

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

М-шаг для неизвестного шума

Пусть $\Lambda = hI_L$, и требуется найти h , максимизирующий следующую функцию:

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

Исключим слагаемые, не содержащие Λ :

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

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

$$\begin{aligned}
& -T \left[\log |\Lambda| + \text{Tr} (\Lambda^{-1} \mathbb{E}_{q(\Xi)} [XX^*]) - 2 \operatorname{Re} \text{Tr} (A^*(\theta) \Lambda^{-1} \mathbb{E}_{q(\Xi)} [XS^*]) \right. \\
& \left. + \text{Tr} (\Lambda^{-1} A(\theta) \mathbb{E}_{q(\Xi)} [SS^*] A^*(\theta)) \right] = \\
& -T \left[\log |hI_L| + \text{Tr} \left(\frac{1}{h} I_L \mathbb{E}_{q(\Xi)} [XX^*] \right) - 2 \operatorname{Re} \text{Tr} \left(\frac{1}{h} I_L A^*(\theta) \mathbb{E}_{q(\Xi)} [XS^*] \right) \right. \\
& \left. + \text{Tr} \left(\frac{1}{h} I_L A(\theta) \mathbb{E}_{q(\Xi)} [SS^*] A^*(\theta) \right) \right] = \\
& -T \left[L \log h + \frac{1}{h} \left(\text{Tr} (\mathbb{E}_{q(\Xi)} [XX^*]) - 2 \operatorname{Re} \text{Tr} (A^*(\theta) \mathbb{E}_{q(\Xi)} [XS^*]) \right. \right. \\
& \left. \left. + \text{Tr} (A(\theta) \mathbb{E}_{q(\Xi)} [SS^*] A^*(\theta)) \right) \right].
\end{aligned}$$

Обозначим через B выражение

$$\text{Tr} (\mathbb{E}_{q(\Xi)} [XX^*]) - 2 \operatorname{Re} \text{Tr} (A^*(\theta) \mathbb{E}_{q(\Xi)} [XS^*]) + \text{Tr} (A(\theta) \mathbb{E}_{q(\Xi)} [SS^*] A^*(\theta)).$$

Получаем функцию

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

Производная этой формулы имеет следующий вид:

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

Эта производная равна нулю, если:

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

Соответственно:

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

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

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

MAP-оценивание ковариации сигналов

Можно превратить EM в MAP-EM, 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 (EM-алгоритм с ММП-оценкой ее не ищет практически, неидентифицируемость или слабая идентифицируемость масштаба, вероятно). Будем считать, что ковариация сигналов $\Gamma_{K \times K}$ имеет обратное комплексное распределение Вишарта:

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

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

В таком случае, оптимизации подвергается следующая функция:

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

Точный шаг для обновления параметров будет определяться так:

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

Чем больше ν , тем больше уверенности в априорных данных. Плотность для Γ пропорциональна следующей величине (нормировочная константа в данном случае не важна):

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