ECM-алгоритм, стохастическая модель сигнала

§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, 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 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:

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

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, также как и шумы на сенсорах, предполагаются стохастическими: $S_t \sim CN(\mathbf{O}_{M\times 1}, \mathbf{P}), t = \overline{1,G}$, $N_t \sim CN(\mathbf{O}_{L\times 1}, \mathbf{\Lambda})$. Матрицы $\mathbf{P}$ и $\mathbf{\Lambda}$ предполагаются диагональными, т.е. и сигналы, и шумы, являются некоррелированными. 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 $\Psi = (\theta, \mathbf{P})$, 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

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

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

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

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

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

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

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

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

В рамках данной задачи:

$$\begin{cases} \Sigma_{x_{o_t},x_{o_t}} = A_{ot}\mathbf{P}A_{ot}^H + \mathbf{\Lambda}_{ot} \\ \Sigma_{x_{o_t},x_{m_t}} = A_{ot}\mathbf{P}A_{mt}^H \\ \Sigma_{x_{m_t},x_{o_t}} = A_{mt}\mathbf{P}A_{ot}^H \\ \Sigma_{x_{m_t},x_{m_t}} = A_{mt}\mathbf{P}A_{mt}^H + \mathbf{\Lambda}_{mt} \\ \mu_{x_{o_t}} = \mathbf{O}_{L_{o_t}\times 1} \\ \mu_{x_{m_t}} = \mathbf{O}_{L_{m_t}\times 1} \end{cases} \tag{7}$$

$$\begin{cases} \mu_{x_{m_t}|x_{o_t}} = A_{mt}\mathbf{P}A_{ot}^H(A_{ot}\mathbf{P}A_{ot}^H + \mathbf{\Lambda}_{ot})^{-1} \cdot x_{o_t} \\ \Sigma_{x_{m_t}|x_{o_t}} = A_{mt}\mathbf{P}A_{mt}^H + \mathbf{\Lambda}_{mt} - A_{mt}\mathbf{P}A_{ot}^H(A_{ot}\mathbf{P}A_{ot}^H + \mathbf{\Lambda}_{ot})^{-1}A_{ot}\mathbf{P}A_{mt}^H \end{cases} \tag{8}$$

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:

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

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:

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

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

$$-L\log(\pi)-\log\Big(\operatorname{Det}(\Sigma_{X_t|(X_{m_t}|X_{o_t}=x_{o_t})})\Big)-$$
$$-\mathbb{E}_{\Psi^{(\tau-1)}}\Big[X_{m_t}^H(\Sigma^{-1})_{(m_t,m_t)}X_{m_t}+2X_{o_t}^H(\Sigma^{-1})_{(m_t,o_t)}X_{m_t}+X_{o_t}^H(\Sigma^{-1})_{(m_t,m_t)}X_{o_t}\Big|X_{m_t}|X_{o_t}=x_{o_t}\Big] =$$
$$-L\log(\pi)-\log\Big(\operatorname{Det}(\Sigma_{X_t|(X_{m_t}|X_{o_t}=x_{o_t})})\Big)-$$
$$-x_{o_t}^H(\Sigma^{-1})_{(o_t,o_t)}x_{o_t}-2x_{o_t}^H(\Sigma^{-1})_{(m_t,o_t)}\hat{x}_{m_t}^{(\tau)}-\mathbb{E}_{\Psi^{(\tau-1)}}\Big[X_{m_t}^H(\Sigma^{-1})_{(m_t,m_t)}X_{m_t}\Big|X_{m_t}|X_{o_t}=x_{o_t}\Big]$$

Заметим, что:

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

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

$\mathbb{E}(W)\mathbb{E}(W)^H$.

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

$$\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau-1)}}[\log P(X_{o_t},X_{m_t})]=$$
$$-L\log(\pi)-\log\Big(\mathrm{Det}(\Sigma_{X_t|(X_{m_t}|X_{o_t}=x_{o_t})})\Big)-x_{o_t}^H(\Sigma^{-1})_{(o_t,o_t)}x_{o_t}-2x_{o_t}^H(\Sigma^{-1})_{(m_t,o_t)}\hat{x}_{mt}^{(\tau)}-$$
$$\mathrm{Tr}\left([(\Sigma^{-1})_{(m_t,m_t)}]\Sigma_{x_{m_t}|x_{o_t}}^{(\tau)}\right)-\mathrm{Tr}\left((\hat{x}_{mt}^{(\tau)})^H[(\Sigma^{-1})_{(m_t,m_t)}]\hat{x}_{mt}^{(\tau)}\right)$$

Таким образом, оптимизируемая функция примет следующий вид:

$$\mathbb{E}_{X_m|X_o=x_o,\Psi^{(\tau)}}[\log P(X_o,X_m)]=$$
$$\sum_{t=1}^{G}\mathbb{E}_{X_{m_t}|X_{o_t}=x_{o_t},\Psi^{(\tau)}}[\log P(X_{o_t},X_{m_t})]=$$
$$\sum_{t=1}^{G}\Bigg[-L\log(\pi)-\log\Big(\mathrm{Det}(\Sigma_{X_t|(X_{m_t}|X_{o_t}=x_{o_t})})\Big)-x_{o_t}^H(\Sigma^{-1})_{(o_t,o_t)}x_{o_t}-2x_{o_t}^H(\Sigma^{-1})_{(m_t,o_t)}\hat{x}_{mt}^{(\tau)}-$$
$$\mathrm{Tr}\left([(\Sigma^{-1})_{(m_t,m_t)}]\Sigma_{x_{m_t}|x_{o_t}}^{(\tau)}\right)-\mathrm{Tr}\left((\hat{x}_{mt}^{(\tau)})^H[(\Sigma^{-1})_{(m_t,m_t)}]\hat{x}_{mt}^{(\tau)}\right)\Bigg]$$

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\Psi^{(\tau)}=\underset{\Psi}{\mathrm{argmax}}\,\mathbb{E}_{X_m|X_o=x_o,\Psi^{(\tau-1)}}[\log P(X_o,X_m)]=$$
$$\underset{\Psi}{\mathrm{argmax}}\sum_{t=1}^{G}\Bigg[-\log\Big(\mathrm{Det}(\Sigma_{X_t|(X_{m_t}|X_{o_t}=x_{o_t})})\Big)-x_{o_t}^H(\Sigma^{-1})_{(o_t,o_t)}x_{o_t}-2x_{o_t}^H(\Sigma^{-1})_{(m_t,o_t)}\hat{x}_{mt}^{(\tau)}-$$
$$\mathrm{Tr}\left([(\Sigma^{-1})_{(m_t,m_t)}]\Sigma_{x_{m_t}|x_{o_t}}^{(\tau)}\right)-\mathrm{Tr}\left((\hat{x}_{mt}^{(\tau)})^H[(\Sigma^{-1})_{(m_t,m_t)}]\hat{x}_{mt}^{(\tau)}\right)\Bigg]$$

Первый СМ-шаг

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

$$\theta^{(\tau)}=\underset{\theta}{\mathrm{argmax}}\,Q(\theta|\theta^{(\tau-1)})=$$
$$\underset{\theta}{\mathrm{argmax}}\sum_{t=1}^{G}\Bigg[-\log\Big(\mathrm{Det}(\Sigma_{X_t|(X_{m_t}|X_{o_t}=x_{o_t})})\Big)-x_{o_t}^H(\Sigma^{-1})_{(o_t,o_t)}x_{o_t}-2x_{o_t}^H(\Sigma^{-1})_{(m_t)(o_t)}\hat{x}_{mt}^{(\tau)}-$$
$$\mathrm{Tr}\left([(\Sigma^{-1})_{(m_t,m_t)}]\Sigma_{x_{m_t}|x_{o_t}}^{(\tau)}\right)-\mathrm{Tr}\left((\hat{x}_{mt}^{(\tau)})^H[(\Sigma^{-1})_{(m_t,m_t)}]\hat{x}_{mt}^{(\tau)}\right)\Bigg]$$

Второй CM-шаг

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

$$\mathbf{P}^{(\tau)} = \underset{\mathbf{P}}{\operatorname{argmax}}\, Q(\mathbf{P}|\mathbf{P}^{(\tau-1)}) =$$

$$\underset{\mathbf{P}}{\operatorname{argmax}} \sum_{t=1}^{G} \Bigg[-\log\Big(\operatorname{Det}(\Sigma_{X_t|(X_{m_t}|X_{o_t}=x_{o_t})})\Big) - x_{o_t}^H (\Sigma^{-1})_{(o_t,o_t)} x_{o_t} - 2x_{o_t}^H (\Sigma^{-1})_{(m_t)(o_t)} \hat{x}_{m_t}^{(\tau)} -$$

$$\operatorname{Tr}\left([(\Sigma^{-1})_{(m_t,m_t)}] \Sigma^{(\tau)}_{x_{m_t}|x_{o_t}} \right) - \operatorname{Tr}\left((\hat{x}_{m_t}^{(\tau)})^H [(\Sigma^{-1})_{(m_t,m_t)}] \hat{x}_{m_t}^{(\tau)} \right) \Bigg]$$

§3 Неизвестный шум

Е-шаг

М-шаг

Первый СМ-шаг

Второй СМ-шаг

Список источников