A Dual Framework for Low-Rank Tensor Completion

Madhav Nimishakavi[1] Pratik Jawanpuria[2] Bamdev Mishra[2]

[1]Indian Institute of Science [2]Microsoft, India

madhav@iisc.ac.in, {pratik.jawanpuria,bamdevm}@microsoft.com

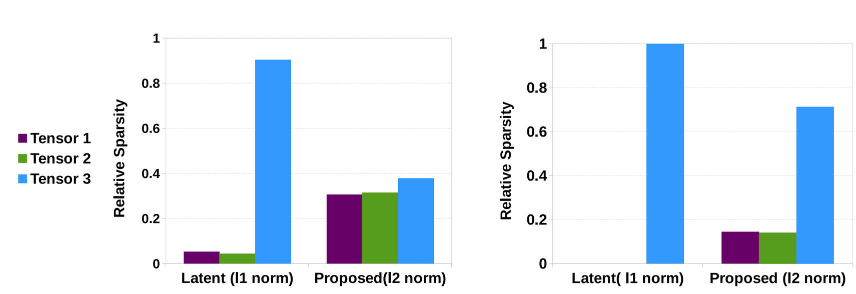
Contributions

- We propose a novel formulation for low-rank tensor completion by using a non-sparse "mixture of tensors" as the regularizer.
- We propose a dual framework for analyzing the problem formulation.
- The dual allows to develop large-scale algorithms by exploiting the versatile Riemannian optimization framework.

Motivation

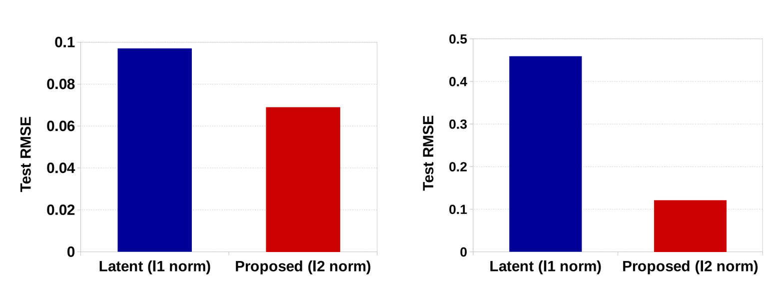
- **Latent Trace Norm:** $\min_{\substack{\mathcal{W}^{(k)}\in\mathbb{R}^{n_1\times n_2\times\ldots n_K}\\ \sum_k \mathcal{W}^{(k)}=\mathcal{W}}} \sum_k \left\|\mathcal{W}_k^{(k)}\right\|_*$
- **Proposed Regularizer:** $\min_{\substack{\mathcal{W}^{(k)}\in\mathbb{R}^{n_1\times n_2\times\ldots n_K}\\ \sum_k \mathcal{W}^{(k)}=\mathcal{W}}} \sum_k \left\|\mathcal{W}_k^{(k)}\right\|_*^2$

where $\mathcal{W}_k$ denotes mode-k unfolding. Latent trace norm is highly sparse across modes.



(a) Ribeira (b) Baboon

Relative Sparsity



(a) Ribeira (b) Baboon

Test RMSE

Non-sparse Latent Trace Norm and Duality

Given is a partially observed tensor $\mathcal{Y}_\Omega \in \mathbb{R}^{n_1\times\ldots\times n_K}$.

- **Proposed Primal Formulation**:

$$\min_{\mathcal{W}^{(k)}\in\mathbb{R}^{n_1\times n_2\times\ldots\times n_K}} \|\mathcal{W}_\Omega - \mathcal{Y}_\Omega\|_F^2 + \sum_k \frac{1}{\lambda_k}\left\|\mathcal{W}_k^{(k)}\right\|_*^2, \qquad (1)$$

where $\sum_{k=1}^{K} \mathcal{W}^{(k)} = \mathcal{W}$.

- **Equivalent Min-Max Formulation**:

$$\min_{\Theta_1,\ldots,\Theta_K} \max_{\mathcal{Z}\in\mathcal{C}} \langle \mathcal{Z}, \mathcal{Y}_\Omega\rangle - \frac{1}{4}\|\mathcal{Z}\|_F^2 - \sum_k \frac{\lambda_k}{2}\langle \Theta_k, \mathcal{Z}_k\mathcal{Z}_k^\top\rangle,$$

where $\mathcal{Z}$ is sparse dual tensor variable and Θ_k are $n\times n$, PSD auxiliary matrices with unit trace.

- **Representer Theorem**: The optimal solution of the primal problem (1) admits a representation of the form:

$$\mathcal{W}^{(k)*} = \lambda_k(\mathcal{Z}\times_k \Theta_k)\forall k,$$

where the set $\mathcal{C}$ constrains $\mathcal{Z}$ to be sparse tensor with $|\Omega|$ non-zero entries.

Fixed-rank Dual Formulation

- Constraining the rank of Θ_k by $\Theta_k = \mathbf{U}_k\mathbf{U}^\top$, where $\mathbf{U}_k \in \mathcal{S}_{r_k}^{n_k}$ and $\mathcal{S}_n^r := \{\mathbf{U}\in\mathbb{R}^{n\times r} : \|\mathbf{U}\|_F = 1\}$ leads to

$$\min_{u\in\mathcal{S}_{r_1}^{n_1}\times\ldots\times\mathcal{S}_{r_K}^{n_K}} g(u),$$

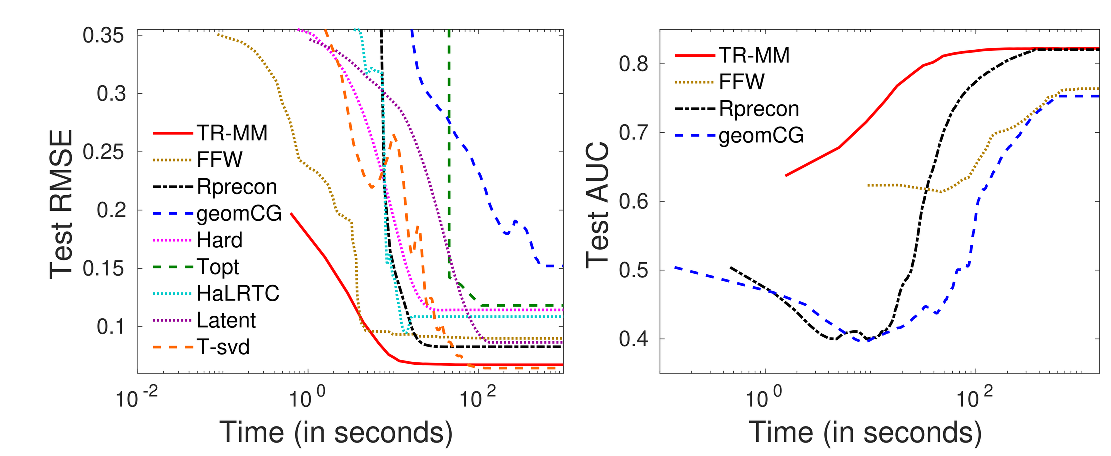
where $u = (\mathbf{U}_1,\ldots,\mathbf{U}_K)$ and g is the convex function

$$g(u) := \max_{\mathcal{Z}\in\mathcal{C}} \langle\mathcal{Z},\mathcal{Y}_\Omega\rangle - \frac{1}{4}\|\mathcal{Z}\|_F^2 - \sum_k \frac{\lambda_k}{2}\left\|\mathbf{U}_k^\top\mathcal{Z}_k\right\|_F^2.$$

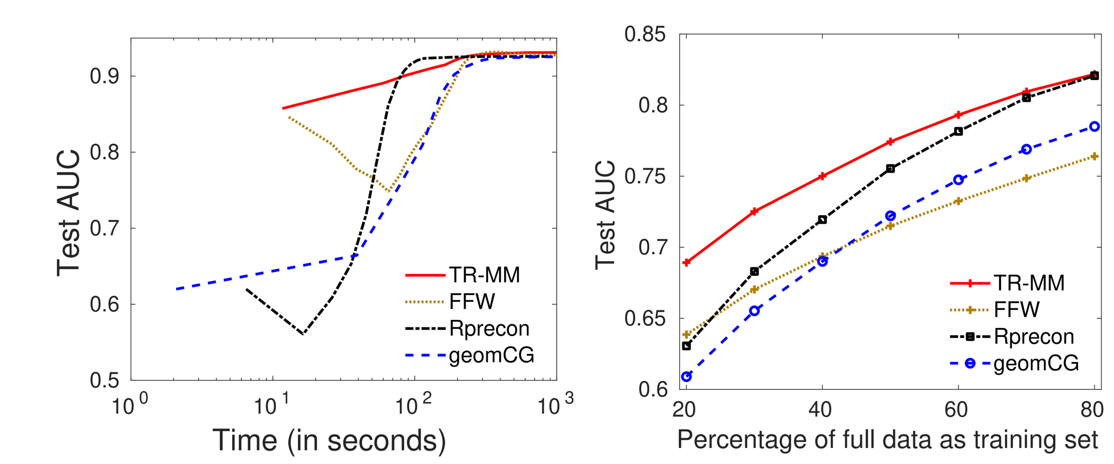
Proposed Riemannian Trust Region algorithm (TR-MM)

Input: $\mathcal{Y}_\Omega$, rank $(r_1,\ldots,r_K)$, regularization parameter λ, and tolerance ϵ.
Initialize : $u = (\mathbf{U}_1,\ldots,\mathbf{U}_K) \in \mathcal{S}_{r_1}^{n_1}\times\ldots\times\mathcal{S}_{r_K}^{n_K}$.
repeat
 1: Solve for $\mathcal{Z}$ by computing $g(u)$
 2: Compute $\nabla_u g(u)$.
 3: Riemannian TR step: Compute the search direction $v = (\mathbf{V}_1,\ldots,\mathbf{V}_K)$ which minimizes the trust-region subproblem and set step size $\alpha = 1$. It makes use of $\nabla_u g(u)$ and its directional derivative $D\nabla_u g(u)[v]$.
 4: Update u as $\mathbf{U}_k \leftarrow (\mathbf{U}_k + \alpha\mathbf{V}_k)/\left\|\mathbf{U}_k+\alpha\mathbf{V}_k\right\|_F$, for all k.
until $\|\nabla_u g(u)\| < \epsilon$
Output: $u^* = (\mathbf{U}_1^*,\ldots,\mathbf{U}_K^*)$ and $\mathcal{Z}^*$.
Solution to primal problem :
 $\mathcal{W}^{(k)*} = \lambda_k(\mathcal{Z}^*\times_k \mathbf{U}_k^*\mathbf{U}_k^{*\top})$ and $\mathcal{W}^* = \sum_k \mathcal{W}^{(k)*}$.

Experiments



(a) Ribeira (b) FB15k-237



(c) YouTube (full) (d) FB15k-237

(a) Evolution of test RMSE on Ribeira; (b) & (c) Evolution of test AUC on FB15k-237 and YouTube, respectively. (d) Variation of test AUC as percentage of training data changes on FB15k-237.

Generalization performance across several applications: hyperspectral-image/video/image completion, movie recommendation, and link prediction.

	TR-MM	FFW	Rprecon	geomCG	Hard	Topt	HaLRTC	Latent	T-svd	BayesCP
RMSE reported										
Ribeira	0.067	0.088	0.083	0.156	0.114	0.127	0.095	0.087	**0.064**	0.154
Tomato	**0.041**	0.045	0.052	0.052	0.060	0.102	0.202	0.046	0.042	0.103
Baboon	**0.121**	0.133	0.128	0.128	0.126	0.130	0.247	0.459	0.146	0.159
ML10M	0.840	0.895	**0.831**	0.844	–	–	–	–	–	–
AUC reported										
YouTube (subset)	**0.957**	0.954	0.941	0.941	0.954	0.941	0.783	0.945	0.941	0.950
YouTube (full)	**0.932**	0.929	0.926	0.926	–	–	–	–	–	–
FB15k-237	**0.823**	0.764	0.821	0.785	–	–	–	–	–	–

Results on outlier robustness experiments.

	x	TR-MM	FFW	Rprecon	geomCG	Hard	Topt	HaLRTC	Latent	T-svd	BayesCP
Ribeira	0.05	**0.081**	0.095	0.157	0.258	0.142	0.169	0.121	0.103	0.146	0.201
(RMSE)	0.10	**0.111**	0.112	0.172	0.373	0.158	0.188	0.135	0.120	0.182	0.204
FB15k-237	0.05	**0.803**	0.734	0.794	0.764	–	–	–	–	–	–
(AUC)	0.10	**0.772**	0.711	0.765	0.739	–	–	–	–	–	–

- Our implementation uses the Manopt toolbox at `manopt.org`.
- Codes: `https://pratikjawanpuria.com/`.