A New Retraction for Accelerating the Riemannian Three-Factor Low-Rank Matrix Completion Algorithm

Zhizhong Li[1,2], Deli Zhao[2], Zhouchen Lin[3,4], Edward Y. Chang[5]
[1]School of Mathematics, Peking University. [2]Advanced Algorithm Research Group, HTC.
[3]Key Lab. of Machine Perception (MOE), School of EECS, Peking University. [4]Cooperative Medianet Innovation Center, China. [5]HTC Research.

Geometric optimization algorithms on matrix manifolds have been applied to the low-rank matrix completion problem in machine learning. Various geometric optimization algorithms, such as LRGeomCG [5], RTRMC [2], ScGrass-CG [4], and R3MC [3], are available for solving this problem. R3MC, the Riemannian three-factor matrix completion algorithm, is one of the state-of-the-art geometric optimization methods for the low-rank matrix completion problem. This paper improves the performance of R3MC by proposing a new retraction with a minimizing property. Accelerated R3MC, which is R3MC equipped with this new retraction, outperforms the original algorithm and other geometric algorithms for matrix completion in our empirical study.

The formulation of the matrix completion problem in this paper is to find a rank-r matrix X such that

$$\min_{X\in\mathbb{R}_r^{n\times m}} f(X) := \frac{1}{|\Omega|}\|\mathcal{P}_\Omega(X)-\mathcal{P}_\Omega(X^*)\|_F^2, \tag{1}$$

where X^* is the partially known matrix to be completed, Ω is the set of indices for known entries of X^* and $|\Omega|$ the cardinality of Ω. $\mathbb{R}_r^{n\times m}$ is the set of rank-r matrices of size $n\times m$, $\|A\|_F$ is the Frobenius norm of matrix A, and $\mathcal{P}_\Omega$ is the orthogonal sampling operator.

The manifold $\mathcal{M}$ in R3MC is one that is homeomorphic to the manifold of rank-r matrices $\mathbb{R}_r^{n\times m}$,

$$\mathcal{M} := (\mathrm{St}(r,n)\times\mathrm{GL}(r)\times\mathrm{St}(r,m))/(\mathcal{O}(r)\times\mathcal{O}(r)), \tag{2}$$

where $\mathrm{St}(r,n) := \{X\in\mathbb{R}^{n\times r}|X^TX = I_r\}$ is a Stiefel manifold, $\mathrm{GL}(r) := \{X\in\mathbb{R}^{r\times r}||X|\neq 0\}$ is a general linear group and $\mathcal{O}(r) := \{X\in\mathbb{R}^{r\times r}|X^TX = I_r\}$ is an orthogonal group. $|X|$ is the determinant of matrix X and I_r is the identity matrix of size $r\times r$. The quotient originates from the following group action

$$\begin{aligned}&(\mathcal{O}(r)\times\mathcal{O}(r))\times\overline{\mathcal{M}}\to\overline{\mathcal{M}}\\&((O_1,O_2),(U,R,V))\mapsto(UO_1,O_1^TRO_2,VO_2),\end{aligned} \tag{3}$$

where $(U,R,V)\in\overline{\mathcal{M}}$, $(O_1,O_2)\in\mathcal{O}(r)\times\mathcal{O}(r)$ and $\overline{\mathcal{M}} := \mathrm{St}(r,n)\times\mathrm{GL}(r)\times\mathrm{St}(r,m)$. $\overline{\mathcal{M}}$ is the *total space* of the quotient manifold $\mathcal{M}$. It is convenient to describe the algorithm using various lifted objects (often with 'bar' notations) in the total space because they are more concrete to manipulate. One example is the lifted cost function $\bar{f}$ for Equation (1),

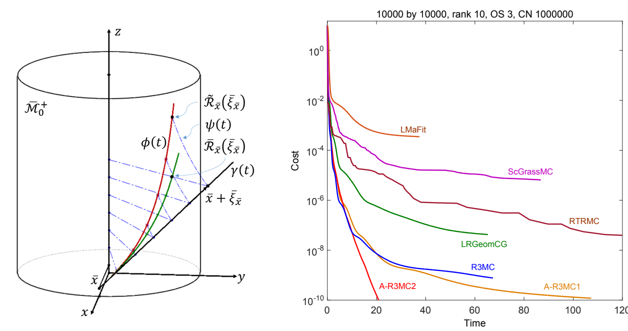
$$\bar{f}(U,R,V) := \frac{1}{|\Omega|}\|\mathcal{P}_\Omega(URV^T)-\mathcal{P}_\Omega(X^*)\|_F^2, \tag{4}$$

where $(U,R,V)\in\overline{\mathcal{M}}$.

Retraction [1, Definition 4.1.1] is one of the most important ingredients for geometric optimization algorithms. It provides a mechanism to move along a direction while constrained to the manifold. We first define the new retraction $\widetilde{\mathcal{R}}$ on the total space $\overline{\mathcal{M}}$ as as

$$\begin{aligned}\widetilde{\mathcal{R}}_{\bar{x}} : T_{\bar{x}}\overline{\mathcal{M}}&\to\overline{\mathcal{M}}\\ \bar{\xi}_{\bar{x}}&\mapsto(P_1,Q_1(R+\bar{\xi}_R)Q_2^T,P_2),\end{aligned} \tag{5}$$

where $\bar{x} = (U,R,V)\in\overline{\mathcal{M}}$, $\bar{\xi}_{\bar{x}} = (\bar{\xi}_U,\bar{\xi}_R,\bar{\xi}_V)$ is the given direction and P_1, Q_1, P_2 and Q_2 are defined by polar decompositions $U+\bar{\xi}_U = P_1Q_1$ and $V+\bar{\xi}_V = P_2Q_2$, respectively. This retraction can induce a retracion $\mathcal{R}$ on the quotient space $\mathcal{M}$. Illustration of the new retraction and the original one is shown in Figure 1(a).



(a) (b)

Figure 1: (a) The new retraction $\widetilde{\mathcal{R}}$ and the original retraction $\overline{\mathcal{R}}$ on the cylinder $\overline{\mathcal{M}}_0^+$. $\bar{x}$ is a point on the cylinder, $\gamma(t)$ a line passing through $\bar{x}$ with direction $\bar{\xi}_{\bar{x}}$. The upper red curve $\phi(t)$ is the retracted curve $\phi(t) := \widetilde{\mathcal{R}}_{\bar{x}}(t\bar{\xi}_{\bar{x}})$ under retraction $\widetilde{\mathcal{R}}$, and the lower green curve is the one generated by the retraction $\overline{\mathcal{R}}$. (b) Comparisons of A-R3MC1/A-R3MC2 with R3MC, LRGeomCG [5], RTRMC [2], ScGrassMC [4] and LMaFit [6]. A-R3MC1 and A-R3MC2 are two variants of the accelerated R3MC.

A special property of $\widetilde{\mathcal{R}}$ can be observed in Figure 1(a). In fact, points on the blue dashed curve $\psi(t)$ connecting $\gamma(t)$ and $\phi(t)$ share the same function value of $\bar{f}$. The invariance of the cost function on curve ψ establishes a correspondence between points on γ and points on ϕ. This correspondence can induce an minimizing property that allows exact minimizations for the line-search steps of R3MC. This minimizing property is stated as follows:

Proposition 1. *Suppose that $\bar{x}\in\overline{\mathcal{M}}$ and $\bar{\eta}_{\bar{x}}\in\mathcal{H}_{\bar{x}}$, where $\mathcal{H}_{\bar{x}}$ is the horizontal space. Let $\phi(t) := \widetilde{\mathcal{R}}_{\bar{x}}(t\bar{\eta}_{\bar{x}})$ be the retracted curve of line $t\mapsto t\bar{\eta}_{\bar{x}}$ and $\gamma(t) : t\mapsto\bar{x}+t\bar{\eta}_{\bar{x}}$ be the tangent line passing through $\bar{x}$. If the solution of $\min_t\bar{f}(\gamma(t))$ is t_*, then the minimum point of $\bar{f}$ restricted on curve ϕ is $\widetilde{\mathcal{R}}_{\bar{x}}(t_*\bar{\eta}_{\bar{x}})$.*

Equipped with this new retraction, R3MC is able to find an optimal step length for the line search step. We call this modified version A-R3MC. The performance comparisons of various algorithms are shown in Figure 1(b). We can see that the new retraction with minimizing property performs well in the experiment.

[1] P. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, 2009.

[2] Nicolas Boumal and Pierre-Antoine Absil. RTRMC: A Riemannian trust-region method for low-rank matrix completion. In *Advances in Neural Information Processing Systems 24*. 2011.

[3] B. Mishra and R. Sepulchre. R3MC: A Riemannian three-factor algorithm for low-rank matrix completion. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, 2014.

[4] Thanh Ngo and Yousef Saad. Scaled gradients on Grassmann manifolds for matrix completion. In *Advances in Neural Information Processing Systems 25*. 2012.

[5] Bart Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214, 2013.

[6] Zaiwen Wen, Wotao Yin, and Yin Zhang. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. *Mathematical Programming Computation*, 4(4): 333–361, 2012.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.