VI-CVA

VI-CVA here denotes the CVA model trained by VI, and the whole training process is mainly established according to the reference:
Yu J, Ye L, Zhou L, et al. Dynamic process monitoring based on variational Bayesian canonical variate analysis[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 52(4): 2412-2422.

Wishart distribution $\mathcal{W}$ is used as the prior of the precision in Multivariate Gaussian.

Model Structure

input :

$\mathbf{x}_t \in \mathbb{R}^{M\times 1}$

$\mathbf{P}^t = [\mathbf{x}_{t-1}^T, \mathbf{x}_{t-2}^T, ..., \mathbf{x}_{t-l}^T]^T \in \mathbb{R}^{Ml\times 1}$

$\mathbf{F}^t = [\mathbf{x}_{t}^T, \mathbf{x}_{t+1}^T, ..., \mathbf{x}_{t+s}^T]^T \in \mathbb{R}^{M(s+1)\times 1}$

$\mathbf{P}^t = [p_1^t, p_2^t, ..., p_{Ml}^t]^T \in \mathbb{R}^{Ml\times 1}$

$\mathbf{F}^t = [f_1^t, f_2^t, ..., f_{M(s+1)}^t]^T \in \mathbb{R}^{M(s+1)\times 1}$

model mapping :

$\mathbf{P} = \mathbf{W}^T\mathbf{z} + \epsilon$

$\mathbf{F} = \mathbf{H}^T\mathbf{z} + \delta$

$\mathbf{z}_t \in \mathbb{R}^{D\times 1}$

$\mathbf{W}^T \in \mathbb{R}^{Ml\times D}$

$\mathbf{H}^T \in \mathbb{R}^{M(s+1)\times D}$

$$\mathbf{W}^T = \begin{pmatrix} -\mathbf{W}_1^T- \\ -\mathbf{W}_2^T- \\ \vdots \\ -\mathbf{W}_{Ml}^T- \end{pmatrix} = \left(\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_{Ml}\right), \mathbf{W}_m \in \mathbb{R}^{D\times 1}$$

$$\mathbf{H}^T = \begin{pmatrix} -\mathbf{H}_1^T- \\ -\mathbf{H}_2^T- \\ \vdots \\ -\mathbf{H}_{M(s+1)}^T- \end{pmatrix} = \left(\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_{M(s+1)}\right), \mathbf{H}_m \in \mathbb{R}^{D\times 1}$$

prior distributions：

$$\mathcal{P}(\mathbf{z}) = \prod_{t=1}^{N} \mathcal{N}(\mathbf{z}_t|0, \mathbf{I})$$

$$\mathcal{P}(\epsilon) = \prod_{m=1}^{Ml} \mathcal{N}(\epsilon_m|0, \tau_m^{-1}); \mathcal{P}(\delta) = \prod_{m=1}^{M(s+1)} \mathcal{N}(\delta_m|0, \psi_m^{-1})$$

$$\mathcal{P}(\tau) = \prod_{m=1}^{Ml} \mathcal{G}(\tau_m|j_m^\tau, k_m^\tau); \mathcal{P}(\psi) = \prod_{m=1}^{M(s+1)} \mathcal{G}(\psi_m|j_m^\psi, k_m^\psi)$$

$$\mathcal{P}(\alpha) = \prod_{m=1}^{Ml} \mathcal{W}(\alpha_m|a_m^\alpha, \mathbf{B}_m^\alpha); \mathcal{P}(\beta) = \prod_{m=1}^{M(s+1)} \mathcal{W}(\beta_m|a_m^\beta, \mathbf{B}_m^\beta)$$

$$\mathcal{P}(\mathbf{W}|\alpha) = \prod_{m=1}^{Ml} \mathcal{N}(\mathbf{W}_m|0, \alpha_m^{-1}); \mathcal{P}(\mathbf{H}|\beta) = \prod_{m=1}^{M(s+1)} \mathcal{N}(\mathbf{H}_m|0, \beta_m^{-1})$$

$$\mathcal{P}(\mathbf{P}|\mathbf{z}, \mathbf{W}, \alpha, \tau) = \prod_{t=1}^{N} \prod_{m=1}^{Ml} \mathcal{N}(p_m^t|\mathbf{W}_m^T\mathbf{z}_t, \tau_m^{-1})$$

$$\mathcal{P}(\mathbf{F}|\mathbf{z}, \mathbf{H}, \beta, \psi) = \prod_{t=1}^{N} \prod_{m=1}^{M(s+1)} \mathcal{N}(f_m^t|\mathbf{H}_m^T\mathbf{z}_t, \psi_m^{-1})$$

variational distributions：

$$\mathcal{Q}(\mathbf{z}) = \prod_{t=1}^{N} \mathcal{N}(\mathbf{z}_t|\mu_t^z, (\Lambda_t^z)^{-1})$$

$$\mathcal{Q}(\mathbf{W}) = \prod_{m=1}^{Ml} \mathcal{N}(\mathbf{W}_m|\mu_m^W, (\Lambda_m^W)^{-1}); \mathcal{Q}(\mathbf{H}) = \prod_{m=1}^{M(s+1)} \mathcal{N}(\mathbf{H}_m|\mu_m^H, (\Lambda_m^H)^{-1})$$

$$\mathcal{Q}(\tau) = \prod_{m=1}^{Ml} \mathcal{G}(\tau_m|\lambda_m^\tau, \nu_m^\tau);\ \mathcal{Q}(\psi) = \prod_{m=1}^{M(s+1)} \mathcal{G}(\psi_m|\lambda_m^\psi, \nu_m^\psi)$$

$$\mathcal{Q}(\alpha) = \prod_{m=1}^{Ml} \mathcal{W}(\alpha_m|\nu_m^\alpha, \mathbf{V}_m^\alpha);\ \mathcal{Q}(\beta) = \prod_{m=1}^{M(s+1)} \mathcal{W}(\beta_m|\nu_m^\beta, \mathbf{V}_m^\beta)$$

Upate Strategy

joint distribution

$$\begin{aligned}\mathcal{P}(\mathbf{P}, \mathbf{F}, \mathbf{H}, \mathbf{W}, \mathbf{z}, \tau, \psi, \alpha, \beta) \\ = \mathcal{P}(\mathbf{P}|\mathbf{W}, \mathbf{z}, \tau)\mathcal{P}(\mathbf{W}|\alpha)\mathcal{P}(\alpha)\mathcal{P}(\tau) \\ \mathcal{P}(\mathbf{F}|\mathbf{H}, \mathbf{z}, \psi)\mathcal{P}(\mathbf{H}|\beta)\mathcal{P}(\beta)\mathcal{P}(\psi)\mathcal{P}(\mathbf{z})\end{aligned}$$

mean-field distribution

$$\begin{aligned}\mathcal{Q}(\mathbf{H}, \mathbf{W}, \mathbf{z}, \tau, \psi, \alpha, \beta) \\ = \mathcal{Q}(\mathbf{W})\mathcal{Q}(\alpha)\mathcal{Q}(\tau)\mathcal{Q}(\mathbf{H})\mathcal{Q}(\beta)\mathcal{Q}(\psi)\mathcal{Q}(\mathbf{z})\end{aligned}$$

Update $\mathcal{Q}(\mathbf{z}_t)$

$\Lambda_t^z = \sum_{m=1}^{Ml} \mathbb{E}_{Q(\tau_m)}(\tau_m)\mathbb{E}_{Q(\mathbf{W}_m)}(\mathbf{W}_m\mathbf{W}_m^T) + \sum_{m=1}^{M(s+1)} \mathbb{E}_{Q(\psi_m)}(\psi_m)\mathbb{E}_{Q(\mathbf{H}_m)}(\mathbf{H}_m\mathbf{H}_m^T) + \mathbf{I}$

$\mu_t^z = (\Lambda_t^z)^{-1}\{\sum_{m=1}^{Ml} \mathbb{E}_{Q(\tau_m)}(\tau_m)\mathbb{E}_{Q(\mathbf{W}_m)}(\mathbf{W}_m)p_m^t + \sum_{m=1}^{M(s+1)} \mathbb{E}_{Q(\psi_m)}(\psi_m)\mathbb{E}_{Q(\mathbf{H}_m)}(\mathbf{H}_m)f_m^t\}$

$\mathbb{E}_{Q(\tau_m)}(\tau_m) = \frac{\lambda_m^\tau}{\nu_m^\tau}$

$\mathbb{E}_{Q(\mathbf{W}_m)}(\mathbf{W}_m\mathbf{W}_m^T) = (\Lambda_m^W)^{-1} + \mu_m\mu_m^T$

update $\mathcal{Q}(\mathbf{H})$, $\mathcal{Q}(\mathbf{W})$

$\Lambda_m^W = \sum_{t=1}^{N} \mathbb{E}_{Q(\tau_m)}(\tau_m)\mathbb{E}_{Q(\mathbf{z}_t)}(\mathbf{z}_t\mathbf{z}_t^T) + \mathbb{E}_{Q(\alpha_m)}(\alpha_m)$

$\mu_m^{\mathbf{W}} = (\Lambda_m^W)^{-1}\mathbb{E}_{Q(\tau_m)}(\tau_m)\sum_{t=1}^N \mathbb{E}_{Q(\mathbf{z}_t)}(\mathbf{z}_t)p_m^t$

$\Lambda_m^H = \sum_{t=1}^N \mathbb{E}_{Q(\psi_m)}(\psi_m)\mathbb{E}_{Q(\mathbf{z}_t)}(\mathbf{z}_t\mathbf{z}_t^T) + \mathbb{E}_{Q(\beta_m)}(\beta_m)$

$\mu_m^{\mathbf{H}} = (\Lambda_m^H)^{-1}\mathbb{E}_{Q(\psi_m)}(\psi_m)\sum_{t=1}^N \mathbb{E}_{Q(\mathbf{z}_t)}(\mathbf{z}_t)f_m^t$

$\mathbb{E}_{Q(\alpha_\mathbf{m})}(\alpha_m) = \nu_m^\alpha \mathbf{V}_m^\alpha$

$\mathbb{E}_{Q(\mathbf{z}_t)}(\mathbf{z}_t\mathbf{z}_t^T) = (\Lambda_t^z)^{-1} + \mu_t^z(\mu_t^z)^T$

update $\mathcal{Q}(\tau)$, $\mathcal{Q}(\psi)$

$\lambda_m^\tau = j_m^\tau + \frac{1}{2}N$
$\nu_m^\tau = k_m^\tau + \frac{1}{2}\sum\limits_{t=1}^N \mathbb{E}_{Q(\mathbf{z}_t)Q(\mathbf{W}_m)}\{(p_m^t - \mathbf{W}_m^T\mathbf{z}_t)^T(p_m^t - \mathbf{W}_m^T\mathbf{z}_t)\}$

$\lambda_m^\psi = j_m^\psi + \frac{1}{2}N$
$\nu_m^\psi = k_m^\psi + \frac{1}{2}\sum\limits_{t=1}^N \mathbb{E}_{Q(\mathbf{z}_t)Q(\mathbf{H}_m)}\{(f_m^t - \mathbf{H}_m^T\mathbf{z}_t)^T(f_m^t - \mathbf{H}_m^T\mathbf{z}_t)\}$

$\mathbb{E}_{\mathcal{Q}(\mathbf{z})}(\mathbf{z}^T\mathbf{z}) = \mathrm{Tr}(\Lambda_z^{-1}) + \mu_z^T\mu_z$

update $\mathcal{Q}(\alpha)$, $\mathcal{Q}(\beta)$

$\nu_m^\alpha = a_m^\alpha + 1$
$(\mathbf{V}_m^\alpha)^{-1} = (\mathbf{B}_m^\alpha)^{-1} + \mathbb{E}_{Q(\mathbf{W}_m)}\{\mathbf{W}_m\mathbf{W}_m^T\}$

$\nu_m^\beta = a_m^\beta + 1$
$(\mathbf{V}_m^\beta)^{-1} = (\mathbf{B}_m^\beta)^{-1} + \mathbb{E}_{Q(\mathbf{H}_m)}\{\mathbf{H}_m\mathbf{H}_m^T\}$

Monitoring Strategy

Monitoring Strategy is completely established according to the aforementioned reference, where SPE and T^2 are constructed. The specific calculations are given as follows:

index T^2

$T^2 = \mathbf{z}_{new}^T \mathbf{z}_{new}$

where

$\mathbf{z}_{new} = \Lambda_t^{\mathbf{z}}(\mathbf{z}_{new}^p + \mathbf{z}_{new}^f)$

$\mathbf{z}_{new}^p = (\mu^W)^T \tau \mathbf{p}_{new}$

$\mathbf{z}_{new}^f = (\mu^H)^T \psi \mathbf{f}_{new}$

$\tau = diag\{[\mathbb{E}_{Q(\tau_m)}(\tau_m)]_{m=1}^{Ml}\}$

$\psi = diag\{[\mathbb{E}_{Q(\psi_m)}(\psi_m)]_{m=1}^{M(s+1)}\}$

index SPE

$SPE = \delta_{new}^T \psi \delta_{new}$

$\delta_{new} = \mathbf{f}_{new} - \mu^H \mathbf{z}_{new}$