Е-шаг

Требуется найти математическое ожидание полного правдоподобия с учетом текущей оценки параметров и апостериорного совместного распределения пропущенных значений в наблюдениях X_m и сигналов S

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

Преобразуем выражение, обозначив через $\mathcal{I}$ условную информацию $X_o = x_o, \Psi^{(\tau-1)}$:

$$\begin{gathered}\mathbb{E}_{(X_m,S)|\mathcal{I}}[\log P(X,S)] = \\ \mathbb{E}_{(X_m,S)|\mathcal{I}}[\log P(X|S) + \log P(S)] = \\ -\mathrm{T}\Big[\log|\mathbf{\Lambda}| + \mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\mathbb{E}_{(X_m,S)|\mathcal{I}}[XX^*]\right) - 2\,\mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\mathrm{A}(\theta)\mathbb{E}_{(X_m,S)|\mathcal{I}}[XS^*]\right) \\ + \mathrm{Tr}\big(\mathbf{\Lambda}^{-1}\mathrm{A}(\theta)\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^*]\mathrm{A}^*(\theta)\big) + \log|\mathbf{\Gamma}| + \mathrm{Tr}\big(\mathbf{\Gamma}^{-1}\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^*]\big)\Big].\end{gathered}$$

Для нахождения условных моментов, указанных в формуле выше, требуется найти апостериорное распределение скрытых переменных. Воспользуемся формулой произведения плотностей:

$$P(X_m, S|\mathcal{I}) = P(X_m|\mathcal{I}) \cdot P(S|X_m, \mathcal{I}). \tag{2}$$

Сначала найдем апостериорное распределение $P(X_m|\mathcal{I})$, причем ввиду того, что индексы, соответствующие пропущенным значениям в наблюдениях, могут отличаться в зависимости от t, будем находить апостериорное распределение для каждого X_{t,m_t}. Создадим разбиение оценки ковариационной матрицы наблюдений $\hat{\mathrm{R}}$ на блоки, индуцированное этим разбиением множества индексов, оно имеет следующий вид:

$$\hat{\mathrm{R}} = \begin{pmatrix} \hat{\mathrm{R}}_{o_t,o_t} & \hat{\mathrm{R}}_{o_t,m_t} \\ \hat{\mathrm{R}}_{m_t,o_t} & \hat{\mathrm{R}}_{m_t,m_t} \end{pmatrix}, \tag{3}$$

где каждый блок определяется как

$$\hat{\mathrm{R}}_{a,b} = (\hat{\mathrm{R}}_{ij})_{i\in a, j\in b}.$$

Для каждого наблюдения, содержащего пропуски, требуется найти апостериорное распределение пропущенных значений, $P(X_{t,m_t}|X_{t,o_t} = x_{t,o_t}, \Psi^{(\tau-1)}), t = 1,\dots,\mathrm{T}, |m_t| > 0$. Обозначим через $\mathcal{I}_t$ условную информацию $X_{t,o_t} = x_{t,o_t}, \Psi^{(\tau-1)}$. Параметры апостериорного распределения $P(X_{t,m_t}|\mathcal{I}_t), t = 1,\dots,\mathrm{T}$ на итерации τ можно найти следующим образом:

$$\begin{cases} \mu_{X_{t,m_t}|\mathcal{I}_t} = \hat{\mathrm{R}}_{m_t,o_t}\left(\hat{\mathrm{R}}_{o_t,o_t}\right)^{-1} \cdot x_{t,o_t}, \\ \Sigma_{X_{t,m_t}|\mathcal{I}_t} = \hat{\mathrm{R}}_{m_t,m_t} - \hat{\mathrm{R}}_{m_t,o_t}\left(\hat{\mathrm{R}}_{o_t,o_t}\right)^{-1}\hat{\mathrm{R}}_{o_t,m_t}, \end{cases} \tag{4}$$

где $\hat{\mathrm{R}}_{o_t,o_t} = \hat{\mathrm{R}}^{(\tau-1)}_{o_t,o_t}$, $\hat{\mathrm{R}}_{o_t,m_t} = \hat{\mathrm{R}}^{(\tau-1)}_{o_t,m_t}$, $\hat{\mathrm{R}}_{m_t,o_t} = \hat{\mathrm{R}}^{(\tau-1)}_{m_t,o_t}$, $\hat{\mathrm{R}}_{m_t,m_t} = \hat{\mathrm{R}}^{(\tau-1)}_{m_t,m_t}$.
Для каждого наблюдения X_t оценим условную ковариационную матрицу $\widetilde{\Sigma}_{X_t} = \mathbb{E}_{X_{m_t}|\mathcal{I}_t}[X_tX_t^*]$:

$$\begin{gathered}\widetilde{\Sigma}_{X_t} = \begin{pmatrix} \mu_{X_{t,o_t}|\mathcal{I}_t} \cdot \mu^*_{X_{t,o_t}|\mathcal{I}_t} & \mu_{X_{t,o_t}|\mathcal{I}_t} \cdot \mu^*_{X_{t,m_t}|\mathcal{I}_t} \\ \mu_{X_{t,m_t}|\mathcal{I}_t} \cdot \mu^*_{X_{t,o_t}|\mathcal{I}_t} & \mu_{X_{t,m_t}|\mathcal{I}_t} \cdot \mu^*_{X_{t,m_t}|\mathcal{I}_t} + \Sigma_{X_{t,m_t}|\mathcal{I}_t} \end{pmatrix} = \\ \mathbb{E}[X_t|\mathcal{I}_t] \cdot \mathbb{E}[X_t|\mathcal{I}_t]^* + \begin{pmatrix} \mathbf{O}_{o_t,o_t} & \mathbf{O}_{o_t,m_t} \\ \mathbf{O}_{m_t,o_t} & \Sigma_{X_{t,m_t}|\mathcal{I}_t} \end{pmatrix},\end{gathered}$$

где $\mathbf{O}_{o_t,o_t}$, $\mathbf{O}_{o_t,m_t}$, $\mathbf{O}_{m_t,o_t}$ – нулевые блочные матрицы. Разбиение указанной матрицы на четыре блока, три из которых состоят из нулей, индукцировано разбиением множества индексов

$\{1,\dots,\mathrm{L}\}$ на множества o_t, m_t.
Оценим условную ковариационную матрицу наблюдений $\widetilde{\Sigma}_X = \mathbb{E}_{(X_m,S)|\mathcal{I}}[XX^*]$:

$$\widetilde{\Sigma}_X = \frac{1}{\mathrm{T}}\sum_{t=1}^{T}\widetilde{\Sigma}_{X_t}. \tag{5}$$

Параметры апостериорного распределения $P(S_t|\mathcal{I}_t, X_{t,m_t}), t = 1,\dots,\mathrm{L}$ можно найти исходя из следующих формул:

$$\begin{cases} \mu_{S_t|\mathcal{I}_t,X_{t,m_t}} = \mathbf{\Gamma}\mathrm{A}^*R^{-1}\widetilde{\Sigma}_{X_t}R^{-1}\mathbb{E}[X_t|\mathcal{I}_t], \\ \Sigma_{S_t|\mathcal{I}_t,X_{t,m_t}} = \mathbf{\Gamma} - \mathbf{\Gamma}\mathrm{A}^*R^{-1}\widetilde{\Sigma}_{X_t}R^{-1}\mathrm{A}\mathbf{\Gamma}, \end{cases} \tag{6}$$

где $\mathrm{A} = \mathrm{A}(\theta^{(\tau-1)}), \mathbf{\Gamma} = \mathbf{\Gamma}^{(\tau-1)}$.
Оценим ковариационную матрицу сигналов с учетом текущей оценки параметров и доступных наблюдений $\widetilde{\Sigma}_S = \mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^*]$.

$$\widetilde{\Sigma}_S = \frac{1}{\mathrm{T}}\Big[\sum_{t=1}^{T}\big(\Sigma_{S_t|\mathcal{I}_t,X_{t,m_t}} + \mu_{S_t|\mathcal{I}_t,X_{t,m_t}}\cdot\mu^H_{S_t|\mathcal{I}_t,X_{t,m_t}}\big)\Big]. \tag{7}$$

Остается оценить кросс-ковариацию $\Sigma_{XS} = \mathbb{E}_{(X_m,S)|\mathcal{I}}[XS^*]$:

$$\begin{aligned} \mathbb{E}_{(X_m,S)|\mathcal{I}}[XS^*] &= \mathbb{E}_{(X_m,S)|\mathcal{I}}[XX^*](\hat{\mathrm{R}}^{(\tau-1)})^{-1}\mathrm{A}(\theta^{(\tau-1)})\mathbf{\Gamma}^{(\tau-1)} \\ &= \widetilde{\Sigma}_X(\hat{\mathrm{R}}^{(\tau-1)})^{-1}\mathrm{A}(\theta^{(\tau-1)})\mathbf{\Gamma}^{(\tau-1)}. \end{aligned} \tag{8}$$

Тест

Параметры апостериорного распределения $P(S_t|X_{t,o_t} = x_{t,o_t}, X_{t,m_t}, k^{(j-1)}, W^{(j-1)}), t = 1,\dots,\mathrm{T}$ можно найти исходя из следующих формул:

$$\begin{cases} \mu_{S_t|X_{t,o_t}=x_{t,o_t},X_{t,m_t},k^{(j-1)},W^{(j-1)}} = W\mathrm{A}^HR^{-1}\widetilde{\Sigma}_{X_t}R^{-1}\mathbb{E}[X_t|X_{t,o_t} = x_{t,o_t}, k^{(j-1)}, W^{(j-1)}], \\ \Sigma_{S_t|X_{t,o_t}=x_{t,o_t},X_{t,m_t},k^{(j-1)},W^{(j-1)}} = W - W\mathrm{A}^HR^{-1}\widetilde{\Sigma}_{X_t}R^{-1}\mathrm{A}W, \end{cases} \tag{9}$$

$$\begin{cases} \mu_{S_t|X_{t,o_t}=x_{t,o_t},X_{t,m_t},k^{(j-1)},W^{(j-1)}} = W\mathrm{A}^H\widetilde{\Sigma}^{-1}_{X_t}\mathbb{E}[X_t|X_{t,o_t} = x_{t,o_t}, k^{(j-1)}, W^{(j-1)}], \\ \Sigma_{S_t|X_{t,o_t}=x_{t,o_t},X_{t,m_t},k^{(j-1)},W^{(j-1)}} = W - W\mathrm{A}^H\widetilde{\Sigma}^{-1}_{X_t}\mathrm{A}W, \end{cases} \tag{10}$$

$$\begin{cases} \mu_{S_t|X_{t,o_t}=x_{t,o_t},X_{t,m_t},k^{(j-1)},W^{(j-1)}} = W\mathrm{A}^HR^{-1}\mathbb{E}[X_t|X_{t,o_t} = x_{t,o_t}, k^{(j-1)}, W^{(j-1)}], \\ \Sigma_{S_t|X_{t,o_t}=x_{t,o_t},X_{t,m_t},k^{(j-1)},W^{(j-1)}} = W - W\mathrm{A}^HR^{-1}\mathrm{A}W, \end{cases} \tag{11}$$

где

$$\widetilde{\Sigma}_{X_t} = \begin{pmatrix} \mu_{X_{t,o_t}|\mathcal{I}_t}\cdot\mu^*_{X_{t,o_t}|\mathcal{I}_t} & \mu_{X_{t,o_t}|\mathcal{I}_t}\cdot\mu^*_{X_{t,m_t}|\mathcal{I}_t} \\ \mu_{X_{t,m_t}|\mathcal{I}_t}\cdot\mu^*_{X_{t,o_t}|\mathcal{I}_t} & \mu_{X_{t,m_t}|\mathcal{I}_t}\cdot\mu^*_{X_{t,m_t}|\mathcal{I}_t} + \Sigma_{X_{t,m_t}|\mathcal{I}_t} \end{pmatrix} =$$
$$\mathbb{E}[X_t|\mathcal{I}_t]\cdot\mathbb{E}[X_t|\mathcal{I}_t]^* + \begin{pmatrix} \mathbf{O}_{o_t,o_t} & \mathbf{O}_{o_t,m_t} \\ \mathbf{O}_{m_t,o_t} & \Sigma_{X_{t,m_t}|\mathcal{I}_t} \end{pmatrix},$$

где $\mathcal{I}_t = X_{t,o_t} = x_{t,o_t}, k^{(j-1)}, W^{(j-1)}$.

Пусть $\mathcal{I}$ — комплекс условий $X_o = x_o, k^{(j-1)}, W^{(j-1)}$. Пусть $\mathcal{I}_t$ — комплекс условий $X_{t,o_t} = x_{t,o_t}, k^{(j-1)}, W^{(j-1)}$.

$$\begin{gathered}\mathbb{E}_{(X_m,S)|\mathcal{I}}[\log P(X,S)] = \\ \mathbb{E}_{(X_m,S)|\mathcal{I}}[\log(P(X|S)P(S))] = \\ \mathbb{E}_{(X_m,S)|\mathcal{I}}[\log P(X|S) + \log P(S)] = \\ -\Big[\mathrm{T}\log|\mathbf{\Lambda}| + \mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\mathbb{E}_{(X_m,S)|\mathcal{I}}[XX^H]\right) - 2\,\mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\mathrm{A}(\theta)\mathbb{E}_{(X_m,S)|\mathcal{I}}[XS^H]\right) \\ + \mathrm{Tr}\big(\mathbf{\Lambda}^{-1}\mathrm{A}(\theta)\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^H]\mathrm{A}^*(\theta)\big) + \mathrm{T}\log|W| + \mathrm{Tr}\big(W^{-1}\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^H]\big)\Big].\end{gathered}$$

Могу ли я рассчитать $\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^H], \mathbb{E}_{(X_m,S)|\mathcal{I}}[XS^H], \mathbb{E}_{(X_m,S)|\mathcal{I}}[XX^H]$ как усредненные величины по t. Ну то есть: вот есть величины $\mathbb{E}_{(X_m,S)|\mathcal{I}_t}[S_t S_t^H], t = 1, \ldots, T$. Я могу посчитать так:

$$\mathbb{E}_{(X_m,S)|\mathcal{I}}[SS^H] = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{(X_m,S)|\mathcal{I}_t}[S_t S_t^H] \tag{12}$$