EM-алгоритм для оценки направления прибытия сигнала

Введем некоторые условные обозначения:

- θ — вектор направлений прибытия сигнала (DOA);
- τ — итерация EM-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;
- N — набор шумов (случайная величина), 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;
- s — набор сигналов (реализация), испускаемых источниками в моменты времени $t = \overline{1, G}$, s_t соответствует сигналу 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 (реализация), полученных датчиками в моменты времени $t = \overline{1, G}$, x_t соответствует сигналу в момент времени t;
- $\mathbf{O}_{D_1 \times D_2}$ — нулевая матрица размера $D_1 \times D_2$;
- Итоговый сигнал, получаемый массивом датчиков:

$$X_t = A(\theta) S_t + N_t, \tag{1}$$

где $S_t \sim CN(\mathbf{O_{M\times1}}, \mathbf{P}), t = \overline{1, G}$, $N_t \sim CN(\mathbf{O_{L\times1}}, \mathbf{\Lambda}), t = \overline{1, G}$, S_t имеет размер $M \times 1$, N_t имеет размер $L \times 1$, $\theta = [\theta_1, ..., \theta_M]^T$ – вектор направлений прибытия сигнала, $A(\theta)$ (далее – A) представляет собой матрицу управляющих векторов размера $L \times M$, $\mathbf{P}$ и $\mathbf{\Lambda}$ 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 EM-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 $P(S|X = x, \theta)$, воспользуемся формулой Байеса:

$$P(S|X = x, \theta) = \frac{P(X|S, \theta)P(S|\theta)}{P(X|\theta)} \tag{2}$$

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

$$X_t = AS_t + N_t$$
$$X_t \sim CN(\mathbf{O_{L\times 1}}, A\mathbf{P}A^H + \mathbf{\Lambda})$$

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

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

$$X_t|S_t = s_t, \theta \sim CN(As_t, \mathbf{\Lambda})$$

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

$$P(S|X = x, \theta) = \prod_{t=1}^{G} \frac{1}{\pi^M |\mathbf{\Lambda}||\mathbf{P}||(A\mathbf{P}A^H + \mathbf{\Lambda})^{-1}|} e^{-(x_t - A\cdot S_t)^H \mathbf{\Lambda}^{-1}(x_t - A\cdot S_t) - S_t^H \mathbf{P}^{-1} S_t + x_t^H |(A\mathbf{P}A^H + \mathbf{\Lambda})^{-1}|^{-1} x_t}, \tag{6}$$

Таким образом:

$$P(S|X = x, \theta) = \prod_{t=1}^{G} \frac{1}{\pi^M |\Sigma_{S_t|X_t}|} e^{-(S_t - \mu_{S_t|X_t})^H \Sigma_{S_t|X_t}^{-1} (S_t - \mu_{S_t|X_t})} \tag{7}$$

Условное распределение также будет комплексным гауссовским, а его параметры будут определяться в соответствии с модифицированной теоремой о нормальной корреляции (см. файл Conditional_Distribution.pdf). Определим кросс-ковариацию между S и X.

$$\begin{gathered}\mathsf{Cov}(S, X) = \mathbb{E}[SX^H] = \mathbb{E}[S(AS + N)^H] = \mathbb{E}[S(AS)^H] + \mathbb{E}[SN^H] = \mathbb{E}[S(AS)^H] = \\ \mathbb{E}[SS^H A^H] = \mathbb{E}[SS^H]A^H = \mathbf{P}A^H\end{gathered} \tag{8}$$

Данные переходы возможны ввиду свойств ковариации и независимости N и S. $\mathsf{Cov}(X, S) = (\mathsf{Cov}(S, X))^H$:

$$\mathsf{Cov}(X, S) = A\mathbf{P} \tag{9}$$

$$\mu = \begin{pmatrix} \mathbf{O}_{M\times 1} \\ \mathbf{O}_{L\times 1} \end{pmatrix}, \Sigma = \begin{pmatrix} \mathbf{P} & \mathbf{P}A^H \\ A\mathbf{P} & A\mathbf{P}A^H + \mathbf{\Lambda} \end{pmatrix}. \tag{10}$$

В соответствии с выкладками, полученными в файле Conditional_Distribution.pdf, параметры апостериорной плотности будут определяться по следующим формулам:

$$\begin{cases} \Sigma_{S_t|X_t} = \Sigma_{S_t} - \Sigma_{S_t,X_t}\Sigma_{X_t}^{-1}\Sigma_{X_t,S_t} \\ \mu_{S_t|X_t} = \mu_{S_t} + \Sigma_{S_t,X_t}\Sigma_{X_t}^{-1}(x_t - \mu_{X_t}) \end{cases} \tag{11}$$

Подствим в эту формулу полученные нами значения:

$$\begin{cases} \Sigma_{S_t|X_t} = \mathbf{P} - \mathbf{P}A^H(A\mathbf{P}A^H + \mathbf{\Lambda})^{-1}A\mathbf{P} \\ \mu_{S_t|X_t} = \mathbf{P}A^H(A\mathbf{P}A^H + \mathbf{\Lambda})^{-1}x_t \end{cases} \tag{12}$$

М-шаг

Требуется найти наилучшую оценку параметров, решив следующую задачу оптимизации:

$$\theta^{(\tau+1)} = \underset{\theta}{\operatorname{argmax}}\, \mathbb{E}[\log P(X,\, S|\theta^{(\tau)})|X, \theta^{(\tau)}] \tag{13}$$

Обозначим приведенное выше условное математическое ожидание через Q_t. Заметим, что $\log P(X,\, S|\theta^{(\tau)}) = \log P(X|S, \theta^{(\tau)})P(S|\theta^{(\tau)}) = \log P(S|X, \theta^{(\tau)})P(X|\theta^{(\tau)})$.
Можно заметить, что $P(X|S, S) = P(X, S)$. Работать с плотостью $P(X|S, S)$ удобнее: кросс-коварициия между S и $X|S$ будет представлять из себя нулевую матрицу. Найдем совместную плотность $P(X|S,\, S|\theta^{(\tau)})$:

$$P(X|S,\, S|\theta^{(\tau)}) = \prod_{t=1}^{G} \frac{1}{\pi^{M+L}|\Sigma|} e^{-(Z_t - \mu_{Z_t})^H \Sigma^{-1}(Z_t - \mu_{Z_t})}, \tag{14}$$

где:

$$Z_t = \begin{pmatrix} S_t \\ X_t|S_t \end{pmatrix}, \mu_{Z_t} = \begin{pmatrix} \mu_{S_t} \\ \mu_{X_t|S_t} \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{S_t} & O_{M\times L} \\ O_{L\times M} & \Sigma_{X_t|S_t} \end{pmatrix}$$

Найдем логарифм совместной плотности (т.е. полное правдоподобие) $\log P(X|S,\, S|\theta^{(\tau)})$:

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

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

$$P(X_t|S_t,\, S_t|\theta^{(\tau)}) = \frac{1}{\pi^{M+L}|\Sigma|} e^{-(Z_t - \mu_{Z_t})^H \Sigma^{-1}(Z_t - \mu_{Z_t})} \tag{15}$$

$$\begin{gathered} Q_t = \mathbb{E}[\log P(X_t|S_t,\, S_t|\theta^{(\tau)})|X_t, \theta^{(\tau)}] \\ = \mathbb{E}[-(M+L)\log(\pi) - \log|\Sigma| - (Z_t - \mu_{Z_t})^H \Sigma^{-1}(Z_t - \mu_{Z_t})|X_t = x_t, \theta^{(\tau)}] = \\ = -(M+L)\log(\pi) - \log|\Sigma| - \mathbb{E}[(Z_t - \mu_{Z_t})^H \Sigma^{-1}(Z_t - \mu_{Z_t})|X_t = x_t, \theta^{(\tau)}] = -(M+L)\log(\pi) - \log|\Sigma| \\ -\mathbb{E}[(X_t|S_t - \mu_{X_t|S_t})^H \Sigma_{X_t|S_t}^{-1}(X_t|S_t - \mu_{X|S})|X_t = x_t, \theta^{(\tau)}] - \mathbb{E}[(S_t - \mu_{S_t})^H \Sigma_{S_t}^{-1}(S_t - \mu_{S_t})|X_t = x_t, \theta^{(\tau)}] \\ = -(M+L)\log(\pi) - \log|\Sigma| - (x_t - \mu_{X|S})^H \Sigma_{X_t|S_t}^{-1}(x_t - \mu_{X_t|S_t}) \\ -\mathbb{E}[(S_t - \mu_{s_t})^H \Sigma_{S_t}^{-1}(S_t - \mu_{S_t})|X_t = x_t, \theta^{(\tau)}] \end{gathered}$$

$$\begin{gathered} Q_t = -(M+L)\log(\pi) - \log|\Sigma| - (x_t - \mu_{X_t|S_t})^H \Sigma_{X_t|S_t}^{-1}(x_t - \mu_{X_t|S_t}) \\ -\mathbb{E}[(S_t - \mu_{S_t})^H \Sigma_{S_t}^{-1}(S_t - \mu_{S_t})|X_t = x_t, \theta^{(\tau)}] \end{gathered} \tag{16}$$

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

$$S_t|(X_t = x_t, \theta^{(\tau)}) \sim CN(\mu_{S_t|X_t}, \Sigma_{S_t|X_t}) \Rightarrow S_t - \mu_{S_t}|(X_t = x_t, \theta^{(\tau)}) \sim CN(\mu_{S_t|X_t} - \mu_{S_t}, \Sigma_{S_t|X_t})$$

$$\Rightarrow [\Sigma^{-1}_{S_t,S_t}]^{\frac{1}{2}}(S_t - \mu_{S_t})|(X_t = x_t, \theta^{(\tau)}) \sim CN([\Sigma^{-1}_{S_t}]^{\frac{1}{2}}(\mu_{S_t|X_t} - \mu_{S_t}), [\Sigma^{-1}_{S_t}]^{\frac{1}{2}}\Sigma_{S_t|X_t}([\Sigma^{-1}_{S_t}]^{\frac{1}{2}})^H)$$

Учтем, что для комплексных векторов Y выполняется следующее соотношение:

$$\mathbb{E}[YY^H] = \mathbb{E}[Y]\mathbb{E}[Y^H] + \Sigma_{YY}. \tag{17}$$

$$\begin{aligned}
\mathbb{E}[(S_t - \mu_{S_t})^H\Sigma^{-1}_{S_t}(S_t - \mu_{S_t})|X_t = x_t, \theta^{(\tau)}] = \mathbb{E}[[[\Sigma^{-1}_{S_t}]^{\frac{1}{2}}(S_t - \mu_{S_t})]^H[\Sigma^{-1}_{S_t}]^{\frac{1}{2}}(S_t - \mu_{S_t})|X_t = x_t, \theta^{(\tau)}] = \\
= \operatorname{Tr}\Big(E[[\Sigma^{-1}_{S_t}]^{\frac{1}{2}}(S_t - \mu_{S_t})[[\Sigma^{-1}_{S_t}]^{\frac{1}{2}}(S_t - \mu_{S_t})]^H]|X_t = x_t, \theta^{(\tau)}\Big) = \\
= \operatorname{Tr}\Big([\Sigma^{-1}_{S_t}]^{\frac{1}{2}}\Sigma_{S_t|X_t}([\Sigma^{-1}_{S_t}]^{\frac{1}{2}})^H\Big) + \operatorname{Tr}\Big([[\Sigma^{-1}_{S_t}]^{\frac{1}{2}}(\mu_{S_t|X_t} - \mu_{S_t})][[\Sigma^{-1}_{S_t}]^{\frac{1}{2}}(\mu_{S_t|X_t} - \mu_{S_t})^H]\Big) = \\
= \operatorname{Tr}\big([\Sigma^{-1}_{S_t}]\Sigma_{S_t|X_t}\big) + (\mu_{S_t|X_t} - \mu_{S_t})^H\Sigma^{-1}_{S_t}(\mu_{S_t|X_t} - \mu_{S_t})
\end{aligned}$$

$$\begin{aligned}
Q_t = -(M+L)\log(\pi) - \log|\Sigma| - (x_t - \mu_{X_t|S_t})^H\Sigma^{-1}_{X_t|S_t,X_t|S_t}(x_t - \mu_{X_t|S_t}) \\
- \operatorname{Tr}\big([\Sigma^{-1}_{S_t}]\Sigma_{S_t|X_t}\big) - (\mu_{S_t|X_t} - \mu_{S_t})^H\Sigma^{-1}_{S_t}(\mu_{S_t|X_t} - \mu_{S_t})
\end{aligned} \tag{18}$$

$$\begin{aligned}
\mathbb{E}[\log P(X,\, S|\theta^{(\tau)})|X, \theta^{(\tau)}] = \sum_{t=1}^{G} Q_t = \sum_{t=1}^{G}\Big(-(M+L)\log(\pi) - \log|\Sigma| - \\
(x_t - \mu_{X_t|S_t})^H\Sigma^{-1}_{X_t|S_t}(x_t - \mu_{X_t|S_t}) - \operatorname{Tr}\big([\Sigma^{-1}_{S_t}]\Sigma_{S_t|X_t}\big) - (\mu_{S_t|X_t} - \mu_{S_t})^H\Sigma^{-1}_{S_t}(\mu_{S_t|X_t} - \mu_{S_t})\Big)
\end{aligned} \tag{19}$$

Как было сказано ранее, первые два слагаемые, составляющие Q_t не зависят от θ, а значит задача о поиске $\operatorname{argmax}_\theta \mathbb{E}[\log P(X,\, S|\theta^{(\tau)})|X, \theta^{(\tau)}]$ сводится к поиску

$$\begin{aligned}
\operatorname*{argmin}_\theta \sum_{t=1}^{G}\Big((x_t - \mu_{X_t|S_t})^H\Sigma^{-1}_{X_t|S_t}(x_t - \mu_{X_t|S_t}) + \operatorname{Tr}\big([\Sigma^{-1}_{S_t}]\Sigma_{S_t|X_t}\big) \\
+(\mu_{S_t|X_t} - \mu_{S_t})^H\Sigma^{-1}_{S_t}(\mu_{S_t|X_t} - \mu_{S_t})\Big).
\end{aligned} \tag{20}$$

Учтем, что $\mu_{S_t} = O_{M\times 1}$, $\mu_{X_t|S_t} = As_t$, в рамках задачи s_t – скрытая переменная, она оценивается так: $\hat{s}_t = \mu_{S_t|X_t}$.

$$\begin{aligned}
\operatorname*{argmin}_\theta \sum_{t=1}^{G}\Big((x_t - \mu_{X_t|S_t})^H\Sigma^{-1}_{X_t|S_t}(x_t - \mu_{X_t|S_t}) + \operatorname{Tr}\big([\Sigma^{-1}_{S_t}]\Sigma_{S_t|X_t}\big) \\
+(\mu_{S_t|X_t} - \mu_{S_t})^H\Sigma^{-1}_{S_t}(\mu_{S_t|X_t} - \mu_{S_t})\Big) = \\
\operatorname*{argmin}_\theta \sum_{t=1}^{G}\big((x_t - A\mu_{S_t|X_t})^H\mathbf{\Lambda}^{-1}(x_t - A\mu_{S_t|X_t}) + \operatorname{Tr}\big([\mathbf{P}^{-1}]\Sigma_{S_t|X_t}\big) \\
+(\mu_{S_t|X_t} - O_{M\times 1})^H\mathbf{P}^{-1}(\mu_{S_t|X_t} - O_{M\times 1})\big) = \\
\operatorname*{argmin}_\theta \sum_{t=1}^{G}\big(x_t^H\mathbf{\Lambda}^{-1}x_t - (A\mu_{S_t|X_t})^H\mathbf{\Lambda}^{-1}x_t - x_t^H\mathbf{\Lambda}^{-1}A\mu_{S_t|X_t} + (A\mu_{S_t|X_t})^H\mathbf{\Lambda}^{-1}A\mu_{S_t|X_t}\big) = \\
\operatorname*{argmin}_\theta \sum_{t=1}^{G}\big(-(A\mu_{S_t|X_t})^H\mathbf{\Lambda}^{-1}x_t - x_t^H\mathbf{\Lambda}^{-1}A\mu_{S_t|X_t} + (A\mu_{S_t|X_t})^H\mathbf{\Lambda}^{-1}A\mu_{S_t|X_t}\big) = \\
\operatorname*{argmin}_\theta \sum_{t=1}^{G}\Big(-\mu^H_{S_t|X_t}A^H\mathbf{\Lambda}^{-1}x_t - x_t^H\mathbf{\Lambda}^{-1}A\mu_{S_t|X_t} + \mu^H_{S_t|X_t}A^H\mathbf{\Lambda}^{-1}A\mu_{S_t|X_t}\Big)
\end{aligned} \tag{21}$$

Теперь стоит подумать о том, как вычислить минимум для этой функции. Для начала определим первую производную для минимизируемой функции. Обозначим выражение, для которого мы ищем argmin, через $\tilde{Q}$.

$$\frac{\partial \tilde{Q}}{\partial \theta_i} = \sum_{t=1}^{G} \frac{\partial}{\partial \theta_i} \left(-x_t^H \mathbf{\Lambda}^{-1} A \mu_{S_t|X_t} - \mu_{S_t|X_t}^H A^H \mathbf{\Lambda}^{-1} x_t + \mu_{S_t|X_t}^H A^H \mathbf{\Lambda}^{-1} A \mu_{S_t|X_t} \right) =$$
$$\sum_{t=1}^{G} \left(-x_t^H \mathbf{\Lambda}^{-1} \left(\frac{\partial A}{\partial \theta_i} \right) \mu_{S_t|X_t} - \mu_{S_t|X_t}^H \left(\frac{\partial A}{\partial \theta_i} \right)^H \mathbf{\Lambda}^{-1} x_t + \mu_{S_t|X_t}^H \left(\frac{\partial A}{\partial \theta_i} \right)^H \mathbf{\Lambda}^{-1} A \mu_{S_t|X_t} \right. \tag{22}$$
$$\left. + \mu_{S_t|X_t}^H A^H \mathbf{\Lambda}^{-1} \left(\frac{\partial A}{\partial \theta_i} \right) \mu_{S_t|X_t} \right)$$

Теперь о том, что из себя представляет производная для матрицы A:

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

Пусть $h = -2j\pi \frac{d}{\lambda}$,

$$\frac{\partial A(\theta)}{\partial \theta_i} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & h \cos(\theta_i) e^{h \sin(\theta_i)} & \dots & 0 \\ \dots & \ddots & \dots & \ddots & \vdots \\ 0 & \dots & h(L-1) \cos(\theta_i) e^{h(L-1) \sin(\theta_i)} & \dots & 0 \end{bmatrix}. \tag{23}$$

Проблема заключается в том, что оптимизируемая функция $\tilde{Q}(\theta)$ – вещественнозначная, $\theta \in \mathbb{R}^M$ а ее производная – комплекснозначная. Будем находить экстремум по следующей схеме:

$$\hat{\theta}^{(\tau+1,k)} = \hat{\theta}^{(\tau+1,k-1)} - \eta Re(\boldsymbol{\nabla} \tilde{Q}(\theta^{(\tau)})), \tag{24}$$

где $\hat{\theta}^{(\tau+1,k)}$ – оценка $\theta^{\tau+1}$ на k-м шаге градиентного спуска, $\hat{\theta}^{(\tau+1,0)} = \theta^{(\tau)}$.

Ранние неверные выкладки

$$-(x_t - AS_t)^H \mathbf{\Lambda}^{-1} (x_t - AS_t) - S_t^H \mathbf{P}^{-1} S_t + x_t^H (A\mathbf{P}A^H + \mathbf{\Lambda})^{-1} x_t = -(S_t - \mu_{S_t|X_t})^H \Sigma_{S_t|X_t}^{-1} (S_t - \mu_{S_t|X_t})$$

$$-x_t^H \mathbf{\Lambda}^{-1} x_t + x_t^H \mathbf{\Lambda}^{-1} AS_t + (AS_t)^H \mathbf{\Lambda}^{-1} x_t - (AS_t)^H \mathbf{\Lambda}^{-1} AS_t - S_t^H \mathbf{P}^{-1} S_t$$
$$+ x_t^H (A\mathbf{P}A^H + \mathbf{\Lambda})^{-1} x_t = -S_t^H \Sigma_{S_t|X_t}^{-1} S_t + \mu_{S_t|X_t}^H \Sigma_{S_t|X_t}^{-1} S_t + S_t^H \Sigma_{S_t|X_t}^{-1} \mu_{S_t|X_t} - \mu_{S_t|X_t}^H \Sigma_{S_t|X_t}^{-1} \mu_{S_t|X_t}$$

$$\begin{cases} -x_t^H \mathbf{\Lambda}^{-1} x_t + x_t^H (A\mathbf{P}A^H + \mathbf{\Lambda})^{-1} x_t = -\mu_{S_t|X_t}^H \Sigma_{S_t|X_t}^{-1} \mu_{S_t|X_t} \\ x_t^H \mathbf{\Lambda}^{-1} AS_t = \mu_{S_t|X_t}^H \Sigma_{S_t|X_t}^{-1} S_t \\ (AS_t)^H \mathbf{\Lambda}^{-1} x_t = S_t^H \Sigma_{S_t|X_t}^{-1} \mu_{S_t|X_t} \\ -S_t^H \mathbf{P}^{-1} S_t - (AS_t)^H \mathbf{\Lambda}^{-1} AS_t = -S_t^H \Sigma_{S_t|X_t}^{-1} S_t \end{cases}$$

$$\begin{cases} -x_t^H \mathbf{\Lambda}^{-1} x_t + x_t^H (A\mathbf{P}A^H + \mathbf{\Lambda})^{-1} x_t = -\mu_{S_t|X_t}^H \Sigma_{S_t|X_t}^{-1} \mu_{S_t|X_t} \\ x_t^H \mathbf{\Lambda}^{-1} AS_t = \mu_{S_t|X_t}^H \Sigma_{S_t|X_t}^{-1} S_t \\ S_t^H A^H \mathbf{\Lambda}^{-1} x_t = S_t^H \Sigma_{S_t|X_t}^{-1} \mu_{S_t|X_t} \\ -S_t^H \mathbf{P}^{-1} S_t - S_t^H A^H \mathbf{\Lambda}^{-1} AS_t = -S_t^H \Sigma_{S_t|X_t}^{-1} S_t \end{cases}$$

$$\begin{cases} -x_t^H(\mathbf{\Lambda}^{-1} - (A\mathbf{P}A^H + \mathbf{\Lambda})^{-1})x_t = -\mu_{S_t|X_t}^H \Sigma_{S_t|X_t}^{-1} \mu_{S_t|X_t} \\ x_t^H \mathbf{\Lambda}^{-1} A S_t = \mu_{S_t|X_t}^H \Sigma_{S_t|X_t}^{-1} S_t \\ S_t^H A^H \mathbf{\Lambda}^{-1} x_t = S_t^H \Sigma_{S_t|X_t}^{-1} \mu_{S_t|X_t} \\ -S_t^H(\mathbf{P}^{-1} + A^H \mathbf{\Lambda}^{-1} A) S_t = -S_t^H \Sigma_{S_t|X_t}^{-1} S_t \end{cases}$$

Предполагая обратимость матрицы $\mathbf{P}^{-1} + A^H \mathbf{\Lambda}^{-1} A$,

$$-S_t^H(\mathbf{P}^{-1}+A^H \mathbf{\Lambda}^{-1} A) S_t = -S_t^H \Sigma_{S_t|X_t}^{-1} S_t \Rightarrow \mathbf{P}^{-1}+A^H \mathbf{\Lambda}^{-1} A = \Sigma_{S_t|X_t}^{-1} \Rightarrow \Sigma_{S_t|X_t} = (\mathbf{P}^{-1}+A^H \mathbf{\Lambda}^{-1} A)^{-1}$$

$$S_t^H A^H \mathbf{\Lambda}^{-1} x_t = S_t^H \Sigma_{S_t|X_t}^{-1} \mu_{S_t|X_t}$$

Равенство выше должно выполняться для любых реализаций S_t, множитель S_t^H является первым множителем в обоих произведениях, соответственно, равенство останется верным после удаления этих множителей.

$$-A^H \mathbf{\Lambda}^{-1} x_t = (\mathbf{P}^{-1} + A^H \mathbf{\Lambda}^{-1} A)^{-1} \mu_{S_t|X_t} \Rightarrow \mu_{S_t|X_t} = -(\mathbf{P}^{-1} + A^H \mathbf{\Lambda}^{-1} A) A^H \mathbf{\Lambda}^{-1} x_t$$

$$\begin{cases} \Sigma_{S_t|X_t} = (\mathbf{P}^{-1} + A^H \mathbf{\Lambda}^{-1} A)^{-1} \\ \mu_{S_t|X_t} = -(\mathbf{P}^{-1} + A^H \mathbf{\Lambda}^{-1} A) A^H \mathbf{\Lambda}^{-1} x_t \end{cases} \tag{25}$$