\log(\pi)-\log(\mathrm{Det}(\mathbf{\Lambda}))+$$
$$+\mathbb{E}_{\Psi^{(\tau-1)}}\Bigg[-\begin{bmatrix}X_{m_t}-A_{m_t}S_t\\ X_{o_t}-A_{o_t}S_t\end{bmatrix}^H\begin{bmatrix}\mathbf{\Lambda}_{(m_t,m_t)} & \mathbf{O}\\ \mathbf{O} & \mathbf{\Lambda}_{(o_t,o_t)}\end{bmatrix}^{-1}\begin{bmatrix}X_{m_t}-A_{m_t}S_t\\ X_{o_t}-A_{o_t}S_t\end{bmatrix}\Bigg|X_{o_t}=x_{o_t}\Bigg]=$$
$$-L\log(\pi)-\log(\mathrm{Det}(\mathbf{\Lambda}))-$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\Bigg[(X_{m_t}-A_{m_t}S_t)^H(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}(X_{m_t}-A_{m_t}S_t)+$$
$$(X_{o_t}-A_{o_t}S_t)^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(X_{o_t}-A_{o_t}S_t)\Bigg|X_{o_t}=x_{o_t}\Bigg]=$$
$$-L\log(\pi)-\log(\mathrm{Det}(\mathbf{\Lambda}))-(X_{o_t}-A_{o_t}S_t)^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(X_{o_t}-A_{o_t}S_t)-$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\Bigg[(X_{m_t}-A_{m_t}S_t)^H(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}(X_{m_t}-A_{m_t}S_t)\Bigg|X_{o_t}=x_{o_t}\Bigg]=$$

Заметим, что:

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

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

$$\mathbb{E}_{\Psi^{(\tau-1)}}\Bigg[X_{m_t}^H(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}X_{m_t}\Bigg|X_{o_t}=x_{o_t}\Bigg]=$$
$$\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}\Bigg[\Big[[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]^{\frac{1}{2}}X_{m_t}\Big]^H\Big[[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]^{\frac{1}{2}}X_{m_t}\Big]\Bigg]=$$
$$\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}[W_t^HW_t]=$$
$$\mathrm{Tr}\left(\mathbb{E}_{\Psi^{(\tau-1)},X_{m_t}|X_{o_t}=x_{o_t}}[W_tW_t^H]\right)=$$
$$\mathrm{Tr}\left(\left[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\right]^{\frac{1}{2}}\mathbf{\Lambda}_{m_t}\left[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\right]^{\frac{1}{2}}\right)+$$
$$\mathrm{Tr}\left(\left[[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]^{\frac{1}{2}}(\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\right]\left[[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]^{\frac{1}{2}}(\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\right]^H\right)=$$
$$\mathrm{Tr}\left(\left[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\right]\mathbf{\Lambda}_{m_t}\right)+\mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})^H\left[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\right](\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\right)$$

$$\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}[\log P(X_{o_t},X_{m_t})]=$$
$$-L\log(\pi)-\log(\mathrm{Det}(\mathbf{\Lambda}))-(x_{o_t}-\mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t}-\mu_{X_{o_t}})-$$
$$\mathrm{Tr}\left([(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]\mathbf{\Lambda}_{m_t}\right)-\mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\right)$$

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

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

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

$$\sum_{t=1}^{G}\Big[-(x_{o_t}-\mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t}-\mu_{X_{o_t}})-\mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\right)\Big]=$$

$$\sum_{t=1}^{G}\Big[-(x_{o_t}-\mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t}-\mu_{X_{o_t}})-(\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\Big]=$$

$$=\sum_{t=1}^{G}\Big(-(\widetilde{x}_t^{(\tau)}-AS_t)^H\mathbf{\Lambda}^{-1}(\widetilde{x}_t^{(\tau)}-AS_t)\Big)=$$

$$-||\mathbf{\Lambda}^{-1/2}(\widetilde{x}^{(\tau)}-AS)||_F^2$$

.

M-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\Psi^{(\tau)}=\arg\max_{\Psi}\mathbb{E}_{X_m|X_o=x_o,\Psi^{(\tau-1)}}[\log P(X_o,X_m)]=$$

$$\arg\max_{\Psi}\sum_{t=1}^{G}\Big[-L\log(\pi)-\log(\mathrm{Det}(\mathbf{\Lambda}))-(x_{o_t}-\mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t}-\mu_{X_{o_t}})-$$

$$\mathrm{Tr}\left([(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]\mathbf{\Lambda}_{m_t}\right)-\mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\right)\Big]=$$

$$\arg\min_{\Psi}||\mathbf{\Lambda}^{-1/2}(\widetilde{x}^{(\tau)}-AS)||_F^2$$

Первый CM-шаг

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

$$\theta^{(\tau)}=\arg\max_{\theta}Q(\theta|\theta^{(\tau-1)})=$$

$$\arg\min_{\theta}||\mathbf{\Lambda}^{-1/2}(\widetilde{x}^{(\tau)}-AS)||_F^2$$

Второй CM-шаг

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

$$S^{\tau}=\arg\min_{S}(\widetilde{x}^{(\tau)}-AS)^H\mathbf{\Lambda}^{-1}(\widetilde{x}^{(\tau)}-AS) \tag{11}$$

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

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

§3 Неизвестный шум

Воспользуемся ECM-алгоритмом для того, чтобы определить значения параметров $\Psi = (\theta, S, \mathbf{\Lambda})$, 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

Оценим вектор углов $\theta^{(0)}$ следующим образом:

1. Выберем число ν, которое будет соответствовать первому компоненту вектора $\theta^{(0)}$:

$$\nu \sim \mathcal{U}([-\pi; \pi]); \tag{13}$$

2. Оценим компоненты вектора $\theta^{(0)}$ так: $\theta_i^{(0)} = (\nu + (i-1) \cdot \frac{2\pi}{M}) \bmod 2\pi, i = \overline{1, M}$. При этом, $a \bmod b = a - b \cdot \lfloor \frac{a}{b} \rfloor$.

Теперь следует составить первоначальную оценку детерминированных сигналов. Предположим, что детерминированные сигналы мы можем представить в форме:

$$S(t) = Be^{j\omega \cdot t + \phi}, \tag{14}$$

где B – амплитуда, ω – частота, ϕ – фаза. Предполагаем, что амплитуда, частота и фаза у сигналов от различных источников могут отличаться, генерируем эти параметры из равномерного распределения. На основе их строим оценку сигналов. Шумы на сенсорах задаются с учетом дисперсии остатков

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

.

E-шаг

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

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

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

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

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

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

$$P(X_o|\Psi) = \prod_{t=1}^{G} \frac{1}{\pi^{L_{o_t}} \operatorname{Det}(\mathbf{\Lambda}_{o_t})} e^{-(X_{o_t} - A_{o_t} S_t)^H (\mathbf{\Lambda}_{o_t})^{-1} (X_{o_t} - A_{o_t} S_t)}, \tag{19}$$

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

$$\begin{cases} \mu^{(\tau)}_{X_{m_t}|X_{o,t}=x_{o_t}} = \mu^{(\tau)}_{X_{m_t}} + \Sigma^{(\tau)}_{X_{m_t},X_{o_t}} \left(\Sigma^{(\tau)}_{X_{o_t},X_{o_t}}\right)^{-1} \cdot (x_{o_t} - \mu^{(\tau)}_{X_{o_t}}) \\ \Sigma^{(\tau)}_{X_{m_t}|X_{o_t}=x_{o,t}} = \Sigma^{(\tau)}_{X_{m_t},X_{m_t}} - \Sigma^{(\tau)}_{X_{m_t},X_{o_t}} \left(\Sigma^{(\tau)}_{X_{o_t},X_{o_t}}\right)^{-1} \Sigma^{(\tau)}_{X_{o_t},X_{m_t}} \end{cases} \tag{20}$$

В рамках данной задачи:

$$\begin{cases} \Sigma_{X_{o_t},X_{o_t}} = \mathbf{\Lambda}^{(\tau)}_{o_t} \\ \Sigma_{X_{o_t},X_{m_t}} = \hat{\Sigma}^{(\tau)}_{X_{o_t},X_{m_t}} \\ \Sigma_{X_{m_t},X_{o_t}} = \hat{\Sigma}^{(\tau)}_{X_{m_t},X_{o_t}} \\ \Sigma_{X_{m_t},X_{m_t}} = \mathbf{\Lambda}^{(\tau)}_{m_t} \\ \mu_{X_{o_t}} = A^{(\tau)}_{o_t} S^{(\tau)}_t \\ \mu_{X_{m_t}} = A^{(\tau)}_{m_t} S^{(\tau)}_t \end{cases} \tag{21}$$

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

$$\begin{cases} \mu^{(\tau)}_{X_{m_t}|X_{O_t}=x_{o_t}} = A^{(\tau)}_{m_t} S^{(\tau)}_t + \hat{\Sigma}^{(\tau)}_{X_{m_t},X_{o_t}} \left(\mathbf{\Lambda}^{(\tau)}_{o_t}\right)^{-1} \cdot (x_o - A^{(\tau)}_{o_t} S^{(\tau)}_t) \\ \Sigma^{(\tau)}_{X_{m_t}|X_{o_t}=x_{o,t}} = \mathbf{\Lambda}^{(\tau)}_{m_t} - \hat{\Sigma}^{(\tau)}_{X_{m_t},X_{o_t}} \left(\mathbf{\Lambda}^{(\tau)}_{o_t}\right)^{-1} \hat{\Sigma}^{(\tau)}_{X_{o_t},X_{m_t}} \end{cases} \tag{22}$$

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

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

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:

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

$$\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}[\log P(X_{o_t}, X_{m_t})] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\left[\log P\Big(X_{o_t}, X_{m_t}\Big| X_{o_t} = x_{o_t}\Big)\right] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\left[\log\left(\frac{1}{\pi^L \operatorname{Det}(\mathbf{\Lambda})} e^{-(X_t-\mu)^H \mathbf{\Lambda}^{-1}(X_t-\mu)}\right)\Big| X_{o_t} = x_{o_t}\right] =$$
$$\mathbb{E}_{\Psi^{(\tau-1)}}\left[-L\log(\pi) - \log(\operatorname{Det}(\mathbf{\Lambda})) - (X_t-\mu)^H \mathbf{\Lambda}^{-1}(X_t-\mu)\Big| X_{o_t} = x_{o_t}\right] =$$
$$-L\log(\pi) - \log(\operatorname{Det}(\mathbf{\Lambda})) +$$
$$+\mathbb{E}_{\Psi^{(\tau-1)}}\left[-(X_t-\mu)^H \mathbf{\Lambda}^{-1}(X_t-\mu)\Big| X_{o_t} = x_{o_t}\right] =$$
$$-L\log(\pi) - \log(\operatorname{Det}(\mathbf{\Lambda})) +$$
$$+\mathbb{E}_{\Psi^{(\tau-1)}}\left[-\begin{bmatrix} X_{m_t} - \mu_{X_{m_t}} \\ X_{o_t} - \mu_{X_{o_t}} \end{bmatrix}^H \mathbf{\Lambda}^{-1} \begin{bmatrix} X_{m_t} - \mu_{X_{m_t}} \\ X_{o_t} - \mu_{X_{o_t}} \end{bmatrix} \Big| X_{o_t} = x_{o_t}\right] =$$

$$-L\log(\pi) - \log(\mathrm{Det}(\mathbf{\Lambda}))+$$

$$+\mathbb{E}_{\Psi^{(\tau-1)}}\Bigg[-\begin{bmatrix} X_{m_t} - A_{m_t}S_t \\ X_{o_t} - A_{o_t}S_t \end{bmatrix}^H \begin{bmatrix} \mathbf{\Lambda}_{(m_t,m_t)} & \mathbf{O} \\ \mathbf{O} & \mathbf{\Lambda}_{(o_t,o_t)} \end{bmatrix}^{-1} \begin{bmatrix} X_{m_t} - A_{m_t}S_t \\ X_{o_t} - A_{o_t}S_t \end{bmatrix} \Big| X_{o_t} = x_{o_t}\Bigg] =$$

$$-L\log(\pi) - \log(\mathrm{Det}(\mathbf{\Lambda}))-$$

$$\mathbb{E}_{\Psi^{(\tau-1)}}\Bigg[(X_{m_t} - A_{m_t}S_t)^H(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}(X_{m_t} - A_{m_t}S_t)+$$

$$(X_{o_t} - A_{o_t}S_t)^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(X_{o_t} - A_{o_t}S_t)\Big|X_{m_t}|X_{o_t} = x_{o_t}\Bigg] =$$

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

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

Заметим, что:

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

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

$$\mathbb{E}_{\Psi^{(\tau-1)}}\Bigg[X_{m_t}^H(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}X_{m_t}\Big|X_{o_t} = x_{o_t}\Bigg] =$$

$$\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}\Bigg[\Big[[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]^{\frac{1}{2}}X_{m_t}\Big]^H\Big[[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]^{\frac{1}{2}}X_{m_t}\Big]\Bigg] =$$

$$\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}[W_t^H W_t] =$$

$$\mathrm{Tr}\left(\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}[W_t W_t^H]\right) =$$

$$\mathrm{Tr}\left(\left[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\right]^{\frac{1}{2}}\mathbf{\Lambda}_{m_t}\left[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\right]^{\frac{1}{2}}\right)+$$

$$\mathrm{Tr}\left(\left[[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]^{\frac{1}{2}}(\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})\right]\left[[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]^{\frac{1}{2}}(\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})\right]^H\right) =$$

$$\mathrm{Tr}\left(\left[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\right]\mathbf{\Lambda}_{m_t}\right) + \mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H\left[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\right](\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})\right)$$

$$\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}[\log P(X_{o_t}, X_{m_t})] =$$

$$-L\log(\pi) - \log(\mathrm{Det}(\mathbf{\Lambda})) - (x_{o_t} - \mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t} - \mu_{X_{o_t}})-$$

$$\mathrm{Tr}\left([(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]\mathbf{\Lambda}_{m_t}\right) - \mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})\right)$$

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

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

$$\sum_{t=1}^{G}\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}[\log P(X_{o_t}, X_{m_t})] =$$

$$\sum_{t=1}^{G}\Bigg[-L\log(\pi) - \log(\mathrm{Det}(\mathbf{\Lambda})) - (x_{o_t} - \mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t} - \mu_{X_{o_t}})-$$

$$\mathrm{Tr}\left([(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]\mathbf{\Lambda}_{m_t}\right) - \mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})\right)\Bigg] =$$

$$\sum_{t=1}^{G}\Bigg[-(x_{o_t} - \mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t} - \mu_{X_{o_t}}) - \mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)} - \mu_{X_{m_t}})\right)\Bigg]$$

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

$$\sum_{t=1}^{G}\Big[-(x_{o_t}-\mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t}-\mu_{X_{o_t}})-\mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\right)\Big]=$$

$$\sum_{t=1}^{G}\Big[-(x_{o_t}-\mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t}-\mu_{X_{o_t}})-(\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\Big]=$$

$$=\sum_{t=1}^{G}\Big(-(\widetilde{x}_t^{(\tau)}-AS_t)^H\mathbf{\Lambda}^{-1}(\widetilde{x}_t^{(\tau)}-AS_t)\Big)=$$

$$-||\mathbf{\Lambda}^{-1/2}(\widetilde{x}^{(\tau)}-AS)||_F^2$$

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\Psi^{(\tau)}=\arg\max_{\Psi}\mathbb{E}_{X_m|X_o=x_o,\Psi^{(\tau-1)}}[\log P(X_o,X_m)]=$$

$$\arg\max_{\Psi}\sum_{t=1}^{G}\Big[-L\log(\pi)-\log(\mathrm{Det}(\mathbf{\Lambda}))-(x_{o_t}-\mu_{X_{o_t}})^H(\mathbf{\Lambda}^{-1})_{(o_t,o_t)}(x_{o_t}-\mu_{X_{o_t}})-$$

$$\mathrm{Tr}\left([(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}]\mathbf{\Lambda}_{m_t}\right)-\mathrm{Tr}\left((\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})^H[(\mathbf{\Lambda}^{-1})_{(m_t,m_t)}](\widetilde{x}_{m_t}^{(\tau)}-\mu_{X_{m_t}})\right)\Big]=$$

$$\arg\min_{\Psi}||\mathbf{\Lambda}^{-1/2}(\widetilde{x}^{(\tau)}-AS)||_F^2$$

Первый СМ-шаг

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

$$\theta^{(\tau)}=\arg\max_{\theta}Q(\theta|\theta^{(\tau-1)})=$$

$$\arg\min_{\theta}||[\mathbf{\Lambda}^{(\tau-1)}]^{-1/2}(\widetilde{x}^{(\tau)}-A(\theta)S^{(\tau-1)})||_F^2$$

Второй СМ-шаг

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

$$S^{\tau}=\arg\min_{S}(\widetilde{x}_t^{(\tau)}-A^{(\tau)}S^{(\tau-1)})^H[\mathbf{\Lambda}^{(\tau-1)}]^{-1}(\widetilde{x}_t^{(\tau)}-A^{(\tau)}S^{(\tau-1)}) \tag{23}$$

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

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

Третий CM-шаг

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

$$\lambda_{jj}^{(\tau)} = \frac{1}{G}\sum_{t=1}^{G} |\widetilde{x}_{j,t}^{(\tau)} - (A^{(\tau)}S^{(\tau)})_{j,t}|^2 \tag{25}$$

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

$$\begin{aligned}
\nabla_{\Lambda^{-1}} Q(\mathbf{\Lambda} \mid \mathbf{\Lambda}^{(\tau-1)}) &= \sum_{t=1}^{G} \nabla_{\mathbf{\Lambda}^{-1}} Q_t(\mathbf{\Lambda} \mid \mathbf{\Lambda}^{(\tau-1)}) \\
&= \sum_{t=1}^{G} \frac{1}{2}\nabla_{\Lambda^{-1}} \log|\mathbf{\Lambda}^{-1}| - \frac{1}{2}\nabla_{\Lambda^{-1}} \operatorname{Tr}\left((\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\Sigma_{X_{m_t}|X_{o_t}}^{(\tau)}\right) \\
&\quad - \sum_{t=1}^{G} \frac{1}{2}\nabla_{\Lambda^{-1}} (\widetilde{x}_t^{(\tau)} - A^{(\tau)}S^{(\tau)})^H \mathbf{\Lambda}^{-1} (\widetilde{x}_t^{(\tau)} - A^{(\tau)}S^{(\tau)}) \\
&= \frac{1}{2}\sum_{t=1}^{G} \nabla_{\Lambda^{-1}} \log|\mathbf{\Lambda}^{-1}| - \nabla_{\Lambda^{-1}} \operatorname{Tr}\left((\mathbf{\Lambda}^{-1})_{(m_t,m_t)}\Sigma_{X_{m_t}|X_{o_t}}^{(\tau)}\right) \\
&\quad - \frac{1}{2}\sum_{t=1}^{G} \nabla_{\mathbf{\Lambda}^{-1}} \operatorname{Tr}\left((\widetilde{x}_t^{(\tau)} - A^{(\tau)}S^{(\tau)})^H \mathbf{\Lambda}^{-1} (\widetilde{x}_t^{(\tau)} - A^{(\tau)}S^{(\tau)})\right)
\end{aligned}$$

$$\begin{aligned}
\nabla_{\Lambda^{-1}} Q(\mathbf{\Lambda} \mid \mathbf{\Lambda}^{(\tau-1)}) &= \frac{1}{2}\Big[\sum_{t=1}^{G} \mathbf{\Lambda} - \widetilde{\mathbf{\Lambda}}_t^{(\tau)}\Big] - \frac{1}{2}\sum_{t=1}^{G} (\widetilde{x}_t^{(\tau)} - A^{(\tau)}S^{(\tau)})(\widetilde{x}_t^{(\tau)} - A^{(\tau)}S^{(\tau)})^H \\
&= \frac{1}{2}\Big[G\mathbf{\Lambda} - \sum_{t=1}^{G} \widetilde{\mathbf{\Lambda}}_t^{(\tau)}\Big] - \frac{1}{2}\sum_{t=1}^{G} (\widetilde{x}_t^{(\tau)} - A^{(\tau)}S^{(\tau)})(\widetilde{x}_t^{(\tau)} - A^{(\tau)}S^{(\tau)})^H
\end{aligned}$$

где $\widetilde{\mathbf{\Lambda}}_t^{(\tau)}$ — 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.

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

Учтем тот факт, что матрица шума оценивается как диагональная:

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

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

Список источников

1. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B (Methodol.) 1977, 39, 1–38

2. Maximum Likelihood Estimation via the ECM Algorithm: A General Framework Xiao-Li Meng; Donald B. Rubin Biometrika, Vol. 80, No. 2 (Jun., 1993), 267-278.