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, 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 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 $X_o = \{X_{ot}\}_{t=1}^G$ и ненаблюдаемой: $X_m = \{X_{m_t}\}_{t=1}^G$. Полученный сигнал 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, 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. Наблюдения X_t, $t = \overline{1, L}$ предполагаются независимыми и одинаково распределенными.

E-шаг

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

$$\mathbb{E}_{(X_m,S)|X_o=x_o,\Psi^{(\tau-1)}}[\log P(X,S)]. \tag{2}$$

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

$$P((X_m, S)|X_o = x_o, \Psi) = P(X_m|X_o = x_o, \Psi) \cdot P(S|X_o = x_o, X_m = \hat{x}_m, \Psi) \tag{3}$$

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

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

$$P(X|\Psi) = \prod_{t=1}^{G} \frac{1}{\pi^L \operatorname{Det}(A\mathbf{P}A^H + \mathbf{\Lambda})} e^{-X_t^H (A\mathbf{P}A^H + \mathbf{\Lambda})^{-1} X_t}, \tag{5}$$

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

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

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

$$\begin{cases} \mu_{X_{m_t}|x_{o_t}} = \hat{\Sigma}_{x_{m_t},x_{o_t}} \hat{\Sigma}_{x_{o_t},x_{o_t}}^{-1} \cdot x_{o_t} \\ \Sigma_{x_{m_t}|x_{o_t}} = \hat{\Sigma}_{x_{m_t},x_{m_t}} - \hat{\Sigma}_{x_{m_t},x_{o_t}} \hat{\Sigma}_{x_{o_t},x_{o_t}}^{-1} \hat{\Sigma}_{x_{o_t},x_{m_t}} \end{cases} \tag{8}$$

Оцениваем пропущенные значения для каждого наблюдения через условное математическое ожидание: $\widetilde{x}_{m_t} = \mu_{x_{m_t}|x_{o_t}}$. $\widetilde{x}_t^{(\tau)}$ – оценка x_t с учетом оценки пропусков. Параметры апостериорного распределения $P(S_t|X_{o_t} = x_{o_t}, X_{m_t} = \widetilde{x}_{m_t}, \Psi)$ можно найти исходя из следующих формул

$$\begin{cases} \mu_{S_t|X_{o_t}=x_{o_t},X_{m_t}=\widetilde{x}_{m_t},\Psi} = \mathbf{P}A^H(A\mathbf{P}A^H + Q)^{-1}\widetilde{x}_t^{(\tau)} \\ \Sigma_{S_t|X_{o_t}=x_{o_t},X_{m_t}=\widetilde{x}_{m_t},\Psi} = \mathbf{P} - \mathbf{P}A^H(A\mathbf{P}A^H + Q)^{-1}A\mathbf{P} \end{cases} \tag{9}$$

Рассчитаем оценку пространственной ковариационной матрицы (пользуемся результатами выкладок по детерминированной модели):

$$\hat{R}^{(\tau)} = \frac{1}{G}\sum_{t=1}^{G}\left[\widetilde{\Sigma}_t^{(\tau)} + \widetilde{x}_t^{(\tau)}(\widetilde{x}_t^{(\tau)})^H\right] \tag{10}$$

где $\widetilde{\Sigma_t}^{(\tau)}$ — 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 $\Sigma^{(\tau)}_{x_{m_t}|x_{o_t}}$. Вернемся к ранее рассмотренному условному математическому ожиданию:

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

Его следует максимизировать, мы можем перейти от логарифма произведения к сумме логарифмов.

$$\log P(X,S|\theta,\mathbf{P}) = -G\log(\mathrm{Det}(\pi\mathbf{\Lambda})) - \sum_{t=1}^{G}(X_t - A(\theta)S_t)^H\mathbf{\Lambda}^{-1}(X_t - A(\theta)S_t)$$
$$-G\log(\mathrm{Det}(\pi\mathbf{P})) - \sum_{t=1}^{G} S_t^H\mathbf{P}^{-1}S_t$$

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\Psi^{(\tau)} = \underset{\Psi}{\operatorname{argmax}}\,\mathbb{E}_{(X_m,S)|X_o=x_o,\Psi^{(\tau-1)}}[\log P(X,S)] =$$
$$\underset{\Psi}{\operatorname{argmax}}\,\mathbb{E}_{(X_m,S)|X_o=x_o,\Psi^{(\tau-1)}}\Big[-G\log(\mathrm{Det}(\pi\mathbf{\Lambda})) - \sum_{t=1}^{G}(X_t - A(\theta)S_t)^H\mathbf{\Lambda}^{-1}(X_t - A(\theta)S_t)$$
$$-G\log(\mathrm{Det}(\pi\mathbf{P})) - \sum_{t=1}^{G} S_t^H\mathbf{P}^{-1}S_t\Big]$$

Первый СМ-шаг

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

$$\theta^{(\tau)} = \underset{\theta}{\operatorname{argmax}}\,Q(\theta|\theta^{(\tau-1)}) =$$
$$\underset{\theta}{\operatorname{argmax}}\,\mathbb{E}_{(X_m,S)|X_o=x_o,\Psi^{(\tau-1)}}\Big[-G\log(\mathrm{Det}(\pi\mathbf{\Lambda})) - \sum_{t=1}^{G}(X_t - A(\theta)S_t)^H\mathbf{\Lambda}^{-1}(X_t - A(\theta)S_t)$$
$$-G\log(\mathrm{Det}(\pi\mathbf{P})) - \sum_{t=1}^{G} S_t^H\mathbf{P}^{-1}S_t\Big]$$

Тогда минимизируемая функция примет следующий вид:

$$\mathcal{J}(\theta) = \sum_{t=1}^{G}\mathbb{E}\Big[(X_t - A(\theta)S_t)^H\mathbf{\Lambda}^{-1}(X_t - A(\theta)S_t)|X_o = x_o\Big] =$$
$$\sum_{t=1}^{G}\mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\mathbb{E}\Big[(X_t - A(\theta)S_t)(X_t - A(\theta)S_t)^H\Big|X_o = x_o\Big]\right) =$$
$$\mathrm{Tr}\left(\mathbf{\Lambda}^{-1}(\hat{R} - A\mathbf{P}^{\tau-1}A^H)\right)$$

$$\underset{\theta}{\operatorname{argmin}} = \operatorname{Tr}\left(\mathbf{\Lambda}^{-1}(\hat{R} - A\mathbf{P}^{\tau-1}A^H)\right) \tag{11}$$

Второй CM-шаг

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

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

Пользуемся тем фактом, что полное правдоподобие раскладывается на сумму $\log P(X|S = s) + \log(S)$. Первый логарифм не зависит от $\mathbf{P}$. Поэтому максимизируем условное математическое ожидание для $\log(S|\Psi)$.

$$\mathcal{K}(\mathbf{P}) = \mathbb{E}_{S|X=\hat{x},\Psi^{(\tau-1)}}\left[-G\log(\operatorname{Det}(\pi\mathbf{P})) - \sum_{t=1}^{G} S_t^H\mathbf{P}^{-1}S_t\right] =$$

$$-G\log(\operatorname{Det}(\pi\mathbf{P})) - \mathbb{E}_{\Psi^{(\tau-1)}}\left[\sum_{t=1}^{G} S_t^H\mathbf{P}^{-1}S_t\Big|X = \widetilde{x}^{(\tau)}\right] =$$

$$-G\log(\operatorname{Det}(\pi)) - G\log(\operatorname{Det}(\mathbf{P})) - \mathbb{E}_{\Psi^{(\tau-1)}}\left[\sum_{t=1}^{G} S_t^H\mathbf{P}^{-1}S_t\Big|X = \widetilde{x}^{(\tau)}\right] =$$

$$-G\log(\operatorname{Det}(\mathbf{P})) - \sum_{t=1}^{G}\operatorname{Tr}\left(\mathbf{P}^{-1}\mathbb{E}[S_tS_t^H\Big|X = \widetilde{x}^{(\tau)}]\right)$$

$$\frac{d}{d\mathbf{P}}\log(\operatorname{Det}(\mathbf{P})) = P^{-1}$$

$X = \widetilde{x}^{(\tau)}$ – оценка наблюдений, полученная с учетом E-шага текущей итерации. Обозначим через M величину $\mathbb{E}[S_tS_t^H\Big|X = \widetilde{x}^{(\tau)}]$.

$$\frac{d}{d\mathbf{P}}\operatorname{Tr}\big(P^{-1}M\big) = -P^{-1}MP^{-1}$$

$$\frac{d\mathcal{K}(\mathbf{P})}{d\mathbf{P}} = -GP^{-1} + P^{-1}MP^{-1}$$

Приравняем производную к нулю (функция по $\mathbf{P}$ выпукла).

$$O = -GP^{-1} + P^{-1}MP^{-1} \Rightarrow M = GP \Rightarrow \mathbf{P}^{(\tau)} = \frac{1}{G}\sum_{t=1}^{G}\left(\Sigma_{S_t|X_t} + \mu_{S_t|X_t}(\mu_{S_t|X_t})^H\right)$$

$$\mathbf{P}^{(\tau)} = \frac{1}{G}\sum_{t=1}^{G}\left(\Sigma_{S_t|X_t} + \mu_{S_t|X_t}(\mu_{S_t|X_t})^H\right) \tag{12}$$

§3 Неизвестный шум

Е-шаг

М-шаг

Первый СМ-шаг

Второй СМ-шаг

Список источников