

$$\begin{aligned}
& \mathbb{E}_{q(\Xi)}[\log \mathbf{P}(X, S)] = \\
& \mathbb{E}_{q(\Xi)}[\log \mathbf{P}(X | S) + \log \mathbf{P}(S)] = \\
& \left[\sum_{t=1}^T \mathbb{E}_{q(\Xi_t)} \log[\mathbf{P}(X_t | S_t)] + \sum_{t=1}^T \mathbb{E}_{q(\Xi_t)}[\log \mathbf{P}(S_t)] \right] = \\
& -T \left[\log |\mathbf{\Lambda}| + \text{Tr}(\mathbf{\Lambda}^{-1} \mathbb{E}_{q(\Xi)}[X X^*]) - 2 \text{Re} \text{Tr}(\mathbf{A}^*(\theta) \mathbf{\Lambda}^{-1} \mathbb{E}_{q(\Xi)}[X S^*]) \right. \\
& \quad \left. + \text{Tr}(\mathbf{\Lambda}^{-1} \mathbf{A}(\theta) \mathbb{E}_{q(\Xi)}[S S^*] \mathbf{A}^*(\theta)) + \log |\mathbf{\Gamma}| + \text{Tr}(\mathbf{\Gamma}^{-1} \mathbb{E}_{q(\Xi)}[S S^*]) \right].
\end{aligned}$$

М-шаг для неизвестного шума

Пусть $\mathbf{\Lambda} = h \mathbf{I}_L$, и требуется найти h , максимизирующий следующую функцию:

$$\begin{aligned}
& -T \left[\log |\mathbf{\Lambda}| + \text{Tr}(\mathbf{\Lambda}^{-1} \mathbb{E}_{q(\Xi)}[X X^*]) - 2 \text{Re} \text{Tr}(\mathbf{A}^*(\theta) \mathbf{\Lambda}^{-1} \mathbb{E}_{q(\Xi)}[X S^*]) \right. \\
& \quad \left. + \text{Tr}(\mathbf{\Lambda}^{-1} \mathbf{A}(\theta) \mathbb{E}_{q(\Xi)}[S S^*] \mathbf{A}^*(\theta)) + \log |\mathbf{\Gamma}| + \text{Tr}(\mathbf{\Gamma}^{-1} \mathbb{E}_{q(\Xi)}[S S^*]) \right].
\end{aligned}$$

Исключим слагаемые, не содержащие $\mathbf{\Lambda}$:

$$\begin{aligned}
& -T \left[\log |\mathbf{\Lambda}| + \text{Tr}(\mathbf{\Lambda}^{-1} \mathbb{E}_{q(\Xi)}[X X^*]) - 2 \text{Re} \text{Tr}(\mathbf{A}^*(\theta) \mathbf{\Lambda}^{-1} \mathbb{E}_{q(\Xi)}[X S^*]) \right. \\
& \quad \left. + \text{Tr}(\mathbf{\Lambda}^{-1} \mathbf{A}(\theta) \mathbb{E}_{q(\Xi)}[S S^*] \mathbf{A}^*(\theta)) \right].
\end{aligned}$$

При этом, $W^{-1} = \frac{1}{h} \mathbf{I}_L$. Проведем преобразования:

$$\begin{aligned}
& -T \left[\log |\mathbf{\Lambda}| + \text{Tr}(\mathbf{\Lambda}^{-1} \mathbb{E}_{q(\Xi)}[X X^*]) - 2 \text{Re} \text{Tr}(\mathbf{A}^*(\theta) \mathbf{\Lambda}^{-1} \mathbb{E}_{q(\Xi)}[X S^*]) \right. \\
& \quad \left. + \text{Tr}(\mathbf{\Lambda}^{-1} \mathbf{A}(\theta) \mathbb{E}_{q(\Xi)}[S S^*] \mathbf{A}^*(\theta)) \right] = \\
& -T \left[\log |h \mathbf{I}_L| + \text{Tr}\left(\frac{1}{h} \mathbf{I}_L \mathbb{E}_{q(\Xi)}[X X^*]\right) - 2 \text{Re} \text{Tr}\left(\frac{1}{h} \mathbf{I}_L \mathbf{A}^*(\theta) \mathbb{E}_{q(\Xi)}[X S^*]\right) \right. \\
& \quad \left. + \text{Tr}\left(\frac{1}{h} \mathbf{I}_L \mathbf{A}(\theta) \mathbb{E}_{q(\Xi)}[S S^*] \mathbf{A}^*(\theta)\right) \right] = \\
& -T \left[L \log h + \frac{1}{h} \left(\text{Tr}(\mathbb{E}_{q(\Xi)}[X X^*]) - 2 \text{Re} \text{Tr}(\mathbf{A}^*(\theta) \mathbb{E}_{q(\Xi)}[X S^*]) \right. \right. \\
& \quad \left. \left. + \text{Tr}(\mathbf{A}(\theta) \mathbb{E}_{q(\Xi)}[S S^*] \mathbf{A}^*(\theta)) \right) \right].
\end{aligned}$$

Обозначим через B выражение

$$\text{Tr}(\mathbb{E}_{q(\Xi)}[X X^*]) - 2 \text{Re} \text{Tr}(\mathbf{A}^*(\theta) \mathbb{E}_{q(\Xi)}[X S^*]) + \text{Tr}(\mathbf{A}(\theta) \mathbb{E}_{q(\Xi)}[S S^*] \mathbf{A}^*(\theta)).$$

Получаем функцию

$$\mathcal{S}(h) = -T \left[L \log h + \frac{1}{h} B \right].$$

Производная этой формулы имеет следующий вид:

$$\mathcal{S}'(h) = \frac{B}{h^2} - \frac{L}{h}.$$

Эта производная равна нулю, если:

$$h = \frac{B}{L}.$$

Соответственно:

$$h^* = \frac{1}{L} \left(\text{Tr}(\mathbb{E}_{q(\Xi)}[XX^*]) - 2 \text{Re Tr}(A^*(\theta)\mathbb{E}_{q(\Xi)}[XS^*]) + \text{Tr}(A(\theta)\mathbb{E}_{q(\Xi)}[SS^*]A^*(\theta)) \right).$$

Ввиду того, что нет гарантии положительности B , нужно брать максимум из двух чисел: вышеуказанной оценки и некоторого малого числа ε .

$$h^{(\tau)} = \max \left(\frac{1}{L} \left(\text{Tr}(\mathbb{E}_{q(\Xi)}[XX^*]) - 2 \text{Re Tr}(A^*(\theta)\mathbb{E}_{q(\Xi)}[XS^*]) + \text{Tr}(A(\theta)\mathbb{E}_{q(\Xi)}[SS^*]A^*(\theta)) \right), \varepsilon \right).$$

МАР-оценивание ковариации сигналов

Можно превратить ЕМ в МАР-ЕМ, если искать не $\log P(X | \theta)$, а $\log P(X | \theta) + \log P(\theta)$, в случае моей задачи разумно задать априор для ковариации сигналов (ЕМ-алгоритм с ММП-оценкой ее не ищет практически, неидентифицируемость или слабая идентифицируемость масштаба, вероятно). Будем считать, что ковариация сигналов $\mathbf{\Gamma}_{K \times K}$ имеет обратное комплексное распределение Вишарта:

$$\mathbf{\Gamma} \sim \mathcal{CW}^{-1}(\Psi, \nu), \quad (1)$$

где Ψ – априорная информация о ковариации сигналов, ν – степень уверенности.

В таком случае, оптимизации подвергается следующая функция:

$$\mathcal{Q}(\Upsilon | \Upsilon^{(\tau-1)}) = \mathbb{E}_{q(\Xi)}[\log P(X, S | \theta)] + \log P(\mathbf{\Gamma}). \quad (2)$$

Точный шаг для обновления параметров будет определяться так:

$$\mathbf{\Gamma} = \frac{\text{TE}[SS^* | q(\Xi)] + \Psi}{\nu + K + T}. \quad (3)$$

Чем больше ν , тем больше уверенности в априорных данных. Плотность для $\mathbf{\Gamma}$ пропорциональна следующей величине (нормировочная константа в данном случае не важна):

$$P(\mathbf{\Gamma}) \propto |\mathbf{\Gamma}|^{-(\nu+K)} \exp(-\text{Tr}(\Psi \mathbf{\Gamma}^{-1})). \quad (4)$$