

Robust Quaternion Estimation with Geometric Algebra

Mauricio Cele Lopez Belon

Abstract. Robust methods for finding the best rotation aligning two sets of corresponding vectors are formulated in the linear algebra framework, using tools like the SVD for polar decomposition or QR for finding eigenvectors. Those are well established numerical algorithms which on the other hand are iterative and computationally expensive. Recently, closed form solutions has been proposed in the quaternion's framework, those methods are fast but they have singularities i.e., they completely fail on certain input data. In this paper we propose a robust attitude estimator based on a formulation of the problem in Geometric Algebra. We find the optimal eigenquaternion in closed form with high accuracy and with competitive performance respect to the fastest methods reported in literature.

Mathematics Subject Classification (2010). Parallel algorithms 68W10; Clifford algebras, spinors 15A66.

Keywords. Geometric Algebra, Quaternion Estimation, Wahba Problem.

1. Introduction

The estimation of rotations has been studied for over half a century [15]. The problem consist on finding the optimal rotation aligning two sets of corresponding vectors. Many effective methods have been developed [1, 8, 12, 14, 20] using 3×3 matrices and quaternions. Formulations based on quaternions solves a max-eigenvalue problem, while formulations based on linear algebra relies on Singular Value Decomposition (SVD). In the last decade formulations based on geometric algebra [4, 13] were introduced but they also rely on linear algebra numerical algorithms such as SVD due to the lack of native numerical algorithms. Closed form solutions for finding the optimal quaternion has been proposed [16, 17, 19, 21] based on analytic formulas for solving the roots of the quartic polynomial associated with eigenvalue problem.

Accuracy and speed of prominent methods have been compared in [5, 11, 19] evidencing a trade-off between performance and robustness. In particular SVD based methods exhibit the best accuracy but low performance and quaternion based methods are faster but less accurate. Regarding the later methods, the closed form solutions exhibit the best performance so far but they have singularities i.e., they completely fail on certain input data.

In this paper we propose a robust estimator of the best quaternion aligning two sets of corresponding vectors. We maximize a convex quadratic energy functional formulated in the $\mathbb{G}_{3,0,0}$ geometric algebra which allow us to find an optimal quaternion in a robust way without resorting to linear algebra numerical algorithms. Geometric algebra rotors are isomorphic to quaternions, we find geometric algebra to be a more natural choice for studying this problem since rotations and subspaces of \mathbb{R}^3 are treated in the same manner, facilitating meaningful algebraic manipulations. We primarily work with bivectors instead of vectors for the sake of mathematical convenience. Due to mathematical (and geometric) duality of vectors and bivectors in \mathbb{G}_3 our formulation is also valid for vectors.

2. Geometric Algebra \mathbb{G}_3

A geometric algebra \mathbb{G}_3 is constructed over a real vector space \mathbb{R}^3 , with basis vectors $\{e_1, e_2, e_3\}$. The associative geometric product is defined on vectors so that the square of any vector a is a scalar $aa = a^2 \in \mathbb{R}$ and the geometric product of two vectors a and b is $ab = a \cdot b + a \wedge b$ and $ba = b \cdot a - a \wedge b$. From the vector space \mathbb{R}^3 , the geometric product generates the geometric algebra \mathbb{G}_3 with elements $\{X, R, A, \dots\}$ called multivectors.

For a pair of vectors a and b , a symmetric inner product $a \cdot b = b \cdot a$ and antisymmetric outer product $a \wedge b = -b \wedge a$ can be defined implicitly by the geometric product. It is easy to prove that $a \cdot b = \frac{1}{2}(ab + ba)$ is a scalar, while the quantity $a \wedge b = \frac{1}{2}(ab - ba)$, called a bivector or 2-vector, is a new algebraic entity that can be visualized as the two-dimensional analogue of a direction i.e., a planar direction. Similar to vectors, bivectors can be decomposed in a bivector basis $\{e_{12}, e_{13}, e_{23}\}$ where $e_{ij} = e_i \wedge e_j$.

The outer product of three vectors $a \wedge b \wedge c$ generates a 3-vector also known as the pseudoscalar, because the trivector basis consist of single element $e_{123} = e_1 \wedge e_2 \wedge e_3$. Similarly, the scalars are regarded as 0-vectors whose basis is the number 1. It follows that the outer product of k -vectors is the completely antisymmetric part of their geometric product: $a_1 \wedge a_2 \wedge \dots \wedge a_k = \langle a_1 a_2 \dots a_k \rangle_k$ where the angle bracket means k -vector part, and k is its grade. The term grade is used to refer to the number of vectors in any exterior product. This product vanishes if and only if the vectors are linearly dependent. Consequently, the maximal grade for nonzero k -vectors is 3. It follows that every multivector X can be expanded into its k -vector parts and the entire

algebra can be decomposed into k -vector subspaces:

$$\mathbb{G}_3 = \sum_{k=0}^n \mathbb{G}_3^k = \{X = \sum_{k=0}^n \langle X \rangle_k\}$$

This is called a *grading* of the algebra.

Reversing the order of multiplication is called reversion, as expressed by $(a_1 a_2 \dots a_k)^\sim = a_k \dots a_2 a_1$ and $(a_1 \wedge a_2 \wedge \dots \wedge a_k)^\sim = a_k \wedge \dots \wedge a_2 \wedge a_1$, and the reverse of an arbitrary multivector is defined by $\tilde{X} = \sum_{k=0}^n \langle \tilde{X} \rangle_k$.

2.1. Rotors

Rotations are even grade multivectors known as rotors. We denote the subalgebra of rotors as \mathbb{G}_3^+ . A rotor R can be generated as the geometric product of an even number of vectors. A reflection of any k -vector X in a plane with normal n is expressed as the sandwich product $(-1)^k n X n$. The most basic rotor R is defined as the product of two unit vectors a and b with angle of $\frac{\theta}{2}$. The rotation plane is the bivector $B = \frac{a \wedge b}{\|a \wedge b\|}$.

$$ab = a \cdot b + a \wedge b = \cos\left(\frac{\theta}{2}\right) + B \sin\left(\frac{\theta}{2}\right). \quad (2.1)$$

Rotors act on all k -vectors using the sandwich product $X' = R X \tilde{R}$, where \tilde{R} is the reverse of R and can be obtained by reversing the order of all the products of vectors.

2.2. Bivector products

We define the commutator product of two bivectors p_j and q_j as $p_j \times q_j = \frac{1}{2}(p_j q_j - q_j p_j)$. The commutator product of bivectors in \mathbb{G}_3 can be interpreted as a cross-product of bivectors i.e., the resulting bivector $B = p_j \times q_j$ is orthogonal to both p_j and q_j . The commutator product allow us to define the geometric product of two bivectors as $AB = A \cdot B + A \times B$. The inner product of bivectors differs from the inner product of vectors on the sign, since the square of bivectors is negative, the inner product of bivectors is a negative scalar e.g., $(ae_{12} + be_{13} + ce_{23}) \cdot (de_{12} + ee_{13} + fe_{23}) = -ad - be - cf$.

2.3. Quaternions

A quaternion $Q = w + \vec{v}$ consists of two parts: a scalar part w and vector part \vec{v} which denotes the axis of rotation. The vector part \vec{v} is defined in a basis of complex vectors $\{i, j, k\}$ that squares to -1 and anticommute i.e., $i^2 = j^2 = k^2 = -1$ and $ijk = -1$. In geometric algebra they corresponds to bivectors:

$$\begin{aligned} i &= e_{23} & j &= e_{13} & k &= e_{12} \\ ijk &= e_{23}e_{13}e_{12} = -1 \end{aligned} \quad (2.2)$$

Rotors can easily be transformed to quaternions and vice versa. A rotor $R = w + B$ corresponds with a quaternion $Q = w + \vec{v}$, where $B = \alpha e_{12} + \beta e_{13} + \gamma e_{23}$ and $\vec{v} = \gamma i + \beta j + \alpha k$.

3. Geometric Algebra Rotor Estimation

Given two sets of n corresponding bivectors $P = \{p_j\}_{j=1}^n$ and $Q = \{q_j\}_{j=1}^n$, we attempt to maximize the following energy function:

$$E(R) = \max_{R \in \mathbb{G}_3^+} \sum_j c_j \|p_j + \tilde{R}q_j R\|^2 \quad (3.1)$$

s.t. $R\tilde{R} = 1$

where $\{c_j\}_{j=1}^n$ are scalar weights such that $\sum_j c_j = 1$. It is a quadratic maximization problem with a non-linear constraint in R . Notice that the term $\|p_j + \tilde{R}q_j R\|^2$ is dual to the traditional least squares error $\|q_j - Rp_j \tilde{R}\|^2$. Duality is stated in the sense that the optimal R is a critical point of both energies. Notice also that $p_j + \tilde{R}q_j R$ is equivalent to $Rp_j + q_j R$ i.e., by multiplying by R on the left and using the fact that $R\tilde{R} = 1$. The equivalent problem is:

$$E(R) = \max_{R \in \mathbb{G}_3^+} \sum_j c_j \|Rp_j + q_j R\|^2 \quad (3.2)$$

s.t. $R\tilde{R} = 1$

which exposes the that R is only quadratic in 3.2.

We can take the multivector derivative of the energy $E(R)$ with respect to the rotor R to find the critical multivector of it i.e., solve for $\partial E(\tilde{R}) = 0$ using e.g., Newton's method (provided the energy is convex). The multivector derivative is an extrinsic operator which produces gradient multivectors outside the energy's tangent bundle. That implies the critical multivector R of such gradient might lie outside the rotor manifold. In general it would be needed to either work directly on the Lie algebra of $\text{SO}(3)$ or add a Lagrange multiplier term $L(R) = \lambda(R\tilde{R} - 1)$ to $E(R)$ to constrain critical multivector to $\text{SO}(3)$. However we can avoid that inconvenience in \mathbb{G}_3 since a rotor can be orthogonally projected to manifold by simply normalization. In other words, once the critical multivector is found it can be projected to the manifold by normalizing it.

Note that since p_j and q_j are bivectors then Rp_j and $q_j R$ are also bivectors. Since the square of a bivector X is a scalar i.e., $\|X\|^2 = X\tilde{X}$, we have:

$$\begin{aligned} \|Rp_j + q_j R\|^2 &= (Rp_j + q_j R)(\tilde{p}_j \tilde{R} + \tilde{R} \tilde{q}_j) \\ &= (Rp_j + q_j R)(-p_j \tilde{R} - \tilde{R} q_j) \\ &= -\langle Rp_j p_j \tilde{R} \rangle - \langle Rp_j \tilde{R} q_j \rangle - \langle q_j R p_j \tilde{R} \rangle - \langle q_j R \tilde{R} q_j \rangle \end{aligned}$$

Where we recognize that everything inside $\langle \rangle$ is a scalar. So the multivector derivative of a scalar-valued function gives a gradient function (equivalent to a Jacobian matrix in vector calculus). Also note that $\langle \rangle$ acts in a similar way to the trace operator in matrix algebra. So it can be shown that any cyclic permutation of the bivector products inside the $\langle \rangle$ produces the same

scalar. That's a useful fact to take into account when taking the multivector derivative. For example:

$$\begin{aligned}
-\partial\langle Rp_j p_j \tilde{R} \rangle &= -\partial\langle \dot{R} * p_j p_j \tilde{R} \rangle - \partial\langle \dot{\tilde{R}} * Rp_j p_j \rangle \\
&= -\langle \partial R * p_j p_j \tilde{R} \rangle - \langle \partial \tilde{R} * Rp_j p_j \rangle \\
&= \mathcal{P} \left[-p_j p_j \tilde{R} - \widetilde{Rp_j p_j} \right] \\
&= -p_j p_j \tilde{R} - p_j p_j \tilde{R} \\
&= -2p_j p_j \tilde{R}
\end{aligned} \tag{3.3}$$

where $\mathcal{P}[\]$ is a grade projection to even grades (i.e., grades 0 and 2) of the resulting multivector which, in this case, is simply the identity function.

Doing something similar to all other terms we get:

$$\partial E_j(\dot{R}) = -2p_j p_j \tilde{R} - 4p_j \tilde{R} q_j - 2\tilde{R} q_j q_j \tag{3.4}$$

taking the reverse to express it in terms of R (recall that $\tilde{X} = -X$, $\widetilde{XY} = \tilde{Y}\tilde{X}$ and $\tilde{\tilde{X}} = X$) we get $\partial E_j(\dot{R}) = -Rpp - 2qRp - qqR$ from which follows the eigenrotor equation for $\partial E(\dot{R}) = 0$:

$$\lambda R = 2 \sum_j c_j q_j R p_j \tag{3.5}$$

where $\lambda = -\sum_j c_j (p_j^2 + q_j^2)$ which shows that the eigenrotor R associated with the eigenvalue λ is the optimal one. However this equation only holds for exact matches, i.e., when the input bivectors can be aligned by a rotor. In general that is not the case since some arbitrary noise is expected on input bivectors therefore we are looking for the best alignment in the least squares sense. The eigenrotor equation can be stated in matrix algebra as:

$$(A + \lambda I)R = 0 \tag{3.6}$$

where A is a matrix derived from $-2\sum_j c_j q_j R p_j$ and $\lambda = -\sum_j c_j (p_j^2 + q_j^2)$. Note that λ is a positive number since the square of a bivector is a negative number and $c_j \geq 0$. Let's define the 4×4 symmetric matrix $M = A + \lambda I$ where M is singular when the input bivectors can be aligned by a rotor and otherwise is positive definite i.e., M is a positive semi-definite matrix.

The matrix $A = \sum_j A_j$ can be derived from $-2c_j q_j R p_j$ as follows. First split R as $w+B$ where w is a scalar and B is a bivector, second split geometric

product of bivectors on inner and commutator products

$$\begin{aligned}
 -2c_j q_j (w + B)p_j &= -2c_j (wq_j p_j + q_j Bp_j) \\
 &= -2c_j (w(q_j \cdot p_j) + w(q_j \times p_j) + \\
 &\quad (q_j \cdot B)p_j + B \cdot (p_j \times q_j) - B(p_j \cdot q_j) + q_j(B \cdot p_j))
 \end{aligned} \tag{3.7}$$

the last rearrange uses the identity $qBp = (q \cdot B)p + B \cdot (p \times q) - B(p \cdot q) + q(B \cdot p)$. Now that everything is in terms of inner products and commutators we can rewrite that equation in matrix form using the correspondence $X \cdot Y = -\vec{X}^T \vec{Y}$ for bivectors X, Y and 3×1 column vectors \vec{X} and \vec{Y} . We also use the fact that bivector's commutator corresponds to vector's cross-product $X \times Y = \vec{X} \times \vec{Y}$:

$$A_j = 2c_j \begin{bmatrix} q_j^T p_j & (p_j \times q_j)^T \\ p_j \times q_j & q_j p_j^T + p_j q_j^T - (p_j^T q_j)I \end{bmatrix} \tag{3.8}$$

$$\tag{3.9}$$

To proof the positive semi-definiteness of $M = \sum_j A_j + \lambda I$ we proceed to express $E(R)$ in matrix algebra as the quadratic form $E(R) = R^T M R$ as follows: the constraint $Rp_j + q_j R$ can be rewritten as $(w + B)p_j + q_j(w + B)$. Expanding the geometric product of bivectors in terms of the inner product and the commutator product we get:

$$w(p_j + q_j) + B \cdot (p_j + q_j) + (q_j - p_j) \times B \tag{3.10}$$

In matrix language we can define the following matrix system $M_j R = 0$:

$$\begin{aligned}
 M_j R &= \begin{bmatrix} 0 & -s_j^T \\ s_j & [d_j]_{\times} \end{bmatrix} \begin{bmatrix} w \\ B \end{bmatrix} = \begin{bmatrix} -s_j^T B \\ ws_j + d_j \times B \end{bmatrix} \\
 &\quad d_j = q_j - p_j \quad s_j = p_j + q_j
 \end{aligned} \tag{3.11}$$

where d_j and s_j are 3×1 column vectors holding bivector's coefficients, M_j is a skew-symmetric 4×4 real matrix, so that $M_j^T = -M_j$. The rotor R is represented as 4×1 column vector made of the scalar w and the 3×1 column vector B holding the bivector's components. The 3×3 matrix $[d_j]_{\times}$ is representing the skew-symmetric cross-product matrix as usually defined for vectors in \mathbb{R}^3 .

Therefore we can express $E(R)$ as the quadratic form:

$$E(R) = R^T M R$$

where $M = \sum_j^n c_j M_j^T M_j$. Note that since M_j is skew-symmetric, the product $M_j^T M_j$ is symmetric and positive semi-definite. Consequently the matrix M is also symmetric positive semi-definite. It follows that all eigenvalues of M are real and $\lambda_i \geq 0$.

$$\begin{aligned}
 M_j^T M_j &= \begin{bmatrix} \|s_j\|^2 & (s_j \times d_j)^T \\ s_j \times d_j & s_j s_j^T - [d_j]_{\times}^2 \end{bmatrix} \\
 &\quad d_j = q_j - p_j \quad s_j = p_j + q_j
 \end{aligned} \tag{3.12}$$

By the spectral theorem the maximizer of $E(R)$ is the eigenvector of M associated with the largest eigenvalue which is non-negative number.

4. Convexity

The multivector derivative gives the gradient of the energy $\partial E(\dot{R})$, which is equal to zero at critical points and gives the eigenrotor equation $\lambda R = 2 \sum_j c_j q_j R p_j$ which implies the matrix equation $(A + \lambda I)R = 0$. From the eigenrotor equation follows that $M = A + \lambda I$ is singular when input bivectors can be aligned by a rotor R associated with eigenvalue λ . Otherwise, e.g., in presence of noisy input bivectors, the matrix M is positive definite. That means the energy $E(R)$ is convex.

The convexity of the energy $E(R)$ can be proved by showing that its Hessian matrix of second partial derivatives is positive semi-definite. The Hessian matrix of $E(R)$ is $\frac{\partial^2 E(R)}{\partial R^2} = \sum_j^n c_j M_j^T M_j$ which is symmetric, moreover since M_j is skew-symmetric matrix, the product $M_j^T M_j$ is symmetric positive semi-definite. Then follows that $\sum_j^n c_j M_j^T M_j$ is positive semi-definite and therefore convex, provided that the sum of weights is convex i.e., $\sum_j^n c_j = 1$.

5. Optimal Quaternion using 4D Geometric Algebra

The most time consuming task of the estimation is to compute the eigenrotor of M associated with the greatest eigenvalue. In this section we show how to find the largest eigenvalue and its corresponding eigenrotor i.e., the required quaternion, using the 4D Geometric Algebra \mathbb{G}_4 which is robust, efficient and accurate.

Let us define four vectors $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4$ corresponding to the columns of matrix M :

$$M = \sum_j^n c_j M_j^T M_j = \begin{bmatrix} \mathbf{m}_1 & \mathbf{m}_2 & \mathbf{m}_3 & \mathbf{m}_4 \end{bmatrix} \quad (5.1)$$

The matrix system that we want to solve is $MR = \lambda R$ for some λ corresponding to the largest eigenvalue of M . We can write the system in its homogeneous form:

$$\begin{bmatrix} \mathbf{m}_1 - \lambda e_1 & \mathbf{m}_2 - \lambda e_2 & \mathbf{m}_3 - \lambda e_3 & \mathbf{m}_4 - \lambda e_4 \end{bmatrix} \begin{bmatrix} w \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (5.2)$$

The matrix in equation 5.2, called *characteristic matrix*, is of rank 3 and thus singular i.e., the vectors $(\mathbf{m}_1 - \lambda e_1)$, $(\mathbf{m}_2 - \lambda e_2)$, $(\mathbf{m}_3 - \lambda e_3)$ and $(\mathbf{m}_4 - \lambda e_4)$ are linearly dependent. It follows that its outer product must be zero:

$$(\mathbf{m}_1 - \lambda e_1) \wedge (\mathbf{m}_2 - \lambda e_2) \wedge (\mathbf{m}_3 - \lambda e_3) \wedge (\mathbf{m}_4 - \lambda e_4) = 0 \quad (5.3)$$

Which is called the *characteristic outer product* and is equivalent to the characteristic polynomial $P(\lambda) = \det(M - \lambda I)$. A simple way to find the largest eigenvalue is using Newton-Raphson method $\lambda_{i+1} = \lambda_i - P(\lambda)/P'(\lambda)$. This method is indeed robust given that all eigenvalues are non-negative real numbers, it wasn't however for methods based on Davenport's matrix [2] which eigenvalues can be negative. We found that $\sum_j c_j (\|p_j\|^2 + \|q_j\|^2)$ is a robust guess for Newton-Raphson iteration because it is close to largest eigenvalue (under certain input it is the actual largest eigenvalue, as discussed earlier) and thus converges to it in few iterations. For the sake of completeness we show the first derivative of 5.3:

$$\begin{aligned} P'(\lambda) = & -e_1 \wedge (\mathbf{m}_2 - \lambda e_2) \wedge (\mathbf{m}_3 - \lambda e_3) \wedge (\mathbf{m}_4 - \lambda e_4) \\ & - (\mathbf{m}_1 - \lambda e_1) \wedge e_2 \wedge (\mathbf{m}_3 - \lambda e_3) \wedge (\mathbf{m}_4 - \lambda e_4) \\ & - (\mathbf{m}_1 - \lambda e_1) \wedge (\mathbf{m}_2 - \lambda e_2) \wedge e_3 \wedge (\mathbf{m}_4 - \lambda e_4) \\ & - (\mathbf{m}_1 - \lambda e_1) \wedge (\mathbf{m}_2 - \lambda e_2) \wedge (\mathbf{m}_3 - \lambda e_3) \wedge e_4 \end{aligned}$$

Following [3] we solve the system 5.2 using outer products. For a linear system $A\mathbf{x} = b$ the authors in [3] defines N linear equations of the form $A_j^T \mathbf{x} = b_j$ each of which corresponds to a dual hyper-plane of the form $\mathbf{a}_j \equiv A_j - b_j e_0$ where the solution \mathbf{x} must lie on. The e_0 is an *homogeneous* basis vector needed for enabling projective geometry, enlarging the base space to $N + 1$, which interpretation is to be the offset of the hyper-plane. So solution of the linear system is the intersection of N dual hyper-planes, which is given by its outer product.

$$\alpha(\mathbf{x} + e_0)^* = \mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \dots \wedge \mathbf{a}_N \quad (5.4)$$

Where the symbol $*$ is the dual operator of the $N + 1$ space and α is a weight factor. After taking the dual of $\alpha(\mathbf{x} + e_0)^*$ and divide by the coefficient of e_0 (which is α), the solution \mathbf{x} can be read off the coefficients of the 1-vector.

We know that the null space of $(M - \lambda I)$ is of rank one. Algebraically this means that one of its column vectors is redundant i.e., it can be written in term of the others. In linear algebra this means that the system has infinitely many solutions. The geometric interpretation is that all hyper-planes intersect in a line passing through the origin. Since all solutions lie on the same line they only differ by a scalar term, therefore a particular solution can be found by fixing a scalar. The homogeneous component e_0 of the hyperplane is such scalar (analog to the hyper-plane's offset as in [3]) since it affects only the scale of the solution not its attitude. In linear algebra language it is equivalent to set one offset value at the right hand side of 5.2, however that system can't be solved in linear algebra because it requires to invert a singular matrix.

We define the dual hyper-planes passing through the origin as:

$$\mathbf{a}_i \equiv \mathbf{m}_i - \lambda e_i \quad (5.5)$$

Although it is enough to fix the scale of a single hyper-plane to get a solution it is inconvenient to do so, as will be explained in Section 7. It is

more robust to fix the scale of all hyper-planes to some $\gamma \neq 0$ which is a scalar value. Intersection can then be found by taking the outer product as:

$$\alpha(\mathbf{x} + 1)^* = (\mathbf{a}_1 + \gamma) \wedge (\mathbf{a}_2 + \gamma) \wedge (\mathbf{a}_3 + \gamma) \wedge (\mathbf{a}_4 + \gamma) \quad (5.6)$$

Distributing the outer product and keeping the terms of grade-3 we get:

$$\alpha(\mathbf{x} + 1)^* = \gamma(\mathbf{a}_1 \wedge \mathbf{a}_3 \wedge \mathbf{a}_4 + \mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \mathbf{a}_4 + \mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \mathbf{a}_3 + \mathbf{a}_2 \wedge \mathbf{a}_3 \wedge \mathbf{a}_4) \quad (5.7)$$

Here the symbol $*$ is the dual operator of the 4D space (not 5D as in [3] which allow us to be more efficient) and α is a weight factor. After taking the dual of $\alpha(\mathbf{x} + 1)^*$ the eigenvector \mathbf{x} can be read off the coefficients of the 1-vector. Notice that solution needs to be normalized.

6. Optimal Computation of M

The symmetric matrix $M_j^T M_j$ has a simple form in terms of p_j and q_j :

$$M_j^T M_j = 2 \begin{bmatrix} p_j^T q_j & (p_j \times q_j)^T \\ p_j \times q_j & p_j q_j^T + q_j p_j^T - p_j^T q_j I_{3 \times 3} \end{bmatrix} + (\|p_j\|^2 + \|q_j\|^2) I_{4 \times 4} \quad (6.1)$$

All terms of 6.1 can be derived from the covariance matrix $D = p_j q_j^T$ plus the quantity $\|p_j\|^2 + \|q_j\|^2$. Since matrix D is of 3×3 its computation is more efficient than the whole 6.1. Details can be found in Section 8.

7. Robustness and Singularities

As stated before, setting the scale of a single hyperplane is enough to get a valid eigenrotor. However, the choice of which hyperplane to constraint is problematic. For instance, assuming input vectors without noise, constraining the hyperplane corresponding to the w component of the rotor won't work because when the angle of rotation is π the $w = \cos(\pi/2) = 0$ and so its scale cannot be constrained. Similarly, constraining one of the hyper-planes corresponding to B_1 , B_2 or B_3 won't work because when the angle of rotation is 0 or 4π the $\sin(2\pi) = 0$ and so on. Then, it is complex task to avoid all problematic situations. By constraining all hyper-planes as in 5.6 our geometric algebra method does not suffer from any of those singularities.

Some of the above mentioned singularities are present in classic methods such as QUEST [14] and FOMA [10] but also on methods derived from those, including recent descendants based on analytic formulas [16–19,21]. All those methods are based on finding the eigenvector corresponding to largest eigenvalue of Davenport's matrix [2]. Since that is an indefinite matrix, some eigenvalues are positive and some negative, the Newton-Raphson can fail to find the max eigenvalue (which can be negative). So significant effort has been put on finding fast and robust analytic solutions to the quartic polynomial but no advances has been made on improving robustness on finding the associated eigenvector besides doing *sequential rotations* i.e., a π angle rotation of input data accomplished by changing the signs of two columns of the covariance matrix of [14]. The QUEST method performs sequential

rotations one axis at a time, until an acceptable reference coordinate system is found.

8. Algorithms

The pseudo-code of proposed method is shown in Algorithm 1.

Algorithm 1 Fast Rotor Estimation

Require: $P = \{p_j\}_{j=1}^n, Q = \{q_j\}_{j=1}^n, C = \{c_j\}_{j=1}^n$

- 1: $S = D = 0, \gamma = 1$
- 2: **for** $j = 1$ **to** n **do**
- 3: $S = S + c_j(p_j \cdot p_j + q_j \cdot q_j)$
- 4: $D = D + c_j p_j q_j^T$
- 5: **end for**
- 6: $\mathbf{m}_1 = (\frac{1}{2}S + Tr(D))e_1 + (D_{12} - D_{21})e_2 + (D_{20} - D_{02})e_3 + (D_{01} - D_{10})e_4$
- 7: $\mathbf{m}_2 = (D_{12} - D_{21})e_1 + (2D_{00} + \frac{1}{2}S - Tr(D))e_2 + (D_{01} + D_{10})e_3 + (D_{20} + D_{02})e_4$
- 8: $\mathbf{m}_3 = (D_{20} - D_{02})e_1 + (D_{01} + D_{10})e_2 + (2D_{11} + \frac{1}{2}S - Tr(D))e_3 + (D_{12} + D_{21})e_4$
- 9: $\mathbf{m}_3 = (D_{01} - D_{10})e_1 + (D_{20} + D_{02})e_2 + (D_{12} + D_{21})e_3 + (2D_{22} + \frac{1}{2}S - Tr(D))e_4$
 {Newton-Raphson}
- 10: $\lambda_0 = S$
- 11: **repeat**
- 12: $\lambda_{i+1} = \lambda_i - P(\lambda_i)/P'(\lambda_i)$
- 13: **until** $\|\lambda_{i+1} - \lambda_i\| < \epsilon$
- 14: $\mathbf{a}_1 = \mathbf{m}_1 - \lambda e_1$
- 15: $\mathbf{a}_2 = \mathbf{m}_2 - \lambda e_2$
- 16: $\mathbf{a}_3 = \mathbf{m}_3 - \lambda e_3$
- 17: $\mathbf{a}_4 = \mathbf{m}_4 - \lambda e_4$
- 18: $\mathbf{X} = \gamma(\mathbf{a}_1 \wedge \mathbf{a}_3 \wedge \mathbf{a}_4 + \mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \mathbf{a}_4 + \mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \mathbf{a}_3 + \mathbf{a}_2 \wedge \mathbf{a}_3 \wedge \mathbf{a}_4)$
- 19: $R = \text{normalize}(\langle \mathbf{X}^* \rangle_1)$
- 20: **return** $R(0) + R(1)e_{12} + R(2)e_{13} + R(3)e_{23}$

An optimized C++ code using the Eigen library [6] and GAALOP [7] is publicly available on GitHub [9]

9. Comparisons

We selected several representative methods e.g. FLAE [19], SVD [1], QUEST [14], FA3R [16], Symbolic [17], Q-Method [14] and FOMA [10] for comparison. The Eigen library [6] was employed to implement all of them. The tests were ran on a MacBook Pro laptop with Intel Core i7 CPU running at 2,5 GHz. The Clang C++ compiler was used with -Ofast option enabled.

We first show that some methods fails on producing meaningful rotations when certain input data is provided. Table 1 shows four cases which causes error. First two cases consist of having input vectors rotated $\pi/2$ and π radians without noise. Other case is when the 3D vectors are projected to the YZ -plane and rotated around the X -axis without noise. The last case is similar but projected to the XZ -plane. The root mean-squared error (RMSE) is used to describe the difference of accuracy.

Algo	$\pi/2$	π	YZ Plane	XZ Plane	XY Plane
FLAE	4.41×10^{-28}	696	392	1328	3.25×10^{-28}
Symbolic	1.78×10^{-26}	1053	392	1321	4.70×10^{-27}
FA3R	4.00×10^{-28}	5.41×10^{-29}	2.25×10^{-09}	2.76×10^{-08}	4.98×10^{-08}
Proposed	7.32×10^{-18}	1.79×10^{-18}	4.05×10^{-28}	3.43×10^{-28}	2.22×10^{-24}
Q-Method	3.78×10^{-28}	2.89×10^{-29}	7.10×10^{-28}	1.45×10^{-28}	9.57×10^{-29}
QUEST	82.6	2.18×10^{-29}	6.09×10^{-26}	9.93×10^{-28}	8.36×10^{-29}
FOMA	6.08×10^{-28}	2.95×10^{-28}	3.93×10^{-26}	2.13×10^{-27}	1.19×10^{-28}
SVD	4.41×10^{-28}	7.51×10^{-29}	3.35×10^{-28}	9.83×10^{-29}	1.09×10^{-28}

TABLE 1. RMSE of one million vector alignments of a thousand input vectors without noise. Some methods failed when 3D vectors were rotated $\pi/2$ and π and when vectors are projected to the coordinate planes

We also compared the robustness of the methods when input data has extreme noise applied to the length of input vectors. Figure 1 shows that some methods often fails to find meaningful rotations. We took SVD as reference. Notice that the amount of RMSE is almost the same on successes and on failures, evidencing that almost all methods builds, in essence, the Q-Method's matrix.

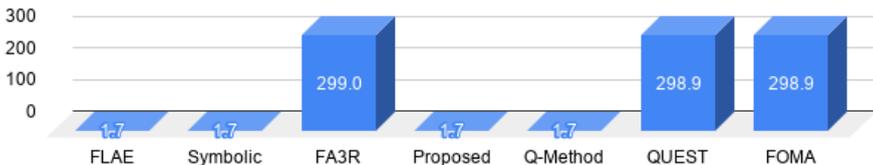


FIGURE 1. Sum of RMSE of one million vector alignments of a thousand input vectors. We applied Gaussian random noise to angles and random noise to lengths

Among the compared methods SVD, Q-Method and the ours are the most robust. Performance is shown in Figure 2 compared to the other methods. Our experiments shows that when the number of input vectors is up to 1000 our method is faster than Q-Method. Beyond that point the trend shows that Q-Method is faster.

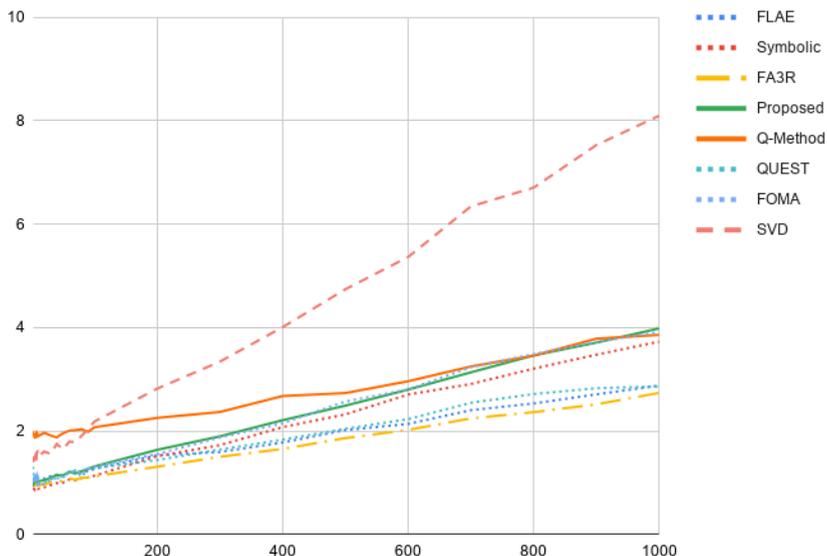


FIGURE 2. Performance comparison: number of vectors v.s. execution time (μs)

10. Conclusion

We presented a novel method for estimating the best quaternion aligning two sets of corresponding vectors and bivectors. Since we maximize a convex energy functional we only deal with non-negative eigenvalues, which makes Newton-Raphson iteration robust. We find the eigenvector corresponding to the largest eigenvalue intersecting hyper-planes in the language of geometric algebra. Results shows that our method is a good alternative in terms of robustness and accuracy. Its speed is competitive when the number of input vectors is relatively small.

References

- [1] K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(5):698–700, Sep. 1987.
- [2] P. B. Davenport. A vector approach to the algebra of rotations with applications. 1968.
- [3] S. De Keninck and L. Dorst. Geometric algebra levenberg-marquardt. In M. Gavrilova, J. Chang, N. M. Thalmann, E. Hitzler, and H. Ishikawa, editors, *Advances in Computer Graphics*, pages 511–522, Cham, 2019. Springer International Publishing.

- [4] L. Dorst and J. Lasenby, editors. *Guide to Geometric Algebra in Practice*. Springer, 2011.
- [5] D. Eggert, A. Lorusso, and R. Fisher. Estimating 3-d rigid body transformations: a comparison of four major algorithms. *Machine Vision and Applications*, 9(5):272–290, Mar 1997.
- [6] G. Guennebaud, B. Jacob, et al. Eigen v3. <http://eigen.tuxfamily.org>, 2010.
- [7] D. Hildenbrand, P. Charrier, C. Steinmetz, and J. Pitt. Gaalop home page. Available at <http://www.gaalop.de>, 2014.
- [8] B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.
- [9] M. C. Lopez Belon. Project on github. Available at <https://github.com/mauriciocele/fast-rotor-estimation>, 2019.
- [10] L. Markley. Attitude determination from vector observations: A fast optimal matrix algorithm. *Journal of the Astronautical Sciences*, 41, 07 1993.
- [11] L. Markley. 30 years of wahba’s problem. 02 1999.
- [12] D. Mortari. EULER-q algorithm for attitude determination from vector observations. *Journal of Guidance, Control, and Dynamics*, 21(2):328–334, 1998.
- [13] C. B. U. Perwass. *Geometric algebra with applications in engineering*, volume 4 of *Geometry and Computing*. Springer, Berlin; Heidelberg, 2009.
- [14] M. D. Shuster and S. D. Oh. Three-axis attitude determination from vector observations. *Journal of Guidance, Control, and Dynamics*, 4(1):70–77, 1981.
- [15] G. Wahba. A Least Squares Estimate of Satellite Attitude, 1965.
- [16] J. Wu, M. Liu, Z. Zhou, and R. Li. Fast rigid 3d registration solution: A simple method free of svd and eigen-decomposition, 2018.
- [17] J. Wu, M. Liu, Z. Zhou, and R. Li. Fast symbolic 3d registration solution, 2018.
- [18] J. Wu, Z. Zhou, J. Chen, H. Fourati, and R. Li. Fast Complementary Filter for Attitude Estimation Using Low-Cost MARG Sensors. *IEEE Sensors Journal*, 16(18):6997–7007, 2016.
- [19] J. Wu, Z. Zhou, B. Gao, R. Li, Y. Cheng, and H. Fourati. Fast Linear Quaternion Attitude Estimator Using Vector Observations. *IEEE Transactions on Automation Science and Engineering*, X(X):1–13, 2017.
- [20] Y. Yang. Attitude determination using Newton’s method on Riemannian manifold. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 229(14):2737–2742, 2015.
- [21] Y. Yang and Z. Zhou. An analytic solution to wahba’s problem. *Aerospace Science and Technology*, 30(1):46–49, Oct 2013.

Mauricio Cele Lopez Belon
Madrid, Spain
e-mail: mclopez@outlook.com