Bayesian Tensor Ring Decomposition for Low Rank Tensor Completion

Zerui Tao [1] Qibin Zhao [2]

[1]Graduate School of Engineering, Tokyo University of Angriculture and Technology

[2]Tensor Learning Team, RIKEN AIP

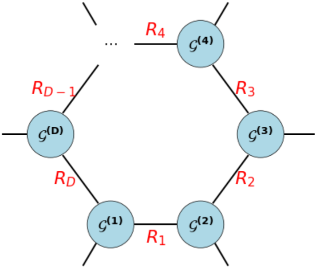
Janurary 8th, 2021

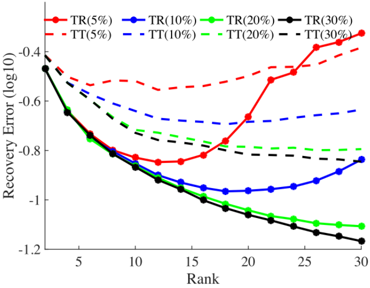
Outline

Tensor Ring Completion (Wang, et al., ICCV 2017)

For a partially observed tensor $\boldsymbol{\mathcal{Y}}$, the tensor ring completion problem is,

$$\hat{\boldsymbol{\mathcal{X}}} = \arg\min_{\boldsymbol{\mathcal{X}}} \|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{X}}\|_{\Omega}^{2}, \quad s.t., \quad \boldsymbol{\mathcal{X}} = \ll \boldsymbol{\mathcal{G}}^{(1)}, \cdots, \boldsymbol{\mathcal{G}}^{(D)} \gg .$$



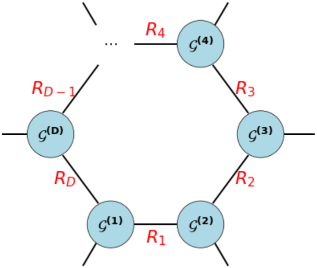


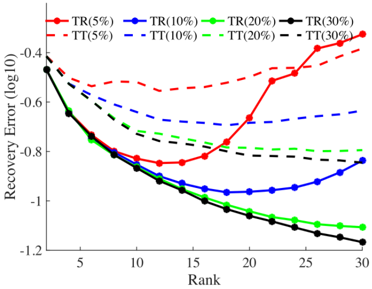
How to choose the ranks R?

Tensor Ring Completion (Wang, et al., ICCV 2017)

For a partially observed tensor $\boldsymbol{\mathcal{Y}}$, the tensor ring completion problem is,

$$\hat{\boldsymbol{\mathcal{X}}} = \arg\min_{\boldsymbol{\mathcal{X}}} \|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{X}}\|_{\Omega}^2 , \quad s.t., \quad \boldsymbol{\mathcal{X}} = \ll \boldsymbol{\mathcal{G}}^{(1)}, \cdots, \boldsymbol{\mathcal{G}}^{(D)} \gg .$$



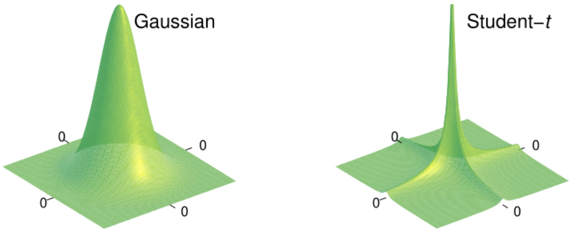


How to choose the ranks R?

Sparse Bayesian Learning (Tipping, JMLR 2001)

The automatic relevance determination (ARD) prior,

$$p(\boldsymbol{G} \mid \boldsymbol{u}) = \mathcal{N}(\boldsymbol{0}, \mathrm{diag}(1/\boldsymbol{u}^{-1})), \quad u_i \sim \Gamma(a_0^u, b_0^u).$$



Related Work

Bayesian CP Decomposition (Zhao, et al., TPAMI 2015)

Automatically select CP rank, but is less expressive than TR.

Tensor Ring Completion (Wang, et al., ICCV 2017; Yuan, et al., AAAI 2019)

Empirically select ranks or add regularization terms (still needs hyper-parameters).

Unable to infer the true rank.

Bayesian Tensor Ring Decomposition

Using the ARD prior, the Bayesian tensor ring decomposition (BTRD) is

$$\mathcal{Y}_\Omega | \{\mathcal{G}^{(i)}\}_i, \tau \sim \prod_{i_1=1}^{I_1} \cdots \prod_{i_D=1}^{I_D} \mathcal{N}(y_{\mathbf{i}} | \mathrm{tr}(\boldsymbol{G}^{(1)}[i_1] \cdots \boldsymbol{G}^{(D)}[i_D]), \tau^{-1})^{\mathcal{O}_{\mathbf{i}}},$$

$$\tau \sim \Gamma(c_0, d_0),$$

$$\boldsymbol{G}^{(d)}[i_d] | \boldsymbol{U}^{(d)}, \boldsymbol{U}^{(d+1)} \sim \mathcal{MN}(0, \boldsymbol{U}^{(d)}, \boldsymbol{U}^{(d+1)}), \quad \boldsymbol{u}_i^{(d)} \sim \Gamma(a_0, b_0),$$

for $d = 1, \ldots, D$.

Prune small factors during training.

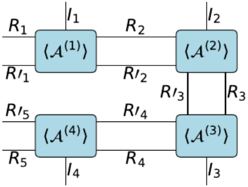
Variational Inference

If we use a *factorized* variational posterior, since the model is conjugate, the update rule is,

$$\ln q_j^*(\Theta_j) = \langle \ln p(\mathcal{Y}_\Omega, \Theta) \rangle_{q(\Theta \setminus \Theta_j)} + const.$$

Compute the expectations by tensor contraction diagram,

$$\langle (\boldsymbol{G}^{\neq 5}[\overline{\mathbf{i}_{-5}}])^{\mathsf{T}} \circ (\boldsymbol{G}^{\neq 5}[\overline{\mathbf{i}_{-5}}])^{\mathsf{T}} \rangle =$$



,

where

$$\boldsymbol{\mathcal{A}}^{(d)} = \boldsymbol{\mathcal{G}}^{(d)} \circ \boldsymbol{\mathcal{G}}^{(d)} \in \mathbb{R}^{I_d \times R_d \times R_{(d+1)} \times R_d \times R_{(d+1)}}.$$

Simulation Study

Infer the true rank of synthetic data.

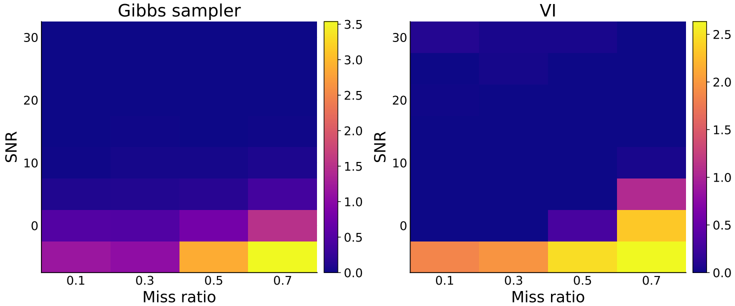


Figure 1: Rank estimation error.

Image Inpainting

Missing BTRD-GS BTRD-VI TRALS TRLRF TTWOPT BCPF

Figure 2: Image inpainting for missing ratio 0.5, 0.7, 0.9.

Thanks for your listening!