A Dual Framework for Low-Rank Tensor Completion

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

madhav@iisc.ac.in, {jawanpur,bamdevm}@amazon.com

[1]Indian Institute of Science [2]Amazon.com

Introduction

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

Novel Formulation

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

- **Primal Formulation**:

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

where $\sum_{k=1}^K \mathcal{W}^{(k)} = \mathcal{W}$ and $\mathcal{W}_k$ denotes mode-k unfolding.

- We exploit the characterization:

$$\|\mathbf{X}\|_*^2 = \min_{\Theta\in\mathcal{P}^d,\mathrm{range}(\mathbf{X})\subseteq\mathrm{range}(\Theta)} \langle\Theta^\dagger, \mathbf{X}\mathbf{X}^\top\rangle,$$

where Θ is an auxiliary variable with $\mathrm{trace}(\Theta)=1$.

- **Equivalent Min-Max Formulation**:

$$\min_{\theta\in\times_{k=1}^K\mathcal{P}^{n_k}} \max_{\mathcal{Z}\in\mathcal{D}} \langle\mathcal{Z},\mathcal{Y}_\Omega\rangle - \frac{1}{4}\|\mathcal{Z}\|_F^2 - \sum_k\frac{\lambda_k}{2}\langle\Theta_k, \mathcal{Z}_k\mathcal{Z}_k^\top\rangle,$$

where $\mathcal{Z}$ is sparse dual tensor variable and Θ_k are auxiliary variables for all k.

Optimization on Spectrahedron Manifold $\mathcal{S}$

- 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

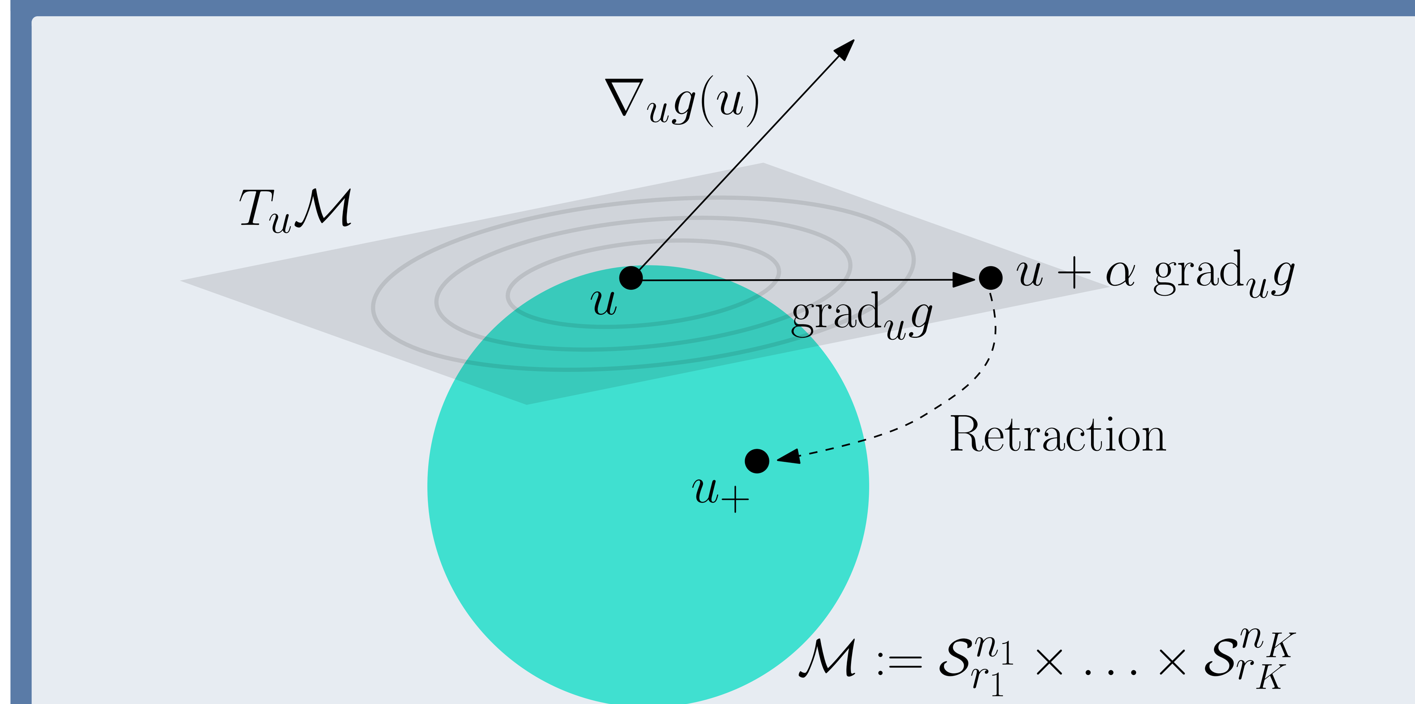
$$\min_{u\in\mathcal{S}_{r_1}^{n_1}\times\ldots\times\mathcal{S}_{r_K}^{n_K}} g(u),$$

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

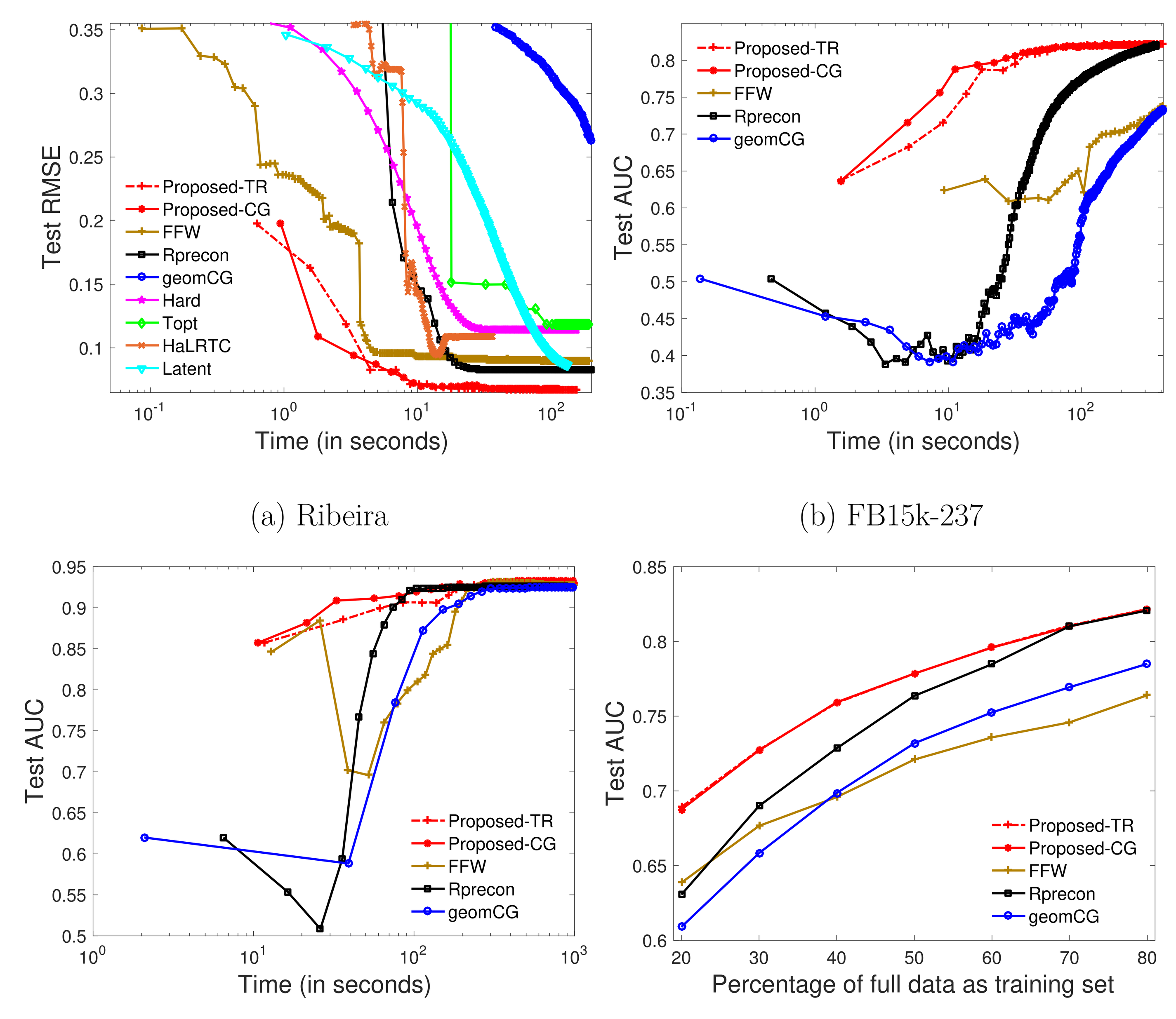
$$g(u) := \max_{\mathcal{Z}\in\mathcal{D}} \langle\mathcal{Z},\mathcal{Y}_\Omega\rangle - \frac{1}{4}\|\mathcal{Z}\|_F^2 - \sum_k\frac{\lambda_k}{2}\|\mathbf{U}_k^\top\mathcal{Z}_k\|_F^2.$$

- Our algorithms use the Manopt toolbox at `manopt.org`.
- Codes: `madhavcsa.github.io`

A schematic view of the optimization on a manifold $\mathcal{M}$.



Results



(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.

Mean test RMSE for hyperspectral-image completion, video completion and recommendation problems.

	Ribeira	Image	MovieLens10M
Proposed-TR	**0.067**	**0.068**	0.840
Proposed-CG	0.068	**0.068**	0.841
FFW	0.088	0.120	0.895
Rprecon	0.083	0.085	**0.831**
geomCG	0.156	0.085	0.844

Mean test AUC for link prediction problem.

	YouTube (subset)	YouTube (full)	FB15k-237
Proposed-TR	**0.957**	**0.932**	**0.823**
Proposed-CG	**0.957**	**0.932**	**0.823**
FFW	0.954	0.929	0.764
Rprecon	0.941	0.926	0.821
geomCG	0.941	0.926	0.785

Proposed Riemannian Conjugate Gradient (CG) and Trust Region (TR) algorithms

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 CG step: Compute the conjugate direction $v = (\mathbf{V}_1,\ldots,\mathbf{V}_K)$ and step size α using the Armijo line search, which makes use of $\nabla_u g(u)$.

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 (retraction step).
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 $\mathcal{W}^* = \sum_k\mathcal{W}^{(k)*}$.