Bayesian Temporal Matrix Factorization

Model description[1]:

- Assumption over observations:

$$y_{it} \sim \mathcal{N}\left(\boldsymbol{w}_i^\top \boldsymbol{x}_t, \tau_i^{-1}\right), \quad (i,t) \in \Omega \tag{1}$$

- Prior setting of factor matrices and precision:

$$\boldsymbol{w}_i \sim \mathcal{N}\left(\boldsymbol{\mu}_w, \Lambda_w^{-1}\right), \tag{2}$$

$$\boldsymbol{x}_t \sim \begin{cases} \mathcal{N}\left(\boldsymbol{0}, I_R\right), & \text{if } t \in \{1, 2, \ldots, h_d\}, \\ \mathcal{N}\left(A^\top \boldsymbol{v}_t, \Sigma\right), & \text{otherwise}, \end{cases} \tag{3}$$

$$\tau_i \sim \text{Gamma}\left(\alpha, \beta\right). \tag{4}$$

- Prior setting of hyperparameters:

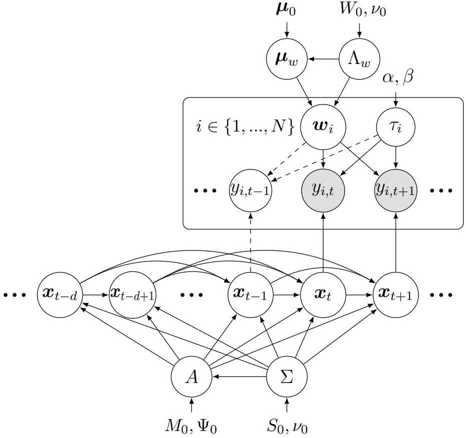
$$\boldsymbol{\mu}_w | \Lambda_w \sim \mathcal{N}\left(\boldsymbol{\mu}_0, (\beta_0 \Lambda_w)^{-1}\right), \, \Lambda_w \sim \mathcal{W}\left(W_0, \nu_0\right), \tag{5}$$

$$A \sim \mathcal{MN}_{(Rd)\times R}\left(M_0, \Psi_0, \Sigma\right), \, \Sigma \sim \mathcal{IW}\left(S_0, \nu_0\right), \tag{6}$$

[1] $\mathcal{N}(\cdot)$: Gaussian/Normal distribution; $\mathcal{W}(\cdot)$: Wishart distribution; $\mathcal{MN}(\cdot)$: Matrix normal distribution; $\mathcal{IW}(\cdot)$: Inverse Wishart distribution; Gamma$(\cdot)$: Gamma distribution.

Bayesian Temporal Matrix Factorization

Bayesian network:



- Graphical model of BTMF (time lag set: $\{1, 2, \ldots, d\}$). The shaded node ($y_{i,t}$) are the observed data in Ω.