East Asian Journal on Applied Mathematics Vol. **16**, No. 4, pp. 888-906
doi: 10.4208/eajam.2025-229.010226 November 2026

Riemannian Gradient Method for Wasserstein Barycenters under the Affine-Invariant Geometry

Donghee Yang[1], Yoon Mo Jung[2] and Sangwoon Yun[3,*]

[1] *Institute of Basic Science, Sungkyunkwan University, Suwon-si 16419, Korea.*
[2] *Department of Mathematics, Sungkyunkwan University, Suwon-si 16419, Korea.*
[3] *Department of Mathematics Education, Sungkyunkwan University, Seoul 03063, Korea.*

Received 16 November 2025; Accepted (in revised version) 1 February 2026.

Abstract. We study a Riemannian gradient method for the L_2-Wasserstein least squares problem of Gaussian measures under the affine-invariant geometry. The variable of L_2-Wasserstein least squares problem lies in the set of positive definite matrices, which, equipped with the affine-invariant metric, forms a Hadamard manifold. The same set with usual Euclidean metric is also a Hadamard manifold, with constant sectional curvature equal to 0. Hence, the gradient descent method proposed in [S. Kum, S. Yun, J. Korean Math Soc. **56** (2019)] can be considered as a Riemannian gradient method with respect to usual Euclidean inner product. This method is known to have a sublinear convergence rate and requires a singular value decomposition at each iteration, which is computationally expensive. In this paper, we adapt the Riemannian gradient method under the affine-invariant geometry for solving the L_2-Wasserstein least squares and prove its local linear convergence. This method does not require a singular value decomposition. We numerically show that the proposed method is more efficient than the gradient descent method mentioned.

AMS subject classifications: 90C31, 90C25
Key words: Riemannian gradient method, Wasserstein barycenter, geodesic convexity, affine-invariant metric.

1. Introduction

The minimization problem

$$\inf_{\rho\in\prod(\mu,\nu)}\int_{\mathbb{R}^m\times\mathbb{R}^m}\|x-y\|d\rho(x,y),$$

where $\prod(\mu,\nu)$ is the set of transport plans between the probability measures μ and ν, is the Monge-Kantorovich problem for probability measures on $\mathbb{R}^m$ with finite second moments.

*Corresponding author. *Email addresses:* `donny@skku.edu` (D. Yang), `yoonmojung@skku.edu` (Y.M. Jung), `yswmathedu@skku.edu` (S. Yun)

The multimarginal optimal transport problem is expressed as

$$\min_{\sigma\in\prod(\mu_1,\dots,\mu_n)} \int_{(\mathbb{R}^m)^n} \left(\sum_{i=1}^{n} w_i \left\| x_i - \sum_{j=1}^{n} w_j x_j \right\|^2 \right) d\sigma(x_1,\dots,x_n),$$

where each μ_i is a probability measure on $\mathbb{R}^m$ with a finite second moment, $(w_1,\dots,w_n)^T$ is a probability vector in $\mathbb{R}^n$, and $\prod(\mu_1,\dots,\mu_n)$ denotes the set of probability measures on $(\mathbb{R}^m)^n$ with marginals $\mu_1,\dots,\mu_n$.

This problem is closely related to the following minimization problem:

$$\min_{\mu} \sum_{i=1}^{n} w_i d_W^2(\mu,\mu_j), \tag{1.1}$$

where d_W denotes the Wasserstein distance. The problem (1.1) has been studied intensively. Agueh and Carlier [2] proved the existence and uniqueness of a minimizer of the problem (1.1) under the absolutely continuity of some μ_j.

The Wasserstein distance between zero mean Gaussian measures μ and ν is given by

$$d_W(X,S) = \sqrt{\operatorname{tr}(X+S) - 2\operatorname{tr}(X^{1/2}SX^{1/2})^{1/2}},$$

where X and S are the covariance matrices of μ and ν, respectively [9]. The Wasserstein barycenter problem of Gaussian measures is an important problem for various applications such as quantum information, statistics, and optimal transport [6, 8, 10, 16]. This problem can be formulated as the following minimization problem — i.e. L_2-Wasserstein least squares problem:

$$\min_{X\succ 0} f(X) := \sum_{j=1}^{n} \frac{1}{2} w_j d_W^2(X,S_j), \tag{1.2}$$

where the weights satisfy $\sum_{j=1}^{n} w_j = 1$ with $w_j \geq 0$, and each S_j is an $m\times m$ positive definite matrix. This problem is also known as the n-coupling problem. We denote the set of $m\times m$ positive definite matrices by $\mathscr{S}_{++}^m$ and the set of $m\times m$ symmetric matrices by $\mathscr{S}^m$.

Kum and Yun [11] proposed a gradient projection method and an accelerated gradient projection method to solve the problem (1.2). They proved that the $\mathscr{O}(1/k)$ sublinear convergence rate for the gradient projection method and the $\mathscr{O}(1/k^2)$ sublinear convergence rate for the accelerated version.

The variable of the problem (1.2) is a positive definite matrix. The set of positive definite matrices with usual Euclidean metric — i.e. $\langle U, V\rangle = \operatorname{tr}(UV)$ for $U,V\in\mathscr{S}^m$, is a Hadamard manifold. The same set with the affine-invariant metric — i.e. $\langle U, V\rangle_{ai,X} = \operatorname{tr}(X^{-1}UX^{-1}V)$ for $U,V\in\mathscr{S}^m$, is also a Hadamard manifold. The gradient projection methods in [11] can be considered as a Riemannian gradient method in which the retraction is implemented by the projection operator. However, these methods require a singular value decomposition (SVD) at each iteration, and the computational cost of this SVD is expensive.

In this paper, we propose a Riemannian gradient method under the affine-invariant metric to solve the problem (1.2). The method uses the exponential map as a retraction