Model Settings in Experiments

"Matrix and Tensor Model for Spatiotemporal Traffic Data Imputation and Forecasting"

How to Select Hyperparameters in the Imputation Methods?

- LATC can be described as the following optimization problem:

$$\begin{aligned} &\min_{\boldsymbol{\mathcal{X}},\boldsymbol{Z}} \ \|\boldsymbol{\mathcal{X}}\|_{r,*}+\frac{\gamma}{2}\|\boldsymbol{Z}\|_{\boldsymbol{A},\mathcal{H}} \\ &\text{s.t.} \begin{cases} \boldsymbol{\mathcal{X}}=\mathcal{Q}(\boldsymbol{Z}) \\ \mathcal{P}_{\Omega}(\boldsymbol{Z})=\mathcal{P}_{\Omega}(\boldsymbol{Y}) \end{cases} \end{aligned} \tag{1}$$

  where we have some hyperparameters in the solution algorithms:

  - Hyperparameter λ: This value determines the algorithm's convergence, which is usually set as small values, e.g., $\lambda \in \{10^{-4}, 10^{-5}\}$.
  - Hyperparameter γ: This is the weight parameter of temporal autoregression. This value can be determined according to the importance of temporal modeling. If we want to reinforce temporal correlations of time series, we can use a relatively large value. In our experiments (on both Seattle and Portland traffic datasets), we set $\gamma \in \{\frac{\lambda}{10}, \frac{\lambda}{5}, \lambda, 5\lambda, 10\lambda\}$.
  - Truncation parameter r: This value is important for the singular value thresholding, making the first r singular values fixed in the thresholding process. Generally speaking, this value can be set relatively small, but it is determined according to the sparsity of data. In our experiments (on both Seattle and Portland traffic datasets),

Comparison with Baseline Models

Section 4.1 Univariate Traffic Time Series Imputation

Section 4.2 Spatiotemporal Traffic Data Imputation

Section 4.3 Large-Scale Traffic Data Imputation

Section 4.4 Extreme Missing Traffic Data Imputation

Section 4.5 Traffic Forecasting From Sparse Data