

ЕМ-алгоритм для оценки направления прибытия сигнала

Введем некоторые условные обозначения:

- θ – вектор направлений прибытия сигнала (DOA);
- τ – итерация ЕМ-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 сигналов (случайная величина), испускаемых источниками в моменты времени $t = \overline{1, G}$, S_t соответствует сигналу в момент времени t ;
- s – набор сигналов (реализация), испускаемых источниками в моменты времени $t = \overline{1, G}$, s_t соответствует сигналу в момент времени t ;
- N – набор шумов (случайная величина), связанных с датчиками в моменты времени $t = \overline{1, G}$, N_t соответствует шуму в момент времени t ;
- 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 ;
- 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 ;
- Y – латентные переменные (случайная величина) (X_m, S) ;
- y – латентные переменные (реализация) (x_m, s) ;
- ψ – параметры комплексного нормального распределения X_m ;
- $\Omega = (\psi, \theta)$;
- $O_{D_1 \times D_2}$ – 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 $S_t \sim CN(0, \Gamma_s), t = \overline{1, G}$, $N_t \sim CN(0, \Gamma_n), t = \overline{1, G}$, S_t имеет размер $M \times 1$, N_t имеет размер $L \times 1$, $\theta = [\theta_1, \dots, \theta_M]$ – вектор направлений прибытия сигнала, $A(\theta)$ (далее – A) представляет собой матрицу управляющих векторов размера $L \times M$, Γ_s и Γ_n предполагаются диагональными.

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

Воспользуемся ЕМ-алгоритмом для того, чтобы определить значения параметров θ , значения сигналов $S_t, t = \overline{1, G}$ рассматриваются как латентные переменные. Пусть X , S и N набор итоговых сигналов полученных L датчиками за моменты времени $t = \overline{1, G}$ и набор выпущенных M источниками сигналов и набор шумов за моменты времени $t = \overline{1, G}$, соответственно. X , S и N представляют из себя матрицы размеров $G \times L$, $G \times M$ и $G \times L$ соответственно.

Е-шаг

Требуется найти апостериорное распределение $P(Y|X_o = x_o, \theta)$, воспользуемся формулой Байеса:

$$P(Y|X_o = x_o, \theta) = P(X_m, S|X_o = x_o, \theta) = \frac{P(X, S|\theta)}{P(X_m, S|\theta)} = \frac{P(X_o, X_m, S|\theta)}{P(X_m, S|\theta)} = \frac{P(X_o, X_m, S|\theta)}{P(X_m|S = s, \theta)P(S|\theta)} \quad (2)$$

$$P(S|\theta) = \prod_{t=1}^G \frac{1}{\pi^M |\Gamma_s|} e^{-S_t^H \Gamma_s^{-1} S_t}, \quad (3)$$

$$\begin{aligned} X_t &= AS_t + N_t \\ X_t &\sim CN(0, A\Gamma_s A^H + \Gamma_n) \end{aligned}$$

$$P(X|\theta) = \prod_{t=1}^G \frac{1}{\pi^L |A\Gamma_s A^H + \Gamma_n|} e^{-X_t^H (A\Gamma_s A^H + \Gamma_n)^{-1} X_t}, \quad (4)$$

Теперь следует определиться с тем, каким будет условное распределение $P(X|S = s, \theta)$

$$X_t|S_t = s_t, \theta \sim CN(As_t, \Gamma_n)$$

$$P(X|S = s, \theta) = \prod_{t=1}^G \frac{1}{\pi^L |\Gamma_n|} e^{-(X_t - As_t)^H \Gamma_n^{-1} (X_t - As_t)}, \quad (5)$$

Теперь следует найти $P(X_m|S = s, \theta)$. Введем новые обозначения: пусть A_o – матрица, образованная теми строками матрицы A , которые соответствуют работающим сенсорам. A_m – матрица, образованная теми строками матрицы A , которые соответствуют неисправным сенсорам. Γ_m –

ковариационная матрица шума на неисправных сенсорах, Γ_o – ковариационная матрица шума на исправных сенсорах, L_1 – число исправных сенсоров, L_2 – число неисправных сенсоров.

$$P(X_m|S = s, \theta) = \prod_{t=1}^G \frac{1}{\pi^{L_2} |\Gamma_m|} e^{-(X_{m,t} - A_m s_t)^H \Gamma_m^{-1} (X_{m,t} - A_m s_t)}, \quad (6)$$

$$\begin{cases} \Sigma_{Y|X_o} = \Sigma_{Y,Y} - \Sigma_{Y,X_o} \Sigma_{X_o,X_o}^{-1} \Sigma_{X_o,Y} \\ \mu_{Y|X_o} = \mu_Y + \Sigma_{Y,X_o} \Sigma_{X_o,X_o}^{-1} \cdot (x_o - \mu_{X_o}), \end{cases} \quad (7)$$

$$\Sigma_{X_o,X_o} = A_o \Gamma_s A_o^H + \Gamma_o \quad (8)$$

$$\Sigma_{Y,Y} = \begin{pmatrix} \Sigma_{X_m,X_m} & \Sigma_{X_m,S} \\ \Sigma_{S,X_m} & \Sigma_{S,S} \end{pmatrix} \quad (9)$$

$$\Sigma_{S,S} = \Gamma_s \quad (10)$$

$$\Sigma_{X_m,X_m} = A_m \Gamma_s A_m^H + \Gamma_m \quad (11)$$

$$\Sigma_{X_m,S} = A_m \Gamma_s, \Sigma_{S,X_m} = \Gamma_s A_m^H \quad (12)$$

$$\Sigma_{Y,Y} = \begin{pmatrix} A_m \Gamma_s A_m^H + \Gamma_m & A_m \Gamma_s \\ \Gamma_s A_m^H & \Gamma_s \end{pmatrix} \quad (13)$$

$$\Sigma_{X_o,Y} = \begin{pmatrix} \Sigma_{X_o,X_m} \\ \Sigma_{X_o,S} \end{pmatrix} = \begin{pmatrix} A_o \Gamma_s A_m^H \\ A_o \Gamma_s \end{pmatrix} \quad (14)$$

$$\Sigma_{Y,X_o} = \Sigma_{X_o,Y}^H \quad (15)$$

$$\begin{cases} \Sigma_{Y|X_o} = \begin{pmatrix} \Sigma_{X_m,X_m} & \Sigma_{X_m,S} \\ \Sigma_{S,X_m} & \Sigma_{S,S} \end{pmatrix} - \begin{pmatrix} \Sigma_{X_o,X_m} \\ \Sigma_{X_o,S} \end{pmatrix}^H (\Sigma_{X_o,X_o})^{-1} \begin{pmatrix} \Sigma_{X_o,X_m} \\ \Sigma_{X_o,S} \end{pmatrix} \\ \mu_{Y|X_o} = \begin{pmatrix} \Sigma_{X_o,X_m} \\ \Sigma_{X_o,S} \end{pmatrix}^H (\Sigma_{X_m,X_m})^{-1} \cdot x_o, \end{cases} \quad (16)$$

$$\begin{cases} \Sigma_{Y|X_o} = \begin{pmatrix} A_m \Gamma_s A_m^H + \Gamma_m & A_m \Gamma_s \\ \Gamma_s A_m^H & \Gamma_s \end{pmatrix} - \begin{pmatrix} A_o \Gamma_s A_m^H \\ A_o \Gamma_s \end{pmatrix}^H (A_o \Gamma_s A_o^H + \Gamma_o)^{-1} \begin{pmatrix} A_o \Gamma_s A_m^H \\ A_o \Gamma_s \end{pmatrix} \\ \mu_{Y|X_o} = \begin{pmatrix} A_o \Gamma_s A_m^H \\ A_o \Gamma_s \end{pmatrix}^H (A_o \Gamma_s A_o^H + \Gamma_o)^{-1} \cdot x_o, \end{cases} \quad (17)$$

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\theta^{(\tau+1)} = \arg \max_{\theta} E[\log P(X_o, X_m, S|\theta^{(\tau)})|X_o, \theta^{(\tau)}] \quad (18)$$

Обозначим приведенное выше условное математическое ожидание через Q . Заметим, что $\log P(X_m, X_o, S|\theta^{(\tau)}) = \log P(X_o|X_m, S, \theta^{(\tau)})P(X_m, S|\theta^{(\tau)}) = \log P(X_o|Y, \theta^{(\tau)})P(Y|\theta^{(\tau)})$.

Можно заметить, что $P(X_o|Y, Y) = P(X_o, Y)$. Работать с плотностью $P(X_o|Y, Y)$ удобнее: кросс-ковариация между Y и $X_o|Y$ будет представлять из себя нулевую матрицу. Найдём совместную плотность $P(X_o|Y, Y|\theta^{(\tau)})$:

$$P(X_o|Y, Y|\theta^{(\tau)}) = \prod_{t=1}^G \frac{1}{\pi^{M+L}|\Sigma|} e^{-(Z_t - \mu_{Z_t})^H \Sigma^{-1} (Z_t - \mu_{Z_t})}, \quad (19)$$

где:

$$Z_t = \begin{pmatrix} Y_t \\ X_{o,t}|Y_t \end{pmatrix}, \mu_{Y_t} = \begin{pmatrix} \mu_{Y_t} \\ \mu_{X_{o,t}|Y_t} \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{Y_t, Y_t} & O_{(M+L_2) \times (L-L_2)} \\ O_{(L-L_2) \times (M+L_2)} & \Sigma_{X_{o,t}|Y_t, X_{o,t}|Y_t} \end{pmatrix}$$

Найдём логарифм совместной плотности (т.е. полное правдоподобие) $\log P(X_o|Y, Y|\theta^{(\tau)})$:

$$\log P(X_o|Y, Y|\theta^{(\tau)}) = \sum_{t=1}^G (-(M+L) \log(\pi) - \log |\Sigma| - (Z_t - \mu_{Z_t})^H \Sigma^{-1} (Z_t - \mu_{Z_t}))$$

Теперь попробуем раскрыть УМО полного правдоподобия для одного наблюдения, с учетом полученных сигналов и текущей оценки DOA:

$$P(X_{o,t}|Y_t, Y_t|\theta^{(\tau)}) = \frac{1}{\pi^{M+L}|\Sigma|} e^{-(Z_t - \mu_{Z_t})^H \Sigma^{-1} (Z_t - \mu_{Z_t})} \quad (20)$$

$$\begin{aligned} Q_t &= E[\log P(X_{o,t}|Y_t, Y_t|\theta^{(\tau)})|X_{o,t}, \theta^{(\tau)}] \\ &= E[-(M+L) \log(\pi) - \log |\Sigma| - (Z_t - \mu_{Z_t})^H \Sigma^{-1} (Z_t - \mu_{Z_t})|X_{o,t} = x_{o,t}, \theta^{(\tau)}] = \\ &= -(M+L) \log(\pi) - \log |\Sigma| - E[(Z_t - \mu_{Z_t})^H \Sigma^{-1} (Z_t - \mu_{Z_t})|X_{o,t} = x_{o,t}, \theta^{(\tau)}] \\ &= -(M+L) \log(\pi) - \log |\Sigma| - E[(X_{o,t}|Y_t - \mu_{X_{o,t}|Y_t})^H \Sigma_{X_{o,t}|Y_t, X_{o,t}|Y_t}^{-1} (X_{o,t}|Y_t - \mu_{X_{o,t}|Y_t})|X_{o,t} = x_{o,t}, \theta^{(\tau)}] \\ &\quad - E[(Y_t - \mu_{Y_t})^H \Sigma_{Y_t, Y_t}^{-1} (Y_t - \mu_{Y_t})|X_{o,t} = x_{o,t}, \theta^{(\tau)}] \\ &= -(M+L) \log(\pi) - \log |\Sigma| - (x_t - \mu_{X_{o,t}|Y_t})^H \Sigma_{X_{o,t}|Y_t, X_{o,t}|Y_t}^{-1} (x_t - \mu_{X_{o,t}|Y_t}) \\ &\quad - E[(Y_t - \mu_{Y_t})^H \Sigma_{Y_t, Y_t}^{-1} (Y_t - \mu_{Y_t})|X_{o,t} = x_{o,t}, \theta^{(\tau)}] \\ Q_t &= -(M+L) \log(\pi) - \log |\Sigma| - (x_t - \mu_{X_{o,t}|Y_t})^H \Sigma_{X_{o,t}|Y_t, X_{o,t}|Y_t}^{-1} (x_t - \mu_{X_{o,t}|Y_t}) \\ &\quad - E[(Y_t - \mu_{Y_t})^H \Sigma_{Y_t, Y_t}^{-1} (Y_t - \mu_{Y_t})|X_{o,t} = x_{o,t}, \theta^{(\tau)}] \end{aligned} \quad (21)$$

Заметим, что первые два слагаемые, составляющие Q_t , не зависят от θ , соответственно требуемый $\argmax[\cdot]$ можно найти без их учета.

$$Y_t|(X_{o,t} = x_{o,t}, \theta^{(\tau)}) \sim CN(\mu_{Y_t|X_{o,t}}, K_{Y_t|X_{o,t}}) \Rightarrow Y_t - \mu_{Y_t}|(X_{o,t} = x_{o,t}, \theta^{(\tau)}) \sim CN(\mu_{Y_t|X_{o,t}} - \mu_{Y_t}, K_{Y_t|X_{o,t}})$$

$$\Rightarrow [\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} (Y_t - \mu_{Y_t})|(X_{o,t} = x_{o,t}, \theta^{(\tau)}) \sim CN([\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} (\mu_{Y_t|X_{o,t}} - \mu_{Y_t}), [\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} K_{Y_t|X_{o,t}} ([\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}})^H)$$

Учтем, что для комплексных векторов V выполняется следующее соотношение:

$$E[VV^H] = E[V]E[V^H] + K_{VV}. \quad (22)$$

$$\begin{aligned} E[(Y_t - \mu_{Y_t})^H \Sigma_{Y_t, Y_t}^{-1} (Y_t - \mu_{Y_t})|X_{o,t} = x_{o,t}, \theta^{(\tau)}] &= E[[[\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} (Y_t - \mu_{Y_t})]^H [\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} (Y_t - \mu_{Y_t})|X_{o,t} = x_{o,t}, \theta^{(\tau)}] = \\ &= tr(E[[\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} (Y_t - \mu_{Y_t}) [[\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} (Y_t - \mu_{Y_t})]^H |X_{o,t} = x_{o,t}, \theta^{(\tau)}) = \\ &= tr([\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} K_{Y_t|X_{o,t}} ([\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}})^H) + tr([\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} (\mu_{Y_t|X_{o,t}} - \mu_{Y_t}) [[\Sigma_{Y_t, Y_t}^{-1}]^{\frac{1}{2}} (\mu_{Y_t|X_{o,t}} - \mu_{Y_t})]^H) = \\ &= tr([\Sigma_{Y_t, Y_t}^{-1}] K_{Y_t|X_{o,t}}) + (\mu_{Y_t|X_{o,t}} - \mu_{Y_t})^H \Sigma_{Y_t, Y_t}^{-1} (\mu_{Y_t|X_{o,t}} - \mu_{Y_t}) \end{aligned}$$

$$Q_t = -(M + L) \log(\pi) - \log |\Sigma| - (x_{o,t} - \mu_{X_{o,t}|Y_t})^H \Sigma_{X_{o,t}|Y_t, X_{o,t}|Y_t}^{-1} (x_{o,t} - \mu_{X_{o,t}|Y_t}) - \text{tr}([\Sigma_{Y_t, Y_t}^{-1}] K_{Y_t|X_{o,t}}) - (\mu_{Y_t|X_{o,t}} - \mu_{Y_t})^H \Sigma_{Y_t, Y_t}^{-1} (\mu_{Y_t|X_{o,t}} - \mu_{Y_t}) \quad (23)$$

$$E[\log P(X_o, Y|\theta^{(\tau)})|X_{o,t}, \theta^{(\tau)}] = \sum_{t=1}^G Q_t = \sum_{t=1}^G (-(M + L) \log(\pi) - \log |\Sigma| - (x_{o,t} - \mu_{X_{o,t}|Y_t})^H \Sigma_{X_{o,t}|Y_t, X_{o,t}|Y_t}^{-1} (x_{o,t} - \mu_{X_{o,t}|Y_t}) - \text{tr}([\Sigma_{Y_t, Y_t}^{-1}] K_{Y_t|X_{o,t}}) - (\mu_{Y_t|X_{o,t}} - \mu_{Y_t})^H \Sigma_{Y_t, Y_t}^{-1} (\mu_{Y_t|X_{o,t}} - \mu_{Y_t})) \quad (24)$$

Как было сказано ранее, первые два слагаемые, составляющие Q_t не зависят от θ , а значит задача о поиске $\arg \max_{\theta} E[\log P(X_o, Y|\theta^{(\tau)})|X_o, \theta^{(\tau)}]$ сводится к поиску

$$\arg \min_{\theta} \sum_{t=1}^G \left((x_{o,t} - \mu_{X_{o,t}|Y_t})^H \Sigma_{X_{o,t}|Y_t, X_{o,t}|Y_t}^{-1} (x_{o,t} - \mu_{X_{o,t}|Y_t}) + \text{tr}([\Sigma_{Y_t, Y_t}^{-1}] K_{Y_t|X_{o,t}}) + (\mu_{Y_t|X_{o,t}} - \mu_{Y_t})^H \Sigma_{Y_t, Y_t}^{-1} (\mu_{Y_t|X_{o,t}} - \mu_{Y_t}) \right). \quad (25)$$

Учтем, что $\mu_{Y_t} = O_{(M+L_2) \times 1}$, в рамках задачи y_t – скрытая переменная, она оценивается так: $\hat{y}_t = \mu_{Y_t|X_{o,t}}$.

$$\begin{cases} \Sigma_{X_o|Y} = \Sigma_{X_o, X_o} - \Sigma_{X_o|Y} \Sigma_{Y, Y}^{-1} \Sigma_{Y|X_o} \\ \mu_{X_o|Y} = \mu_{X_o} + \Sigma_{X_o|Y} \Sigma_{Y, Y}^{-1} \cdot (y_o - \mu_y), \end{cases} \quad (26)$$