On the minimization over $SO(3)$ Manifolds

Frank O. Kuehnel

RIACS NASA Ames Research Center, MS 269-2, Moffett Field, CA 94035, USA
kuehnel@email.arc.nasa.gov

Abstract. In almost all image-model and model-model registration problems the question arises as to what optimal rigid body transformation applies to bring a physical 3-dimensional model in alignment with the observed one. Data may also be corrupted by noise. Here I will present the exponential and quaternion representations for the $SO(3)$ group. I will present the technique of compounding derivatives and demonstrate that it is most suited for dealing with numerical optimization problems that involve rotation groups.

1 Introduction

A common task in computer vision is matching images or features and estimating essential transformation parameters [1]. In the weak perspective regime the 2-dimensional affine image transformation with 6 parameters is applicable; in general a perspective transformation applies. Amongst

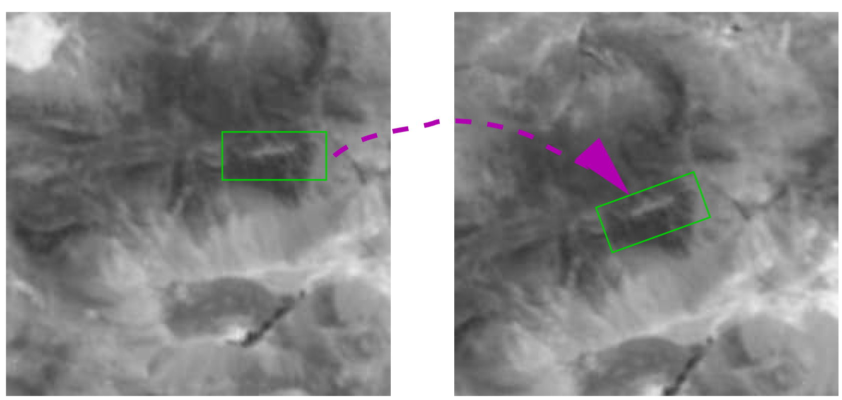


Fig. 1. Fly-by satellite view on a rendered region of duckwater. The main difference between both images is a rotation about the viewing axis.

those parameters are 3 Euler-angles that describe the orientation of the viewer with respect to a world-coordinate system, see Fig. 1 as an example. Then neglecting other free parameters, we could formulate the image matching problem as finding a minimum to the log- likelihood

$$-\log P \propto \sum_p (I_p - \hat{I}_p(\phi, \theta, \psi))^2, \quad (1)$$

where we sum over all pixels in the observed image data I_p and the expected image $\hat{I}_p$. It is well known that numerical optimiztion algorithms with Euler-angle representation have numerical problems. Several alternatives to Euler-angle representations exist. However, a prior it is not clear how well various representations and minimization methods will perfom. The subject of my investigation is the types of suitable rotation group representations and their application in numerical minimization algorithms. I will compare the convergence rate of various methods in the special case of matching point-clouds.

In the next section, I will present suitable rotation group representations, followed by their application within numerical optimization algorithms.

2 $SO(3)$ reprsentations

I briefly introduce the euler, the exponential and the quaternion representation. An extensive introduction to rotation groups and parametrization can be found in [2].

2.1 Euler-angle representation

The rotation matrix $\mathbf{R}$ represents an orthonormal transformation, $\mathbf{R}\mathbf{R}^T = I$, $\det(\mathbf{R}) = 1$, as such it can be decomposed into simpler rotation matrices,

$$\mathbf{R} = \mathbf{R}_z(\phi)\ \mathbf{R}_y(\theta)\mathbf{R}_z(\psi), \quad (2)$$

with

$$\mathbf{R}_z(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{R}_y(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}. \quad (3)$$

The decomposition (2) is not unique. We adhere to the “NASA Standard Aerospace” convention [3] with ψ the precession, θ the nutation and ϕ as the spin. The derivatives with respect to the Euler-angles are easily obtained and will not be given here.

2.2 Exponential representation, (Axis-angle)

The axis-angle representation is frequently used in kinematics [4] and commonly referred to as the exponential representation. The rotation matrix $\mathbf{R}$ is obtained by exponentiation of a matrix $\mathbf{H}$. This generating matrix $\mathbf{H}$ must be antisymmetric. and in 3-dimensions constitutes the cross product operator,

$$\mathbf{H} = [\mathbf{r}]_\times = \begin{pmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{pmatrix}, \quad \mathbf{r} = \boldsymbol{\omega}/\theta, \theta = |\boldsymbol{\omega}|, \tag{4}$$

where $\boldsymbol{\omega}$ is called the Rodrigues vector. We notice that in the limit $\theta \to 0$ the Rodrigues vector is ill-defined and we need proper limit considerations [5]. The rotation matrix $\mathbf{R}$ and its generating matrix $\mathbf{H}$ are related by,

$$\mathbf{R} = \exp(\theta\mathbf{H}) = \mathbf{I} + \sin\theta\mathbf{H} + (1 - cos\theta)\mathbf{H}^2. \tag{5}$$

In minimization problems that utilize gradients we require the knowledge of derivatives with respect to the components of the Rodrigues vector ω_{xyz}. In the following I introduce the index notation: Greek indices run over $\{1, 2, 3\}$ which is synonymous for $\{x, y, z\}$, repeated indices are implicitly summed over[1]. Now the derivatives are easily obtained

$$\frac{d\mathbf{R}}{d\omega_\alpha} = \sin\theta\frac{d\mathbf{H}}{dr_\beta}\frac{dr_\beta}{d\omega_\alpha} + (1 - \cos\theta)\frac{d\mathbf{H}^2}{dr_\beta}\frac{dr_\beta}{d\omega_\alpha} + \cos\theta\mathbf{H}r_\alpha + \sin\theta\mathbf{H}^2 r_\alpha, \tag{6}$$

where

$$\frac{dr_\beta}{d\omega_\alpha} = (\delta_{\alpha\beta} - r_\alpha r_\beta)/\theta, \quad \frac{d\theta}{d\omega_\alpha} = r_\alpha. \tag{7}$$

Hence, using

$$\frac{d\mathbf{H}}{dr_x}r_x + \frac{d\mathbf{H}}{dr_y}r_y + \frac{d\mathbf{H}}{dr_z}r_z = \mathbf{H} \tag{8}$$

and

$$\frac{d\mathbf{H}^2}{dr_x}r_x + \frac{d\mathbf{H}^2}{dr_y}r_y + \frac{d\mathbf{H}^2}{dr_z}r_z = 2\mathbf{H}^2 \tag{9}$$

we obtain the rotation matrix derivative with respect to the Rodrigues vector component

$$\frac{d\mathbf{R}}{d\omega_\alpha} = \frac{d\mathbf{H}}{dr_\alpha}\frac{\sin\theta}{\theta} + \frac{d\mathbf{H}^2}{dr_\alpha}\frac{(1 - \cos\theta)}{\theta} + \left(\mathbf{H}(\cos\theta - \sin\theta/\theta) + \mathbf{H}^2(\sin\theta - 2\frac{1 - \cos\theta}{\theta})\right) r_\alpha. \tag{10}$$

[1] summation over recurrent indicies is call Einstein summation

.

It is emphasized that the natural endowed algebra sturcture of the vector space $\boldsymbol{\omega} \in \mathbb{R}^3$ is not isomorphic to the multiplication in the $SO(3)$ group,

$$\mathbf{R}(\boldsymbol{\omega}_1)\,\mathbf{R}(\boldsymbol{\omega}_2) \quad \neq \quad \mathbf{R}(\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2)\,! \tag{11}$$

2.3 Quaternion representation

Quaternions can compactly represent rotation matrices. We will see that the compounding (non-commutative multiplication) operation is isomorphic to the matrix multiplication in $SO(3)$. Let's introduce general quaternions by

$$\hat{Q} = (q_0, \mathbf{q}), \quad \mathbf{q} = q_1\mathbf{e}_1 + q_2\mathbf{e}_2 + q_3\mathbf{e}_3, \tag{12}$$

where $\mathbf{q}$ is a regular 3-dimensional vector with basis $\{\mathbf{e}_\alpha\}$ in $\mathbb{R}^3$. Addition of quaternions is component-wise, a product (*compounding*) of two quaternions as follows,

$$\hat{P} \circ \hat{Q} = (p_0 q_0 - \mathbf{p}^T\mathbf{q}, p_0\mathbf{q} + q_0\mathbf{p} + \mathbf{p} \times \mathbf{q}). \tag{13}$$

The notation in (13) utilizes the *dot* and *cross* product defined in $\mathbb{R}^3$. It is remarked that the product (13) is non-commutative but associative! In addition we have a conjugate operation,

$$\hat{Q} = (q_0, \mathbf{q}) \;\rightarrow\; \hat{Q}^* = (q_0, -\mathbf{q}), \tag{14}$$

hence we can write the squared norm as,

$$\hat{Q} \circ \hat{Q}^* = (q_0^2 + \mathbf{q}^T\mathbf{q}, 0) \equiv q_0^2 + \mathbf{q}^T\mathbf{q}. \tag{15}$$

We identify a scalar value $c \in \mathbb{R}$ with a quaternion that has "zero vector" components, $(c, 0) \equiv c$ and a vector $\mathbf{q} \in \mathbb{R}^3$ with a quaternion that has "zero scalar" component $(0, \mathbf{q}) \equiv \mathbf{q}$. Further, we will be *sloppy* with the quaternion product notation $\circ$ but implicitly assume that $\hat{Q}\hat{P}$ means $\hat{Q} \circ \hat{P}$.

The claim is that a rotation can be represented by

$$\hat{R} = (\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\mathbf{r}), \quad \hat{R}\hat{R}^* \equiv 1 \tag{16}$$

with θ the magnitude and $\mathbf{r}$ the direction of the Rodirigues vector given in (4) and below, such that

$$\mathbf{R}\,\mathbf{v} = \hat{R}\hat{V}\hat{R}^*, \quad \hat{V} = (0, \mathbf{v}). \tag{17}$$

Proof of (17) is easy and proceeds as follows: Writing equation (5) in the form with half-angles,

$$\begin{aligned}\mathbf{R}\,\mathbf{v} &= \cos^2\frac{\theta}{2} I\mathbf{v} + 2\cos\frac{\theta}{2}\sin\frac{\theta}{2}\mathbf{r}\times\mathbf{v} + \sin^2\frac{\theta}{2}\mathbf{r}\times(\mathbf{r}\times\mathbf{v}) + \sin^2\frac{\theta}{2}\mathbf{r}\mathbf{r}^T\,\mathbf{v} \\ &= (\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\mathbf{r})\circ(0,\mathbf{v})\circ(\cos\frac{\theta}{2}, -\sin\frac{\theta}{2}\mathbf{r}),\end{aligned} \tag{18}$$

which completes the proof!

Any quaternion $\hat{Q}$ with $\hat{Q}\hat{Q}^* = 1$ constitutes a rotation in $SO(3)$, the components of the rotation matrix $\mathbf{R}$ are easily obtained from (18),

$$\hat{Q} = (q_0, q_1, q_2, q_3), \quad q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1 \;\rightarrow$$
$$\mathbf{R}(\hat{Q}) = \begin{pmatrix} 1-2q_2^2-2q_3^2 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_2q_1+q_0q_3) & 1-2q_1^2-2q_3^2 & 2(q_2q_3-q_0q_1) \\ 2(q_3q_1-q_0q_2) & 2(q_3q_2+q_0q_1) & 1-2q_1^2-2q_2^2 \end{pmatrix} \tag{19}$$

Back to our initial statement, that we can easily idenitify an isomorphic structure between quaternions and rotation matrices,

$$\mathbf{R}(\hat{Q}_1)\,\mathbf{R}(\hat{Q}_2) \quad = \quad \mathbf{R}(\hat{Q}_1\circ\hat{Q}_2). \tag{20}$$

2.4 Quaternion differential algebra

As in the case for the Rodrigues vector, we are interested in the differential structure of the rotation matrix with respect to the quaternion components. Since we have 4 quaternion components and only 3 rodrigues vector ones, we also have to address the problem of overparametrization. The quaternion ring is endowed with a commutative operation (addition) and a non-commutative multiplicative operation (composition), it is natural to consider the two differential structures.

Additive differential rotation We can cast the rotation matrix $\mathbf{R}$ in a particular well suited quaternion form, essentially it's restating (18)

$$\mathbf{R} = \mathbf{I}q_0^2 + \mathbf{q}\mathbf{q}^T + 2q_0[\mathbf{q}]_\times + [\mathbf{q}]_\times[\mathbf{q}]_\times, \tag{21}$$

$\mathbf{I}$ is the *unit* matrix. Using tensor notation in (21),

$$I_{\mu\nu} = \delta_{\mu\nu}, \quad \left(\mathbf{q}\mathbf{q}^T\right)_{\mu\nu} = q_\mu q_\nu, \quad ([\mathbf{q}]_\times)_{\mu\nu} = \varepsilon_{\mu\rho\nu}q_\rho, \tag{22}$$

where $\delta_{\mu\nu}$ is the Kronecker-delta and $\varepsilon_{\mu\nu\rho}$ the total antisymmetric tensor. Now we consider a differential rotation $\mathbf{R}(\hat{Q}+d\hat{q})-\mathbf{R}(\hat{Q})$ and obtain

$$\frac{\partial R_{\mu\nu}}{\partial q_0}=2q_0\delta_{\mu\nu}+2\varepsilon_{\mu\rho\nu}q_\rho,\quad \frac{\partial R_{\mu\nu}}{\partial q_\eta}=2(\delta_{\eta\mu}q_\nu+q_\mu\delta_{\eta\nu}+q_0\varepsilon_{\mu\eta\nu}-\delta_{\mu\nu}q_\eta). \tag{23}$$

We used the following tensor contraction in (23)

$$\varepsilon_{\alpha\beta\gamma}\varepsilon_{\alpha\mu\nu}=\delta_{\beta\mu}\delta_{\gamma\nu}-\delta_{\beta\nu}\delta_{\gamma\mu}. \tag{24}$$

Compounded differential rotation Alternatively, we can compose the differential rotation as the compound of the deviation from the *unit*-quaternion $\hat{I}$ (this represents the identity operation, *null*-rotation), by

$$\delta\hat{Q}=(\hat{I}+\delta\hat{q})\circ\hat{Q}-\hat{Q}=\delta\hat{q}\circ\hat{Q},\quad (\hat{I}+\delta\hat{q})(\hat{I}+\delta\hat{q})^*=1. \tag{25}$$

Then, we can write a differential rotation $\mathbf{R}((\hat{I}+\delta\hat{q})\circ\hat{Q})-\mathbf{R}(\hat{Q})$ with the change of $\delta\hat{q}$ (25),

$$\begin{aligned}[R(\hat{Q}+\delta\hat{Q})-R(\hat{Q})]\mathbf{v}&=\delta\hat{Q}\circ(0,\mathbf{v})\circ\hat{Q}^*+\hat{Q}\circ(0,\mathbf{v})\circ\delta\hat{Q}^*\\ &=\delta\hat{q}\circ(0,\mathbf{v}')+(0,\mathbf{v}')\circ\delta\hat{q}^*\\ &=2\delta q_0\mathbf{v}'+2\delta\mathbf{q}\times\mathbf{v}'\\ &=2\delta q_0\,R(\hat{Q})\mathbf{v}+2\delta\mathbf{q}\times(R(\hat{Q})\mathbf{v}),\end{aligned} \tag{26}$$

hence the derivatives have a particularly simple form compared to (23),

$$\frac{\partial R_{\mu\nu}}{\partial\delta q_0}=2R_{\mu\nu},\quad \frac{\partial R_{\mu\nu}}{\partial\delta q_\alpha}=2\varepsilon_{\alpha\rho\mu}R_{\rho\nu}. \tag{27}$$

Overparametrization, connection with the Rodrigues vector The rotation group $SO(3)$ is a 3-parameter group, that is reflected in the number of Euler-angles and the vector components of the Rodrigues vector. Quaternions have 4 real components, seemingly we have increased the number of parameters. However, for a quaternion $\hat{Q}$ to represent a rotation it must lie on the normalized-sphere S^3,

$$\hat{Q}\hat{Q}^*=1. \tag{28}$$

This represents an additional constraint, that is not built into the quaternion-component derivatives (23,27).

We can choose to parameterize the quaternion with the Rodrigues vector (16), then the derivatives of a general rotation quaternion $\hat{q}$ with respect to the independent rodrigues vector $\boldsymbol{\omega}$ components are,

$$\frac{\partial \hat{q}}{\partial \omega_\alpha} = (-\frac{r_\alpha}{2}\sin\frac{\theta}{2}, \frac{r_\alpha}{2}\cos\frac{\theta}{2}\mathbf{r} + \frac{1}{\theta}\sin\frac{\theta}{2}(\delta_{\alpha\beta} - r_\alpha r_\beta)\mathbf{e}_\beta). \tag{29}$$

Here we assume that $0 < \theta$ or equivalently $0 < \mathbf{q}^2$. Equivalently we can express (29) in quaternion component form,

$$\frac{\partial q_0}{\partial \omega_\alpha} = -\frac{1}{2}q_\alpha,\ \frac{\partial q_\eta}{\partial \omega_\alpha} = \frac{1}{2}(q_0 - \operatorname{sinc} x)\frac{q_\alpha q_\eta}{\mathbf{q}^2} + \frac{\operatorname{sinc} x}{2}\delta_{\alpha\eta},\ x = \arctan\sqrt{\frac{\mathbf{q}^2}{q_0^2}}. \tag{30}$$

We notice, that (29,30) show that the quaternion-constraint manifests in non-linear functions.

On the other hand, if we were to parametrize the infinitesimal deviation from a unit-quaternion,

$$\hat{q} = \hat{I} + \delta\hat{q} = (q_0, \mathbf{q}) = (\cos\frac{\delta\theta}{2}, \sin\frac{\delta\theta}{2}\mathbf{r}), \quad \mathbf{r} = \delta\boldsymbol{\omega}/\delta\theta, \quad \delta\theta = |\boldsymbol{\omega}|\,. \tag{31}$$

then the quaternion-constraint is reflected by a rather simple linear dependence,

$$\left.\frac{\partial \delta q_0}{\partial \delta\omega_\alpha}\right|_{\mathbf{q}=0} = 0,\ \left.\frac{\partial \delta q_\eta}{\partial \delta\omega_\alpha}\right|_{\mathbf{q}=0} = \frac{1}{2}\delta_{\alpha\eta}. \tag{32}$$

Notice, that (32) is the limiting case of (30) for $\mathbf{q}^2 \to 0$.

3 Non-linear function minimization over S^3

Given a bounded scalar function with $SO(3)$ group parameters as arguments

$$M \le f(\hat{Q}) = f(R(\hat{Q})), \quad \hat{Q} \in S^3,\ M \in \mathbb{R} \tag{33}$$

we are to find the (not necessarily unique) optimal quaternion $\hat{Q}^{(f)}$ which minimizes the scalar function $f(\hat{Q})$ in (33). In most cases analytical closed form solutions to (33) are not available. Numerical minimization algorithms utilize gradient information and at most deal with quadratic expansions of (33). Therefore it is natural to lay focus on functions of the form

$$f = a + b_{\mu\nu}R_{\mu\nu} + c_{\mu\nu}^{\eta\rho}R_{\mu\nu}R_{\eta\rho}. \tag{34}$$

The arguments of f are implicitly subsumed in the rotation matrix components $R_{\mu\nu}$. Equation (34) is frequently the result of a quasi-Newton approximation to (1). Despite of the seemingly simple structure of Eqn. (34)

with rotation matrix components occuring at most in order 2, hence we could call them $R_{\mu\nu}$*-harmonic functions* , it generally doesn't possess a closed form solution. Only for the particular case presented below the solution is known.

3.1 Pseudo-quadratic functions

The special case of finding a unit quaternion $\hat{Q}_{\min}$ which minimizes

$$f_{\mathrm{s}}(R(\hat{Q})) = a_0 + \sum_{i=1}^{N} \|\mathbf{R}\mathbf{c}_i - \mathbf{d}_i\|^2 \tag{35}$$

can be solved analytically [6, 7]. (35) represents the case where a cloud of N points $\mathbf{c}_i$ in $\mathbb{R}^3$ are mapped by a rotation such that the result most closely resembles the cloud $\mathbf{d}_i$. We can write (35) in tensor notation,

$$f_{\mathrm{s}} = a_0 + d_\mu d_\mu - 2 d_\mu c_\nu R_{\mu\nu} + c_\nu c_\rho \delta_{\mu\eta} R_{\mu\nu} R_{\eta\rho}, \tag{36}$$

where we have omitted the sum over point index i, thus in terms of tensor coefficients (34) we have

$$b_{\mu\nu} = -2 c_\mu d_\nu, \quad c_{\mu\nu}^{\eta\rho} = c_\nu c_\rho \delta_{\mu\eta}.$$

Now, we use the orthogonality relation,

$$c_{\mu\nu}^{\eta\rho} R_{\mu\nu} R_{\eta\rho} = c_\nu c_\rho \delta_{\nu\rho} \tag{37}$$

and obtain an expression that is linear in the matrix elements $R_{\mu\nu}$, hence (35) constitutes only a *pseudo-quadratic* function,

$$f_{\mathrm{s}} = a + b_{\mu\nu} R_{\mu\nu}, \quad b_{\mu\nu} = -2 \sum_{i=1}^{n} \mathbf{d}_i \mathbf{c}_i^T, \quad a = a_0 + \mathbf{c}_i^T \mathbf{c}_i + \mathbf{d}_i^T \mathbf{d}_i. \tag{38}$$

According to [7] we can find a solution to the $R_{\mu\nu}$ linear problem (38) by simply decomposing the general rank-2 tensor $b_{\mu\nu}$ and restating the problem as in (35). Then we solve for $\hat{Q}_{\min}$ by noticing that f_s is quadratic in the quaternion components and therefore can be written as

$$f_{\mathrm{s}} = a_0 + (q_0, \mathbf{q}) \sum_{i=1}^{n} B_i^T B_i (q_0, \mathbf{q})^T, \tag{39}$$

with 4×4 matricies B_i,

$$B_i = \begin{bmatrix} 0 & (\mathbf{c}_i - \mathbf{d}_i)^T \\ \mathbf{d}_i - \mathbf{c}_i & [\mathbf{d}_i + \mathbf{c}_i]_\times \end{bmatrix}.$$

A solution to (35) is given by the eigenvector $\hat{q}$ of $B_i^T B_i$ with minimal eigenvalue.

4 Gradient search methods on the S^3 manifold

In the absence of analytical solutions we resort to numerical minimization methods. To measure the minimization search performance of various methods we focus on the *pseudo-quadratic* function f_s and argue that the results are representative even in the general scenario of (34).

4.1 Iterative quadratic expansion in the Rodrigues vector $\boldsymbol{\omega}$ or Euler-angles

We start with a Rodrigues vector $\boldsymbol{\omega}_0$ and expand the function f_s to quadratic order in $\boldsymbol{\omega}$ about $\boldsymbol{\omega}_0$. We keep the rotation matrix entries $R_{\mu\nu}$ to second order, (this basically means that we don't use the orthogonality relationship),

$$f_s(\Delta\boldsymbol{\omega}) \approx \tilde{a} + \underbrace{(b_{\mu\nu} + 2R_{\mu\rho}c_\rho c_\nu)\frac{\partial R_{\mu\nu}}{\partial\omega_\alpha}}_{\mathbf{b}^T}\Delta\omega_\alpha + \Delta\omega_\alpha \underbrace{\frac{\partial R_{\mu\nu}}{\partial\omega_\alpha}c_\nu c_\rho \frac{\partial R_{\mu\rho}}{\partial\omega_\beta}}_{\mathbf{C}}\Delta\omega_\beta, \tag{40}$$

where $\Delta\boldsymbol{\omega} = \boldsymbol{\omega} - \boldsymbol{\omega}_0$. Omitted are all contributions from $\partial^2 R_{\mu\nu}/\partial\omega_\alpha\partial\omega_\beta$. We need to calculate the exponential derivatives

$$\frac{\partial R_{\mu\nu}}{\partial\omega_\alpha} = \frac{\partial R_{\mu\nu}}{\partial q_{0,\eta}}\frac{\partial q_{0,\eta}}{\partial\omega_\alpha}, \tag{41}$$

which is exactly expression (10). Having obtained the minimal $\Delta\boldsymbol{\omega}$ we then update the rodrigues vector

$$\boldsymbol{\omega}_0 \rightarrow \boldsymbol{\omega}_1 = \boldsymbol{\omega}_0 + \Delta\boldsymbol{\omega},$$

and iteratively proceed with a local expansion (40) until convergence is reached.

expansion with Euler-angles The same outline as depicted above applies to a local quadratic expansion with Euler-angles. Instead of the Rodrigues vector derivatives (41) we need to calculate the Euler-angle derivatives,

$$\frac{\partial R_{\mu\nu}}{\partial\phi}, \quad \frac{\partial R_{\mu\nu}}{\partial\theta} \quad \text{and} \quad \frac{\partial R_{\mu\nu}}{\partial\psi}. \tag{42}$$

Then the update procedure follows the rule,

$$\{\phi_0, \theta_0, \psi_0\} \rightarrow \{\phi_1, \theta_1, \psi_1\} \; = \; \{\phi_0 + \Delta\phi, \theta_0 + \Delta\theta, \psi_0 + \Delta\psi\}\,.$$

4.2 Quaternion-path gradient search

We construct a product sequence (path) from a suitable starting point $\hat{Q}^{(0)}$,

$$\hat{Q}^{(k)} = (\hat{I} + \Delta\hat{q}^{(k)}) \circ (\hat{I} + \Delta\hat{q}^{(k-1)}) \circ \ldots \circ \hat{Q}^{(0)}, \tag{43}$$

with the constraint

$$\hat{Q}^{(k)} \in S^3 \quad \forall k, \tag{44}$$

meaning that $\hat{Q}^{(k)}$ at any step represents a rotation.

Newton-Raphson, Lagrange multiplier technique One way to constrain quaternions on a unit sphere is by a Lagrange multiplier technique. As a test example we aim to obtain a quaternion sequence to find the minimum of the pseudo-quadratic functional (35). As usual with the Lagrange multiplier technique, given f_{s} and the normalization contraint for $\delta\hat{q}$, we can deal with the function g

$$g = f_{\mathrm{s}} + \lambda((1+\delta q_0)^2 + \delta\mathbf{q}^2 - 1). \tag{45}$$

Now, if we write derivatives of g with respect to the multiplicative differential quaternions, we obtain

$$\frac{dg}{d\delta\hat{q}} = 0 \;\Rightarrow\; \begin{cases} b_{\mu\nu}R_{\mu\nu} + \lambda(1+\delta q_0) = 0 \\ \varepsilon_{\alpha\mu\rho}b_{\rho\nu}R_{\mu\nu} + \lambda\delta q_\alpha = 0, \quad \alpha \in \{1,2,3\} \end{cases} \tag{46}$$

Assuming R being constant for the gradient search, we find an easy solution to (46), i.e. starting with $R(\hat{Q}^{(0)})$, in the first step, then we can solve for the update $\hat{q}$,

$$\hat{q} = \hat{I} + \delta\hat{q}, \quad -\lambda = \frac{b_{\mu\nu}R_{\mu\nu}}{q_0}, \; \frac{q_\alpha}{q_0} = \varepsilon_{\alpha\mu\rho}\frac{R_{\mu\nu}b_{\rho\nu}}{R_{\eta\xi}b_{\eta\xi}}. \tag{47}$$

Having found $\hat{q}$ and properly normalized, we successively obtain a new starting point $\hat{Q}^{(1)} = \hat{q} \circ \hat{Q}^{(0)}$ with rotation matrix $R(\hat{Q}^{(1)})$ for the next search step and solve for another quaternion until convergence is reached. Experimental results follow.

Quasi-Newton method through infinitesimal exponential parametrization Instead of imposing the quaternion constraint by a Lagrange multiplier technique we could parametrize the deviation quaternion $(\hat{I} + \Delta\hat{q})$ in

(43) by its Rodrigues-vector and assume that we are in the linear regime (32),

$$f_{\mathrm{s}} \approx \tilde{a} + \underbrace{(b_{\mu\nu} + 2R_{\mu\rho}c_\rho c_\nu)\frac{\partial R_{\mu\nu}}{\partial\delta\omega_\alpha}}_{\mathbf{b}^T} \Delta\omega_\alpha + \Delta\omega_\alpha \underbrace{\frac{\partial R_{\mu\nu}}{\partial\delta\omega_\alpha}c_\nu c_\rho \frac{\partial R_{\mu\rho}}{\partial\delta\omega_\beta}}_{\mathbf{C}} \Delta\omega_\beta, \quad (48)$$

where $\tilde{a} = a - \mathbf{c}_i^T \mathbf{c}_i$. Using (27) and (32),

$$\frac{\partial R_{\mu\nu}}{\partial\delta\omega_\alpha} = \frac{\partial R_{\mu\nu}}{\partial\delta q_{0,\eta}}\frac{\partial\delta q_{0,\eta}}{\partial\delta\omega_\alpha} = 2(R_{\mu\nu}\frac{\partial\delta q_0}{\partial\delta\omega_\alpha} - \varepsilon_{\eta\mu\rho}R_{\rho\nu}\frac{\partial\delta q_\eta}{\partial\delta\omega_\alpha}) = -\varepsilon_{\alpha\mu\rho}R_{\rho\nu}, \quad (49)$$

and we obtain the expression for $\mathbf{b}$

$$\mathbf{b}_\alpha = -\varepsilon_{\alpha\mu\rho}\left(R_{\rho\nu}b_{\mu\nu} - 2R_{\mu\xi}c_\xi R_{\rho\nu}c_\nu\right). \quad (50)$$

To emphasize the structure of the symmetric matrix $\mathbf{C}$ we write,

$$\mathbf{D}_{\alpha\mu} = \frac{\partial R_{\mu\nu}}{\partial\delta\omega_\alpha}c_\nu = -\varepsilon_{\alpha\mu\rho}R_{\rho\nu}c_\nu, \quad \mathbf{C}_{\alpha\beta} = \mathbf{D}_{\alpha\mu}\mathbf{D}^T_{\mu\beta}. \quad (51)$$

We find $\Delta\boldsymbol{\omega}$ in equation (48) is quickly solved using SVD,

$$\frac{\partial f_{\mathrm{s}}}{\partial\Delta\boldsymbol{\omega}} = 0, \rightarrow \ \Delta\boldsymbol{\omega} = -\frac{1}{2}\mathbf{C}^{-1}\mathbf{b}, \quad (52)$$

and calculate the finite step update quaternion $(\hat{I} + \Delta\hat{q}^{(k)})$ using (31). With a new quaternion $\hat{Q}^{(k+1)}$ as obtained in (43), we calculate rotation matrices and derivatives thereof. This process is repeated until suffient convergence is reached.

5 Experimental results

Here I present numerical experiments on the convergence of the various gradient minimization methods. We compare the standard methods as described in section 4.1 with the ones utilizing the quaternion-path methods, see section 4.2. The iterative Euler-angle expansion and the iterative Rodrigues vector expansion are referred to as “Standard 1” and “Standard 2”. The two methods utilizing quaternion-path gradients are referred to as “Lagrange” and “Compounding”.

We initialize all experiments by a random selection of N points $\{\mathbf{c}_i\}$ and choosing a random transformation $\mathbf{R}(\hat{Q}^f)$ to obtain a point cloud $\{\mathbf{d}_i\}$ according to (35). Then we randomly initialize our first guess $\mathbf{R}(\hat{Q}^{(0)})$

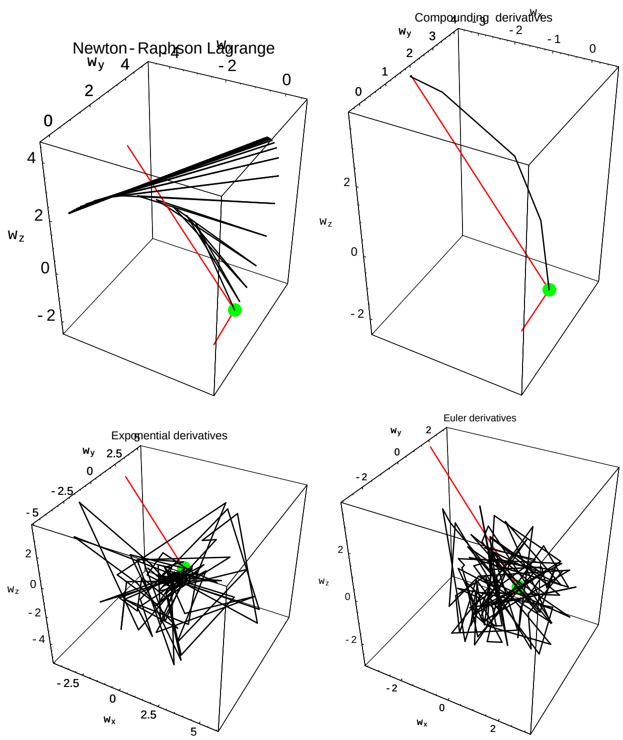
and employ the minimization methods. The graphs show the minimization trajectory in the Rodrigues vector space. The green dota represents the initial guess, the end-points of the two red lines the two possible exact solutions.

We track the convergence by counting the number of steps it takes to obtain the accuracy f_s. This corresponds to the residual value of mismatching the euclidian distance (35). We clearly notice that in all scenarios the compounding quaternion derivative method is vastly superior and has super-linear convergence. Also it never failed in all tested cases. Most rarely does the Euler-angle method succeed.

6 Conclusions

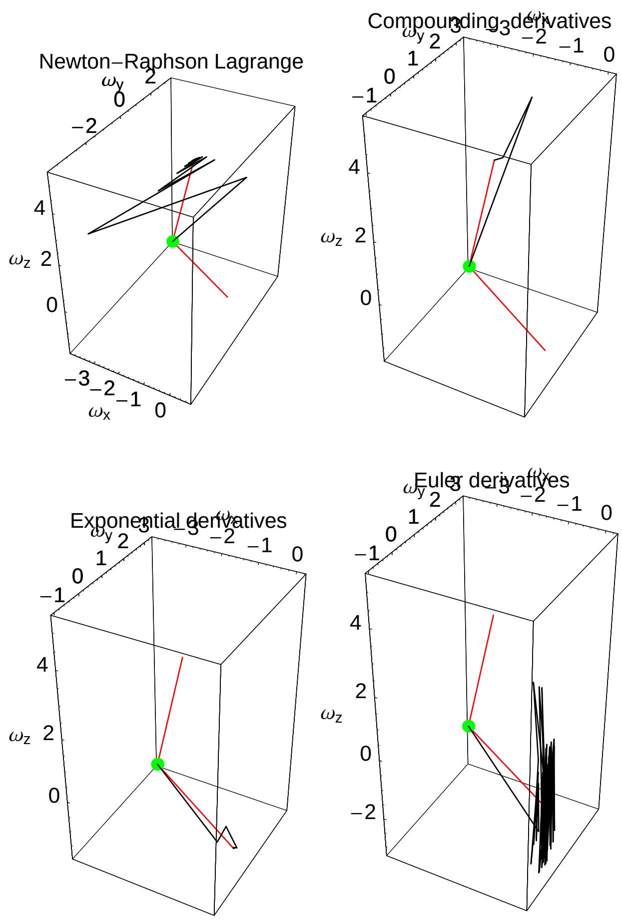
I have concisely introduced important rotation group $SO(3)$ representations and presented their differential structure. I've developed the new idea of a product (compounding) path-quaternions and demonstrated its advantages application towards numerical minimization methods. In the special point-matching case this method is vastly superior than all its alternatives. It is expected that the super-linear convergence of the compounding path-quaternion approach is retained even in the case of more general function types that depend on $SO(3)$ parameters.

6.1 $N = 2$ data points



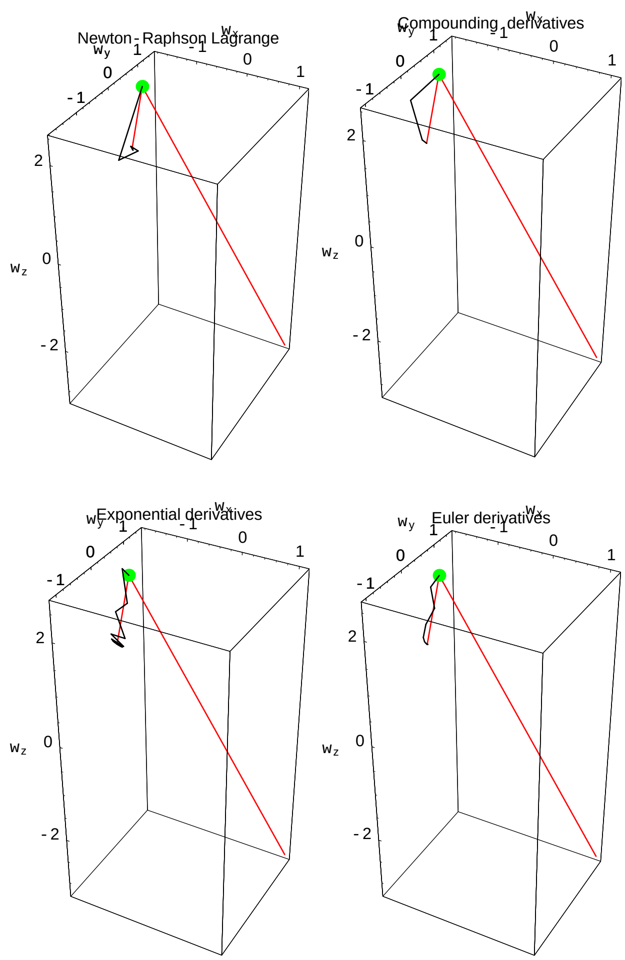
Accuracy	1.	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
Standard 1	∞	∞	∞	∞	∞	∞	∞	∞	∞
Standard 2	∞	∞	∞	∞	∞	∞	∞	∞	∞
Lagrange	∞	∞	∞	∞	∞	∞	∞	∞	∞
Compounding	5	5	5	6	6	6	6	6	7

6.2 $N = 5$ data points



Accuracy	1.	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
Standard 1	∞	∞	∞	∞	∞	∞	∞	∞	∞
Standard 2	5	6	6	7	8	9	9	10	11
Lagrange	15	20	24	29	34	39	44	48	53
Compounding	3	4	4	4	4	5	5	5	5

6.3 $N = 12$ data points



Accuracy	1.	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
Standard 1	7	8	10	12	13	15	17	18	20
Standard 2	22	34	46	59	72	84	97	∞	∞
Lagrange	9	12	14	17	20	23	25	28	31
Compounding	4	4	4	4	4	5	5	5	5

References

1. K. Kanatani, Analysis of 3-D Rotation Fitting, Transactions of pattern analysis and machine intelligence, IEEE, Vol. 16, no. 5, 1994.
2. S. L. Altman, *Rotations, Quaternions and Double Groups.* Clarendon Press, Oxford, 1986.

3. G. J. Minkler, *Aerospace Coordinate Systems and Transformations.* Magellan Book Company, Balitmore, MD, 1990.
4. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physiscs, Mechanics*, ed. L .D. Landau, Butterworth-Heinemann, 3rd edtion, 1976.
5. X. Pennec and J. P. Thirion, *A Framework for Uncertainty and Validation of 3-D Registration Methods based on Points and Frames*, International Journal of Computer Vision, 25(3), pp 203-229, 1997.
6. B. K. P. Horn, Closed Form Solutions of Absolute Orientations Using Unit Quaternions, Journal of Optical Society of America, **A-4**(4), pp 629-642, 1987.
7. J. Weng, T. S. Huang and N. Ahuja, Motion and Structure from Two Perspective Views: Algorithms, Error Analysis, and Error Estimation IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. II, No 5, pp 451-475, 1989.