1 Hypersphere means

A basic operation in Euclidean space is the act of taking the mean of a set of points. For example, k-means, which is a fundamental clustering algorithm, depends on this operation. However, this simple averaging procedure must be adapted to work on manifolds. To see intuitively why this is true, imagine two vectors on the hypersphere. The straight Euclidean mean of the two vectors actually ends up inside the sphere, not on the surface! Taking the Euclidean mean almost certainly results in a vector not on the manifold, so we require some tools from differential geometry to find means on manifolds.

1.1 Karcher Mean

The Karcher mean is a geometric mean of various matrices and an extension of the well known geometric mean of two matrices. It is also referred to as the Riemannian geometric mean. In our investigation, we used it as a way to compute the average distance between two points on a hypersphere while obeying the curvature of the manifold. The process is stated in Algorithm 1. In order to do this we must incorporate a Logarithm map and Exponential map on the manifold.

The Logarithm map is defined as a function that takes a point, ρ_2, on the manifold and maps its projection vector, γ, on the tangent plane at an origin point, T_{p_1}.

$$\rho_2 = Exp_{\rho_1}(\gamma) = cos(|\gamma|)\rho_1 + sin(|\gamma|)\frac{\gamma}{|\gamma|} \tag{1}$$

On the first iteration, ρ_1 is the origin point of the tangent space.

The Exponential map can be thought of as the inverse function of the Logarithm map. It is a function that takes in a vector on the tangent plane at origin point and maps it on the hypersphere. As the difference between the updating vectors on the tangent space begin to show little change, it is thought as the optimal mean between the two points on the manifold.

$$\gamma = Log_{\rho_1}(\rho_2) = \tilde{\rho}\frac{cos^{-1}(\langle\rho_1, \rho_2\rangle)}{\sqrt{\langle\tilde{\rho}, \tilde{\rho}\rangle}} \tag{2}$$

where $\tilde{\rho} = \rho_2 - \langle\rho_2, \rho_1\rangle\rho_1$. Karcher mean, as we have mentioned before, is calculated in a way to obey the curvature of the hypersphere.

1.2 Spherical mean

Unlike the Karcher mean, the spherical mean does not go along the curvature of the hypersphere. In the assignment step of this algorithm, where we adjust the cluster centers, we find the mean by summing up all the vectors in a cluster and normalizing it back onto the manifold.

$$x_l = \frac{\sum_{i \in X_l} \rho_i}{||\sum_{i \in X_l} \rho_i||}, l = 1 \dots K$$

where X_h are the center of each K number of clusters. This is the normalized gravity center of the new cluster.

1.3 Practical comparison

To find the best approach in calculating a cluster center, we put Karcher mean and Spherical K-mean to the test. We tested them in five types of situations in where the angles are either identical, orthogonal, opposite, two random angles, or multiple random angles as shown in Figure 1.

We found that Karcher mean and Spherical K-means produced the same results as long as the step size for Karcher mean was set to one. But both approaches failed in the case where the angles were opposite.

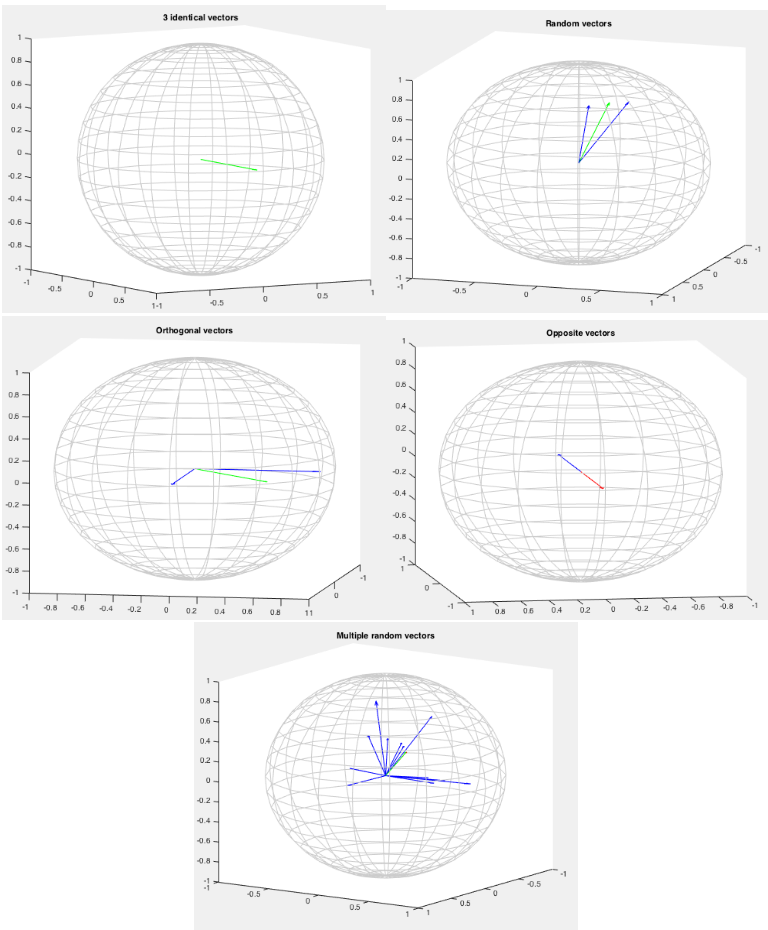

Figure 1: Different test cases comparing accuracy of Karcher mean and Spherical K-means.