Dictionary Learning on Rotations

Tian Cao*

Department of Computer Science
UNC-Chapel Hill

1 Definition of Group

1.1 Group

A group is a set, G, together with an operation $\cdot$ that combines any two elements a and b to form another element, denoted $a \cdot b$ or ab. The qualified group $(G, \cdot)$ must satisfy the following requirements [4],

- **Closure**: for all a, b in G, the result of the operation, $a \cdot b$, is also in G.
- **Associativity**: for all a, b and c in G, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- **Identity element**: there exists an identity element e in G, such that for every element a in G, $e \cdot a = a \cdot e$. Such an element is unique.
- **Inverse element**: for each a in G, there exists an element b in G such that $a \cdot b = b \cdot a = e$, where e is the identity element.

1.2 Group Action

A group action of G on set X is a function

$$G \times X \to X, \ \ (g, x) \mapsto g \cdot x$$

that satisfies the following [4]

- **Compatibility**: $(gh) \cdot x = g \cdot (h \cdot x)$ for all g, h in G and all x in X.(Here, gh denotes the result of applying the group operation of G to the elements g and h.)
- **Identity**: $e \cdot x = x$ for all x in X.

*tiancao@cs.unc.edu

1.3 Lie Group

A group G is a Lie group that is also a finite-dimensional smooth manifold, and in which the group operations of multiplication and inversion are smooth maps [1], i.e. the maps $(g, h) \mapsto gh$ and $g \mapsto g^{-1}$ are smooth.

1.4 Lie Algebra

A Lie algebra is a vector space $\mathfrak{g}$ over some field F together with a binary operation $[\cdot, \cdot] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ called the Lie bracket, which satisfies the following requirement [1],

- **Bilinearity**: $[ax + by, z] = a[x, z] + b[y, z]$, $[z, az + by] = a[z, x] + b[z, y]$, for all scalars a, b in F and all elements x, y, z in $\mathfrak{g}$.
- **Alternating on $\mathfrak{g}$**: $[x, x] = 0$.
- **The Jacobi identity**: $[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$, for all x, y, z in $\mathfrak{g}$.

A Lie algebra is the tangent space at the identity element of a Lie group. A tangent space is a real vector space can be attached to every point x of a differentiable manifold that intuitively contains the possible "direction" at which one can tangentially pass through x. more formal definition of tangent space can be referred to [5, 1].

1.5 Exponential Map and Log Map

The exponential map is a map from the Lie algebra $\mathfrak{g}$ to the corresponding Lie group G, i.e. $\exp : \mathfrak{g} \to G$. If G is a matrix Lie group, the element of $\mathfrak{g}$ and G are squared matrices and the exponential map is given by the matrix exponential [5, 1],

$$\exp(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!} = I + X + \frac{1}{2}X^2 + \dots.$$

The log map is the inverse map $\log = \exp^{-1}$, which maps from the Lie group G into the Lie algebra g.

2 Rotation Groups

2.1 Rotation Group in three dimensions (SO(3))

The 3D rotation group is the group of all rotations about the origin of three-dimensional Euclidean space $\mathbf{R}^3$ under the operation of composition [3]. The composition of rotations corresponds to matrix multiplication. A rotation about the origin is a transformation that preserves the origin, Euclidean distance and orientation [5].

The rotation group satisfies: a. Composing two rotations results in another rotation; b. every rotation has a unique inverse rotation; c. the identity map satisfies the definition of a rotation; d. rotations obey the associative property. So the set of all rotations is a group under composition. Moreover, the rotation group has a natural manifold structure for which the group operations are smooth, thus it is in fact a Lie group. The rotation group is often denoted SO(3), and $\mathrm{SO}(3) = \{A \in \mathbb{R}^{3\times3} | A^T A = I, \det(A) = 1\}$.

Since SO(3) is a Lie group, its Lie algebra, denoted $\mathfrak{so}_3$, consists of all 3×3 skew-symmetric matrices [5].

3 Dictionary Learning for Rotations

Given a set of base points $P = \{p_i\}, p_i \in \mathbb{R}^3, i = 1, \ldots, n$ on a unit sphere, and a set of random rotations $R = \{r_j\}, j = 1, \ldots, m$, we have the corresponding rotated base points $R(P) = \{r_j(p_i)\}, j = 1, \ldots, m, i = 1, \ldots, n$. The standard dictionary learning method which trying to learning a dictionary to reconstruct the coordinate differences of the corresponding points solves

$$\{\hat{D}, \hat{\alpha}\} = \underset{D,\alpha}{\operatorname{argmin}} \frac{1}{2} \|I - D\alpha\|_F^2 + \lambda\Omega(\alpha) \tag{1}$$

where $I = R(P) - P, I \in \mathbb{R}^{3n\times m}$ is the coordinate differences of the training points, $\Omega(.)$ is the regularization on the coefficient α and the commonly used regularization is $\ell1$ norm.

For a given test rotated points $r_k(p_i), i = 1, \ldots, n$, we can reconstruct its coordinate with $D\alpha_k + P$, which is a linear combination of dictionary elements based on the coefficient α.

The 3D rotation group SO(3) act on the points will rotate the points on a manifold of sphere, however, the linear combinations of the coordinates in euclidean space cannot guarantee that the reconstructed points are on the same manifold. In order to keep consistency, we proposed a dictionary learning method for rotations which incorporate the parameters of the Lie algebra and reconstruct the test points with exponential map of the Lie algebra with the rotation parameters.

From section 2., we know that the Lie algebra of SO(3) consists of all skew-symmetric matrices. Thus we only need three parameters for example (a, b, c) to generate a skew-symmetric matrix as

$$M = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}.$$

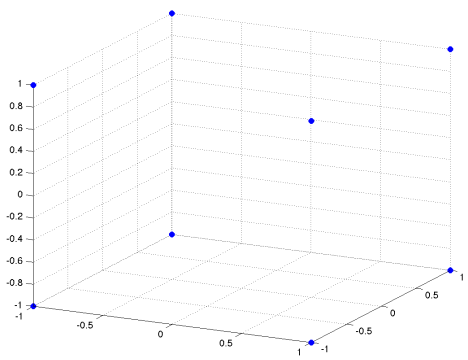
Thus the 3D rotation of point x is $\exp(Mx)$.

The proposed dictionary learning method is as follows,

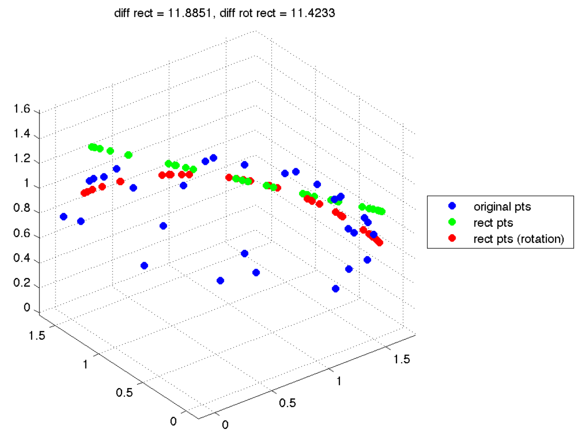
$$\{\hat{\tilde{D}}, \hat{\alpha}\} = \underset{D,\alpha}{\operatorname{argmin}} \frac{1}{2} \|\tilde{I} - \tilde{D}\alpha\|_F^2 + \lambda\Omega(\alpha) \tag{2}$$

where $\tilde{I}$ is the I in (1) combined with the rotation parameters for , thus $\tilde{I} \in \mathbb{R}^{(3n+3)\times m}$, $\tilde{D} = [D_1^T, D_2^T]^T, D_1 \in \mathbb{R}^{3n\times m}, D_2 \in \mathbb{R}^{3\times m}$ where D_1 is the part of dictionary corresponding to coordinate difference while D_2 is the part of dictionary corresponding to rotation parameters.

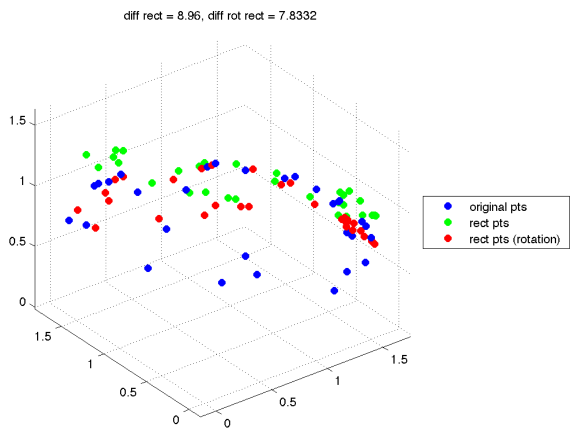
For a given test rotated points $r_k(p_i), i = 1, \ldots, n$, we can first get α_k by coding the test points with learned dictionary part D_1, then we reconstruct the rotation parameters with $D_2\alpha_k$, and the reconstructed rotated points are $\exp(D_2\alpha_k p_i), i = 1, \ldots, n$. Fig. 1 shows the result of point reconstruction using our proposed method.



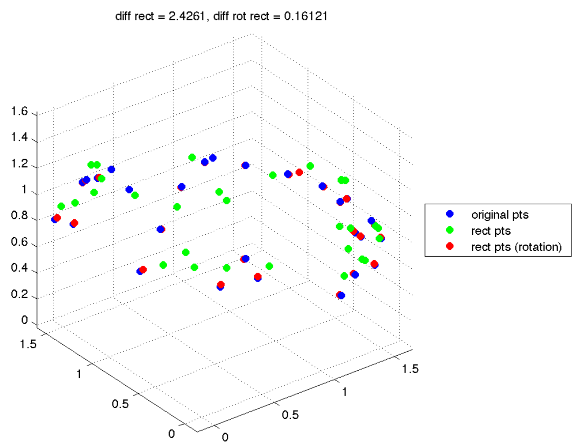
(a) 8 base points in dictionary learning



(b) points reconstruction with 1 dictioanry element



(c) points reconstruction with 2 dictioanry element



(d) points reconstruction with 5 dictioanry element

Figure 1: Examples of point reconstruction using dictionary learned on rotations. a) shows the 8 base points for 3D rotations, the training data is generated by randomly rotate the 8 points on a sphere; b) is the point reconstruction using 1 dictionary element; c) is the point reconstruction using 2 dictionary elements; d) is the point reconstruction using 5 dictionary elements. In b), c) and d), the blue dots represent one randomly rotated base point, green dots means the reconstructed point using D_1 of $\tilde{D}$, the red dot represent the reconstructed points using D_2 of $\tilde{D}$. We can see that the red dots are always on the manifold of sphere while the green dots are not.

4 Connection to Dictionary Learning on Manifold

In [2], the author proposed a dictionary learning method to generalize the standard dictionary learning from Euclidean space to Riemannian manifold which replace the data term $\|I - D\alpha\|_F^2 = \sum_{j=1}^n \|i_j - D\alpha_j\|_2^2$ as $\sum_{j=1}^n \| \sum_{k=1}^m \alpha_{ij} \log_{i_j}(d_k)\|_{x_j}^2$. (this data term defines the distance on manifold instead of on euclidean space, based on my understanding).

Our method does not learning dictionary directly on the manifold however, we learning the dictioanry with some parameters on the manifold and reconstruct the data using the learned parameters.

References

[1] Brian Hall. *Lie groups, Lie algebras, and representations: an elementary introduction*, volume 222. Springer, 2003.

[2] Jeffrey Ho, Yuchen Xie, and Baba Vemuri. On a nonlinear generalization of sparse coding and dictionary learning. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, pages 1480–1488, 2013.

[3] AW Joshi. *Elements of group theory for physicists*. New Age International, 1997.

[4] James S. Milne. Group theory (v3.13), 2013. Available at www.jmilne.org/math/.

[5] Wikipedia. Rotation group so(3) — wikipedia, the free encyclopedia, 2014. [Online; accessed 11-February-2014].