Filomat 39:24 (2025), 8509–8517
https://doi.org/10.2298/FIL2524509N

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Barycenters of Toeplitz matrices and application in clustering

Elham Nobari[a,*], Mohammad Javad Nazari[b]

[a] *Department of Mathematics, University of Science and Technology of Mazandaran, Behshahr, Iran.*
[b] *Department of Computer Science, Amirkabir University of Technology, Hafez Ave., Tehran, Iran.*

Abstract. This paper presents two innovative centering notions, the p-barycenter and the L^p-center of mass, for Toeplitz matrices. The p-barycenter employs a distance function that relies on symbol functions, while the L^p-center of mass is based on the Riemannian distance on the manifold of positive definite matrices. Our proposed methods extend the k-means machine learning algorithm to Toeplitz matrices, thereby enabling potential applications in various fields, including signal processing. Furthermore, when $p = 2$, one of the resulting objects is the geometric mean or Karcher mean, which is also a Toeplitz matrix. These centering notions have great potential for enhancing the performance of clustering algorithms on Toeplitz matrices and can be applied in areas such as image processing, audio signal processing, and time series analysis.

1. Introduction

Toeplitz matrices appear in a wide range of applications and admit interesting and deep theoretical and analytical aspects. There is a great research literature surrounding the Toeplitz matrices [9, 11–13, 16]. In this paper, we consider two notions of barycenters for Toeplitz matrices as described below. In the first one, we introduce a distance function as follows

$$d_p\left(T_n(a), T_n(b)\right) = \left\| a\left(e^{i\theta}\right) - b\left(e^{i\theta)}\right)\right\|_p,$$

where a and b are the symbol functions of a Toeplitz matrices $T_n(a)$ and $T_n(b)$ respectively; see §2. We show the existence and uniqueness of this barycenter. In the second one, inspired by what is mentioned in [1] and using the distance

$$d(A, B) := \left\| \log\left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right)\right\|_F,$$

between two positive definite Toeplitz matrices A, B. We demonstrate that the minimizer of the following problem

$$B_p\left(A_1, \ldots, A_m\right) := \underset{X \in \mathcal{T}_n^+}{\operatorname{argmin}} \frac{1}{p} \sum_{i=1}^{m} d^p\left(X, A_i\right),$$

2020 *Mathematics Subject Classification*. Primary 15B05; Secondary 15B48, 65F99.
Keywords. Toeplitz matrix, L^p-barycenter, p-barycenter, Karcher mean, geometric mean, k-means.
Received: 01 April 2024; Revised: 15 January 2025; Accepted: 18 July 2025
Communicated by Marko Petković
* Corresponding author: Elham Nobari
Email addresses: enobary@gmail.com (Elham Nobari), mohammadjnazari@aut.ac.ir (Mohammad Javad Nazari)
ORCID iDs: https://orcid.org/0000-0001-6625-8336 (Elham Nobari), https://orcid.org/0000-0001-6842-3657 (Mohammad Javad Nazari)

exists, it is unique, and it belongs to the convex hull of the set $A_1, \ldots, A_m$, where A_i, $1 \leq i \leq m$, are positive definite Toeplitz matrices. Yann Cabanes [11], explores the generalization of classical machine learning algorithms to the Riemannian manifold. This approach has been successfully applied to visualize data in two dimensions and various other subjects. Additionally, the concept of barycenters can be utilized for clustering datasets that contain Toeplitz matrices; see [11]. Another topic that has recently received a lot of attention and research is calculating the geometric mean of positive definite matrices [2, 3, 6, 7, 14, 15]. In the case of Toeplitz matrices, the main challenge is whether the geometric mean remains in Toeplitz form or not, as well as is satisfying in the Ando–Li–Mathias (ALM) axioms [3]. For discerning some investigation and progress in this regard, see e.g., [7, 12, 18, 19].

The present work is organized as follows. In Section 2, we state some basic definitions. The notion of p-barycenter and some related results are discussed in Section 3. In Subsection 4.1, the concept of L^p barycenter is presented, and using the geometric mean in Subsection 4.2, we define the geometric mean of Toeplitz matrices which is a Toeplitz matrix itself. The paper concludes with some numerical experiments and comparisons between our results and previous and recent results from the relevant literature.

2. Preliminaries

Let $n \geq 1$. A Toeplitz matrix is an $n \times n$ matrix $A = [a_{ij}]$, where entries along their diagonals are constant, i.e., a matrix of the form:

$$\begin{bmatrix} a_0 & a_1 & \cdots & & a_{n-1} \\ a_{-1} & a_0 & a_1 & \cdots & \\ \vdots & & \ddots & & \vdots \\ & & \ddots & & a_1 \\ a_{-n+1} & & \cdots & a_{-1} & a_0 \end{bmatrix}.$$

The set of all $n \times n$ Toeplitz matrices from a vector space that we denote it by $\mathcal{T}_n$. Let $\mathbb{T} = \{e^{i\theta}, \theta \in \mathbb{R}\}$ be the unit circle in the complex plane. Corresponding to matrix A, the symbol function $a : \mathbb{T} \to \mathbb{C}$ is defined by:

$$a(t) = \sum_{|k|<n} a_k t^k \quad (t \in \mathbb{T}), \tag{1}$$

or

$$a(e^{i\theta}) = \sum_{k=-n+1}^{n-1} a_k e^{ik\theta} \quad (\theta \in \mathbb{R}). \tag{2}$$

Conversely, for each trigonometric polynomial of degree at most $n-1$ akin (1), one can define its associated Toeplitz matrix in manner. We denote the vector space of all trigonometric polynomials of degree at most $n-1$ by $\mathcal{P}_{n-1}$. Hence, there is a one-to-one correspondence between $\mathcal{P}_{n-1}$ and $\mathcal{T}_n$:

$$\begin{aligned} &\mathcal{P}_{n-1} \to \mathcal{T}_n, \\ &\; a \mapsto T_n(a). \end{aligned} \tag{3}$$

Suppose that $a : \mathbb{T} \to \mathbb{C}$ is a Lebesgue measurable function and for each $1 \leq p < \infty$, let

$$\|a\|_p = \left(\frac{1}{2\pi} \int_0^{2\pi} |a(e^{i\theta})|^p d\theta\right)^{\frac{1}{p}}. \tag{4}$$

We denote the set of all Lebesgue measurable functions $a : \mathbb{T} \to \mathbb{C}$ with $\|a\|_\infty < \infty$ by $L^P(\mathbb{T})$. It is Know that $L^P(\mathbb{T})$ equipped with norm $\|.\|$ is a Banach space. Define p-norms

$$\|T_n(a)\|_p := \|a\|_{L^p(\mathbb{T})}, \tag{5}$$

on $\mathcal{T}_n$, the space $\mathcal{T}_n$ with these p-norms is a Banach space too; see [9].

Definition 2.1. *Let $T_n(a)$, $T_n(b) \in \mathcal{T}_n$ and $1 \leq p < \infty$ are given, we define p-distance between $T_n(a)$ and $T_n(b)$ by the formula*

$$d_p(T_n(a), T_n(b)) = \|a(e^{i\theta}) - b(e^{i\theta)})\|_p. \tag{6}$$

Suppose that $T_n(a_1), \dots, T_n(a_m)$ belong to $\mathcal{T}_n$. We denote the convex hull of matrices $T_n(a_1), \dots, T_n(a_m)$ by $S := \mathsf{conv}(T_n(a_1), \dots, T_n(a_m))$.

3. *p*-barycenter

In this section using the norm (5), we define a new notion of center, called p-barycenter. The existence and uniqueness will be proved.

Definition 3.1. *Let $m, n \geq 1$ and $1 \leq p < \infty$. Given Toeplitz matrices $T_n(a_1), \dots, T_n(a_m)$. Suppose that S is convex hull of $T_n(a_1), \dots, T_n(a_m)$. The p-barycenter of $T_n(a_1), \dots, T_n(a_m)$ is:*

$$B_p(T_n(a_m), \dots, T_n(a_m)) := \underset{B \in \mathcal{T}_n}{\operatorname{argmin}} \frac{1}{m} \sum_{i=1}^{m} d_p^2(B, T_n(a_i)). \tag{7}$$

Now, assume that $1 < p < \infty$. It is well-known that the L^p-norm function $x \mapsto \|x\|_p$ is strictly convex. Fix $A \in \mathcal{T}_n$ and let $a(e^{i\theta})$ be its symbol function. The function

$$\begin{aligned} &d_p^2(., A) : \mathcal{T}_n \to \mathbb{R}, \\ &d_p^2(X, A) = \|x(e^{i\theta}) - a(e^{i\theta})\|_p^2, \quad \theta \in \mathbb{R}, \end{aligned} \tag{8}$$

is also strictly convex function; see [10], p.29.
The existence and uniqueness of the optimizer matrix in (7) follows from the following proposition.

Proposition 3.2. *([10], p.71). Let E be a reflexive Banach, and $A \subset E$ be a nonempty, closed, convex subset of E. Let also $\varphi : A \longrightarrow (-\infty, +\infty]$ be a convex lower semi-continuous function such that $\varphi \not\equiv \infty$ (i.e. domain(φ) $\neq \emptyset$) and*

$$\begin{aligned} &\lim_{x \in A} \varphi(x) = +\infty \quad \textit{(no assumption if A is bounded)}, \\ &\|x\| \longrightarrow \infty. \end{aligned}$$

Then φ achieves its minimum on A, i.e. there exists some $x_0 \in A$ such that $\varphi(x_0) = \min_A \varphi$.

Theorem 3.3. *With assumptions of Definition* (3.1), *for $1 < p < \infty$ the problem* (7) *has a unique solution.*

Proof. Put $A = E = \mathcal{T}_n$ and

$$\begin{aligned} &\varphi : \mathcal{T}_n \longrightarrow \mathbb{R} \\ &\varphi(X) = \frac{1}{m} \sum_{i=1}^{m} d_p^2(X, T_n(a_i)), \end{aligned}$$

in the Proposition 3.2. Since φ is a continues and convex function, the required conditions of the Proposition 3.2 are established. So for $1 \leq p \leq \infty$, φ achieves its minimum on $\mathcal{T}_n$, i.e. there exists some $B_0 \in \mathcal{T}_n$ such that $\varphi(B_0) = \min_{\mathcal{T}_n} \varphi$. On the other hand, since for $1 < p < \infty$, d_p^2 is a strictly convex function, the problem (7) has a unique solution. □

In order to make a method to numerically compute the solution of (3.1), we may confine ourselves to the case that B belongs to S. Putting

$$B = t_1 T_n(a_1) + \cdots + t_m T_n(a_m), \tag{9}$$

where

$$t_1, \ldots, t_m \geq 0, \quad t_1 + \cdots + t_m = 1,$$

are unknown. An assuming

$$T_n(a_k(e^{i\theta})) = \begin{bmatrix} a_{0k} & a_{1k} & \cdots & a_{(n-1)k} \\ a_{-1k} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{-(n-1)k} & \cdots & & a_{0k} \end{bmatrix}, \quad 1 \leq k \leq m, \tag{10}$$

we get

$$\begin{aligned} \|B - T_n(a_j)\|_p &= \|\sum_{k=1}^{m} t_k T_n(a_k) - T_n(a_j)\|_p \\ &= \|(t_j - 1)T_n(a_j) + \sum_{k=1, k\neq j}^{m} t_k T_n(a_k)\|_p \\ &= \left(\frac{1}{2\pi}\int_0^{2\pi} |(t_j - 1)a_j(e^{i\theta}) + \sum_{k=1,k\neq j}^{m} t_k a_k(e^{i\theta})|^p\right)^{\frac{1}{p}} \end{aligned} \tag{11}$$

we recall the relation in (2) and we rewrite it as follows

$$a_j(e^{i\theta}) = \sum_{l=-n+1}^{n-1} a_{lj} e^{li\theta},$$

where a_{lj}, $-n+1 \leq l \leq n+1$, are entries of $T_n(a_j)$. Now, we introduce the function $g_j(\theta)$, $1 \leq j \leq m$, as follows:

$$\begin{aligned} g_j(t_1, ..., t_m) &= \|B - T_n(a_j)\|_p^p \\ &= \frac{1}{2\pi}\int_0^{2\pi} (|(t_j - 1)\sum_{l=-n+1}^{n-1} a_{lj}e^{li\theta} + \sum_{k=1,k\neq j}^{m} t_k \sum_{l=-(n-1)}^{l=n-1} a_{lj}e^{li\theta}|)^p. \end{aligned}$$

In order to calculate the functions $g_j(t_1, ..., t_m)$, $1 \leq j \leq m$, we need an appropriate approximating integration method. In Section 5, we use the Simpson's rule with a suitable step size h. Now, the problem (7) reduces to solve the following problem:

$$\min \sum_{\substack{i=1 \\ t_1+...+t_m=1 \\ t_i \geq 0}}^{m} g_j(t_1, \ldots, t_m). \tag{12}$$

This is a constrained optimization problem and can be solved by standard methods. We will solve it with `fmincon` function in Matlab in Section 5.

4. L^p center of mass of Toeplitz matrices

In this section we are going to define another barycenter for positive Toeplitz matrices.

4.1. Main result

It is known that, there is a Riemannian structure on the set $\mathbb{P}_n^+$ of all positive definite matrices given by the scalar product $g(M, N) = \mathrm{trace}(A^{-1}MA^{-1}N)$ on the tangent space to $\mathbb{P}_n^+$ at the given point A.
This structure makes $\mathbb{P}_n^+$ as a complete Riemannian manifold with negative curvature [4–6] in which the distance between two positive matrices A and B is given by

$$d(A, B) := \|\log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})\|_F, \tag{13}$$

where $\|.\|_F$ is the Frobenius norm $\|A\|_F := (\sum\limits_{i,j} |a_{ij}|^2)^{\frac{1}{2}}$.

Definition 4.1. *Assume that $1 < p < \infty$, and $A_1, \ldots, A_m \in \mathcal{T}_n^+$. Also, let S be the convex hull corresponding to $A_1, \ldots, A_m$, and $d(.,.)$ is the distance given* (13). *We define the Riemannian L^p barycenter of $\{A_1, \ldots, A_m\}$ as follows:*

$$B_p(A_1, \ldots, A_m) := \underset{X \in \mathcal{T}_n^+}{\mathrm{argmin}} \sum_{i=1}^{m} d^p(X, A_i) \tag{14}$$

Let $\mathcal{T}_n^+$ be the set of all Toeplitz positive definite matrices. As a closed subset of $\mathbb{P}_n^+$, the set $\mathcal{T}_n^+$ admits a smooth structure as a Riemannian manifold, where $\mathbb{P}_n^+$ is the set of all positive definite matrices of size n. The problem (14) has a unique solution $B \in \mathcal{T}_n^+$ and B belongs to S; see Theorem 3.4 of [1]. In the case $p = 2$, the geometric mean of m positive definite matrices $A_1, \ldots, A_m$ can be defined as their center of mass, that is the unique minimizer of [6]

$$f(x) = \sum_{i=1}^{m} d^2(X, A_i). \tag{15}$$

This mean is known as Karcher mean and extensively has been investigated; e.g., see D. Bini et al; see [6].
An important point to note is that if $A_1, \ldots, A_m$ are $n \times n$ positive definite Toeplitz matrices, the geometric mean or center of mass obtained by solving problem (4.3) may not necessarily be a Toeplitz matrix. Moreover, reference [12, Section 10.3] explores generalized locally Toeplitz structures, which include Toeplitz matrices as a special case. In that work, the symbol function is used to define the geometric mean. Specifically, Theorem 10.2 in [12] demonstrates that the sequence of geometric means $\{G(A_n, B_n)\}_n$ forms a generalized locally Toeplitz sequence, where the symbol is given by the geometric mean of the corresponding symbols. Additionally, Conjecture 10.1 examines a generalization of this theorem, which is extended to the multilevel block case in [2] and proved in great generality in [14]. Furthermore, references [18] and [19] have previously investigated the geometric mean of Toeplitz matrices in detail, particularly in relation to the properties of the ALM, including monotonicity, under a new definition. Unlike the aforementioned references, this paper focuses on the geometric mean in the context of the distance function (4.1) as one of its applications. It is evident that the resulting mean is not necessarily a Toeplitz matrix. However, since the argmin is chosen from the convex combination of Toeplitz matrices, the final result preserves the Toeplitz structure. Moreover, in our approach, since B is a linear combination of Toeplitz matrices, the solution obtained from solving problem (4.2) is guaranteed to be a Toeplitz matrix. But in our approach, as B is a linear combination Toeplitz matrices the solution that comes from solving the problem(14) is a Toeplitz matrix too.
Now that we know $B \in S$, we can consider $B = t_1A_1 + \cdots + t_mA_m$, where $t_1 + \cdots + t_m = 1$ and $t_i \geq 0$, for

$1 \leq i \leq m$.
We have

$$f(t_1, \ldots, t_m) = d^p(X, A_i) = \| \log(A_i^{-\frac{1}{2}} X A_i^{-\frac{1}{2}}) \|_F^p$$
$$= \| \log(\sum_{\substack{j=1 \\ j \neq i}}^{m} t_i A_i^{-\frac{1}{2}} A_j A_i^{-\frac{1}{2}}) + t_i I \|_F^p$$

Therefore, to solve the problem (14), we must solve

$$min \sum_{\substack{i=1 \\ t_1+\ldots+t_m=1 \\ t_i \geq 0}}^{m} f_i(t_1, \ldots, t_m). \tag{16}$$

The problem constrained optimization (16) can be solved via standard methods [8].

4.2. Geometric mean

Let $A_1, \ldots, A_m$ be $n \times n$ Toeplitz positive definite matrices. As we mentioned above, if $p = 2$, then the solution of the problem (16) is the geometric mean of $\{A_1, \ldots, A_m\}$. n order to solve the problem (16) in the case $p = 2$, we use the gradient descent method with a projection step. We know [17] that the derivation of $f(t_1, \ldots, t_m)$ is given by

$$\frac{\partial f_i}{\partial t_k} = 2\,\mathrm{tr}\left(\log\left(A_i^{-\frac{1}{2}} B A_i^{-\frac{1}{2}}\right)\left(A_i^{-\frac{1}{2}} B A_i^{-\frac{1}{2}}\right)^{-1} A_i^{-\frac{1}{2}} A_k A_i^{-\frac{1}{2}}\right)$$

Where $B = t_1 A_1 + \cdots + t_m A_m$.
Set $F(t_1, \ldots, t_m) = \sum_{i=1}^{m} f_i(t_1, \ldots, t_m)$ and $\mathcal{H} = \{(t_1, \ldots, t_m) \mid 1, + \cdots + t_m = 1, \quad t_1, \ldots, t_m \geq 0\}$.
We have

$$\nabla F(t_1, \ldots, t_m) = \begin{bmatrix} \sum_{i=1}^{m} \frac{\partial f_i}{\partial t_1} \\ \vdots \\ \sum_{i=1}^{m} \frac{\partial f_i}{\partial t_m} \end{bmatrix}.$$

Gradient descent method is a well- know algorithm for finding a local minimum of differentiable unctions [8]. In general, if the multi-variable function $F(x)$ is differentiable in a neighborhood f a point t, then $F(x)$ decreases fastest if one goes from t in the direction of negative gradient f F at t. It follows that if

$$t^{(n+1)} = t^{(n)} - \gamma_n \nabla F\left(t^{(n)}\right),$$

for step size $\gamma_n \geq 0$, then $F\left(t^{(n)}\right) \geq F\left(t^{(n+1)}\right)$.
Starting from a point $t^{(0)}$ for a local minimum F, we find a monotonic sequence $F\left(t^{(0)}\right) \geq F\left(t^{(1)}\right) \geq \cdots \geqslant 0$. Hence, as in any gradient descent algorithm the sequence $\{F(t^n)\}$ converges to the optimal value F^*. In other side, for every convergent subsequence $\left\{t^{(m_k)}\right\}$ of $\left\{t^{(m)}\right\}$, we have $F^* = \lim_{n\to\infty} F\left(t^{(m_k)}\right) = F\left(\lim_{k\to\infty} t^{(m_k)}\right)$.
Since, $t^{(m)} \subset B$ is bounded and F is strictly convex (so it does not admit more than one distinct optimizer point), we deduce that the sequence $t^{(m)}$ itself converge to the unique optimizer point t^*; i.e., $F(t^*) = F^*$.
For finding geometric mean we look after $t^* \in \mathcal{H}$. In this regard, we add a projection step to assure that the sequence $\{t_0\}$ remains in $\mathcal{H}$. We now have a two-step update rule, where we first subtract a gradient from the current value of t and then project the resulting vector onto $\mathcal{H}$ (see [20], p. 193, 194).

For computing Euclidean projection of a point $t = (t_1, \ldots, t_m) \in \mathbb{R}^m$ on probability simplex, which is defined via the following problem

$$\begin{aligned} &\min_{x\in\mathbb{R}^m} \frac{1}{2}|x-t|^2 \\ &s.t \quad x^T.1 = 1, \\ &\qquad\quad x \geq 0, \end{aligned} \tag{17}$$

we use the $O(D\log D)$ Algorithm 1 presented in [21].

Algorithm 1 Euclidean projection of a vector onto the probability simplex.

Input: t $\in \mathbb{R}^m$
Sort **t** into **u** : $u_1 \geq u_2 \geq \cdots \geq u_m$
Find $\rho = \max\left\{1 \leq j \leq m : u_j + \frac{1}{j}\left(1 - \sum_{i=1}^{j} u_i\right) > 0\right\}$
Define $\lambda = \frac{1}{\rho}\left(1 - \sum_{i=1}^{\rho} u_i\right)$
Output: x s.t. $x_i = \max\{t_i + \lambda, 0\}, i = 1, \ldots, m.$

Therefore we can compute geometric mean by the following algorithm.

Algorithm 2 Geometric mean.

Input: $A_1, \ldots, A_m, t^{(0)}$
repeat
$\Delta t^{(n)} := -\nabla F\left(t_1^{(n)}, \ldots, t_m^{(n)}\right)$
choose step size γ_n via exact or backtracking line search
$t^{\left(n+\frac{1}{2}\right)} = t^{(n)} + \gamma_n \Delta t^{(n)}$
$t^{(n)} = \operatorname{argmin}_{w\in\mathcal{H}} \frac{1}{2}\left\|w - t^{\left(n+\frac{1}{2}\right)}\right\|$
until stopping criterion is satisfied.
Output: B s.t. $B = t_1^* A_1 + \cdots + t_m^* A_m$.

The stopping criterion can be of the form $\left\|t^{(n+1)} - t^{(n)}\right\|_2 < \epsilon$ where ϵ is a positive small precision.

5. Applications and numerical experiments

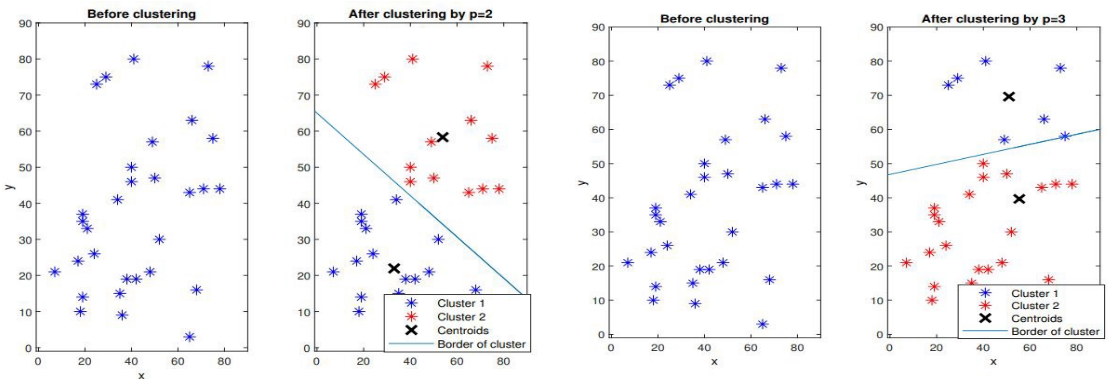
Many well-known machine learning algorithms assume that the data being classified is Euclidean in nature. However, to effectively operate within a Riemannian manifold and respect its unique geometric structure, machine learning algorithms must rely on operations that are specifically defined within these manifolds. In [11], it is explored how machine learning algorithms can be adapted to work within the context of Riemannian manifolds. While some Euclidean based algorithms rely solely on distance as a geometric tool to classify data, adapting these algorithms to work with non-Euclidean Riemannian manifolds simply requires replacing the Euclidean distance with the appropriate distance function of the manifold. One such algorithm [11] is a k-means clustering algorithm that utilizes the the Kähler metric for effective classification (Algorithm 3). Additionally, our approaches to barycenters, as introduced in Sections 3 and 4, can be used to efficiently cluster datasets of Toeplitz matrices.
For illustrative purposes only, we randomly construct 30 symmetric matrices belonging to $\mathcal{T}_2$. For computing the integral (4), we consider the 20 number of subdivisions of the interval $[0, 2\pi]$ to use the Simpson's rule for the numerical integration. Also, for finding the minimizer of (12), we use the `fmincon` function in MATLAB. In Figure 1 we see clustering with $p = 2$ and $p = 3$.

As described in Section 4.2, we can use this approach for computing the geometric mean of positive definite Toeplitz matrices. The cost of computing the geometric mean is based on the method described in this section is $O\left(pmn^4\right)$ arithmetic operations, where n is the size of the matrices, m is the number of

Algorithm k-means algorithm for N cluster

Initialization:
Pick randomly N points in the dataset. They now represent the barycenters of each class.
for $i = 1$ to loop number **do**
  Assign each point of the dataset to the closest barycenter.
  Compute the new barycenter of each class.
end for
return Each point is labeled according to the closest barycenter.

matrices and p is the number of iterations of this algorithm. In [7] other methods have been used to calculate the geometric mean of Toeplitz matrices, the structured geometric mean and the Kähler metric mean. The cost of these approach in term of arithmetic operations for m matrices is respectively $O\left(pn^4 + pmn^3\right)$ and $O\left(mn^4\right)$ arithmetic operations where n, m and p are the same as mentioned before. In [18], another method is introduced based on symbolic function (1), for this method the number of operations equals $O\left(kmn^2\right)$ where n is the size of the matrices, m is the number of matrices and k is the number of subdivision of the interval $[0, 2\pi]$ in order to use Simpson's rule for numerical integration. We construct randomly three positive definite matrices 3×3 belong to $\mathcal{T}_3$. we have assumed that $p = 2$. Fig.2(a) shows these matrices and their means with our method, structured geometric mean (SGM), and the Kähler metric mean (KMM) and symbol based geometric mean (SBGM) in a 3D diagram, note that each 3×3 Toeplitz matrix characterized by its first row, so corresponds to a point of $\mathbb{R}^3$. In order to compare the cost of different methods we let $n = 30, m = 3$ to 10, $p = 1$ and for SBGM method $k = 32$. The results are given in Fig.2(b).

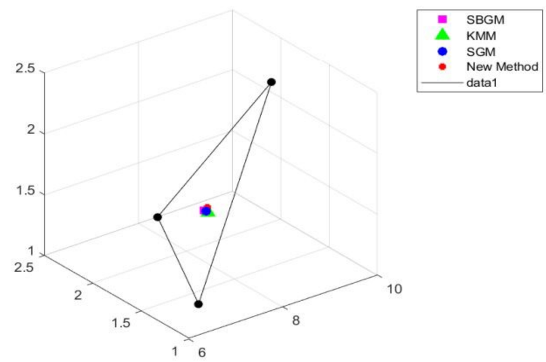


(a) $p = 2$ (b) $p = 3$

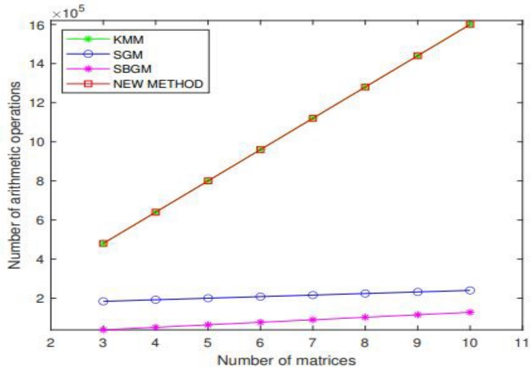
Figure 1: Clustering of 30 2×2 random symmetric Toeplitz matrices

6. Conclusion

In conclusion, our paper has presented two novel concepts of barycenter for Toeplitz matrices and demonstrated their applicability in clustering datasets, including Toeplitz matrices. We have illustrated the use of these barycenters in k-means clustering and also shown that the L^p barycenter can be utilized to calculate the geometric mean of Toeplitz matrices, which itself is a Toeplitz matrix. Our findings have been compared with existing methods in the literature, highlighting the effectiveness of our proposed approach. Overall, this work contributes to the advancement of clustering techniques for Toeplitz matrices and provides a valuable tool for analyzing and interpreting complex datasets.



(a) Representation of the four means for three randomly matrices



(b) Required arithmetic operations for different amounts of matrices

Figure 2: Comparison of the different methods

Acknowledgment

The authors express their gratitude to Yann Cabanes for providing valuable insights on Toeplitz datasets and for sharing his thesis, which proved to be immensely helpful in enhancing the quality of this work.

References

[1] B. Afsari, *Riemannian L^p center of mass: existence, uniqueness, and convexity*, Proc. Amer. Math. Soc., **139** (2011), no. 2, 655–673.
[2] D. Ahmad, M.F. Khan, S. Serra-Capizzano, *Matrix-sequences of geometric means in the case of hidden (asymptotic) structures*, Math., **13** (2025), no. 3, paper 393.
[3] T. Ando, C.-K. Li, R. Mathias, *Geometric means*, Linear Algebra Appl., **385** (2004), 305–334.
[4] R. Bhatia, *Positive Definite Matrices*, Princeton Series in Applied Mathematics, Princeton University Press, 2007.
[5] R. Bhatia and J. Holbrook, *Riemannian geometry and matrix geometric means*, Linear Algebra Appl., **413** (2006), 549–618.
[6] D. A. Bini, B. Iannazzo, *Computing the Karcher mean of symmetric positive definite matrices*, Linear Algebra Appl., **438** (2013), no. 4, 1700–1710.
[7] D. A. Bini, B. Iannazzo, B. Jeuris, R. Vandebril, *Geometric mean of structured matrices*, BIT Numer. Math., **54** (2014), no. 1, 55–83.
[8] S. Boyd, L. Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, 2004.
[9] A. Böttcher, B. Silbermann, *Analysis of Toeplitz operator*, Second edition, Prepared jointly with Alexie Karlovich, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2006.
[10] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext. Springer, New York, 2011.
[11] Y. Cabanes, F. Barbaresco, M. Arnaudon, J. Bigot, *Toeplitz Hermitian positive definite matrix machine learning based on Fisher metric*, Geometric science of information, Lecture Notes in Comput. Sci., 11712, Springer, Cham, (2019), 261–270.
[12] C. Garoni, S. Serra-Capizzano,*Generalized locally Toeplitz sequences: theory and applications*, Vol. I, Springer, Cham, 2017.
[13] U. Grenander, G. Szegö, *Toeplitz forms and their applications*, Univ. of California Press, Berkeley, 1985.
[14] A. Ilyas, M.F. Khan, V. Loi, S. Serra-Capizzano, *Geometric means of HPD GLT matrix-sequences: a maximal result beyond invertibility assumptions on the GLT symbols*, (2025) https://www.arxiv.org/abs/2505.03256.
[15] B. Jeuris, R. Vandebril, B. Vandereycken. *A survey and comparison of contemporary algorithms for computing the matrix geometric mean*, Electron. Trans. Numer. Anal., **39** (2012), 379–402.
[16] J. Lapuyade-Lahorgue, F. Barbaresco, *Radar detection using Siegel distance between autoregressive processes, application to HF and X-band radar*, In Radar Conference, 2008. RADAR '08, IEEE, Rome, May 2008.
[17] M. Moakher, *A differential geometric approach to the geometric mean of symmetric positive definite matrices*, SIAM J. Matrix Anal. Appl. **26** (2005), no. 3, 735–747.
[18] E. Nobari, *A monotone geometric mean for a class of Toeplitz matrices*, Linear Algebra Appl., **511**(2016), 1–18.
[19] E. Nobari, B. Ahmadi Kakavandi, *A geometric mean for Toeplitz and Toeplitz-block block-Toeplitz matrices*, Linear Algebra Appl., **548** (2018), 189–202.
[20] S. Shalev-Shwartz, S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*, Cambridge University Press, 2014.
[21] W. Wang, M. A. Carreira-Perpiñá, *Projection onto the probability simplex: An efficient algorithm with a simple proof, and an application*, preprint arXiv:1309.1541, 2013.