EFFICIENT ITERATIVE ALGORYTHMS FOR MINIMIZING ATTITUDE FUNCTIONS AND THEIR APPLICATION TO THREE-AXIS ATTITUDE DETERMINATION

Michiaki Horii
National Space Development Agency of Japan
2-1-1, Sengen, Tsukuba City, Ibaraki Prefecture, 305-8505 Japan
horii.michiaki@nasda.go.jp

Abstract

Newton's method is applied to solve a minimization problem for general attitude functions and also for least squares type attitude functions. Simple solution formulae are derived and those results are examined with various three-axis attitude determination problems.

Key words: Attitude, Quaternion, Direction cosine matrix, Newton's method, Attitude determination

1. Introduction

A minimization problem for attitude functions is typically encountered in the attitude determination of the spacecraft. Difficulties of those problems are mainly in the fact that the totality of the attitude is a compact manifold and no global coordinates with the number of the freedom of the attitude (three for three-axis attitude) are available. Therefore much effort has been devoted to the treatment of the attitude representation method.

Two major attitude global representations are; the quaternion and the direction cosine matrix (DCM)[1,2]. Quaternion is a four-dimensional vector and one constraint is attached to it. The direction cosine matrix which is regarded as attitude itself comprises a set of nine elements with six orthonormal constraints. One from these two representations is selected and applied to a real application comparing their advantages and losses under the given situation. In recent analyses local coordinates such as Euler angles are rarely used for a global attitude analysis.

An important attitude determination problem was set by Wahba in 1965 and it requires minimization of squares of residuals in vector observations[3]. The problem was solved by various algebraic methods, among which the q-method[1] and the QUEST[6] utilizing the quaternion converted Wahba's problem to an eigenvalue problem of a 4x4-matrix and solved it. Also QUEST gave a covariance matrix to the solution. Another efficient solution to Wahba's problem was given by Markley[7] who treated the problem in the DCM formulation and made use of the singular value decomposition of a 3x3-matrix.

Historically Wahba's problem was algebraically solved shortly after the problem was announced. Among them Brock's necessary condition[4,5] which a solution must satisfy is important and strongly related to the present work.

Recently another type of problem was defined relating to the attitude determination using GPS phase difference measurements[8,9]. The problem is different from Wahba's problem and an iterative solution was derived in the quaternion formulation.

This paper first deals with general minimization problems of attitude functions with both the quaternion and the DCM representations. We set both a general type minimization problem and a least squares (LSQ) type minimization problem. Our intention is to derive simple iterative solutions assuming no explicit function forms. For this purpose we apply Newton's method. After that the obtained formulae will be examined by applying them to the above-mentioned attitude determination problems.

Among the previous works one of the most relevant methods to the present work was that of Gray[10]. He treated a least squred type minimization problem in the DCM representation and adopted local coordinates. He demonstrated that his method was very effectively applicable to many attitude determination problems. This paper generalize his method with explicit expressions for itetative solution in the DCM formulation.

Another very efficient iterative solution to Wahba's problem in the quaternion formulation was derived by Yoshikawa and et al.[11]. They applied Newton's method together with quaternion operations for multiplication. The present paper, however, does not refer to such quaternion operations and use rather the unit length condition of quaternion only.

2. Problem Description

2.1. General Minimization Problem

This paper first deals with a minimization problem for attitude functions in various forms.

2.1.1. Quaternion Formulation

A.1. General Type Formulation

$$\text{minimize} \quad f(q) \qquad (1)$$

subject to

$$q \equiv (q_1, q_2, q_3, q_4)^T \in S^3 \text{ or } q^T q = 1 \qquad (2)$$

A.2. LSQ Type Formulation

$$\text{minimize } f(q) = (\mathbf{y} - \mathbf{h}(q))^T \mathbf{W}(\mathbf{y} - \mathbf{h}(q)) \qquad (3)$$

subject to (2). Here $\mathbf{y}$ is an n-dimensinal vector, $\mathbf{h}$ is a vector-valued function of attitude quaternion q, and $\mathbf{W}$ is an nxn-matrix.

2.1.2. DCM Formulation

B.1. General Type Formulation

$$\text{minimize} \quad f(\mathbf{A}) \qquad (4)$$

subject to

$$\mathbf{A} \text{ (3x3-matrix)} \in SO(3) \text{ or } \mathbf{A}^T \mathbf{A} = \mathbf{I} \qquad (5)$$

B.2. LSQ Type Formulation

minimize

$$f(\mathbf{A}) = (\mathbf{y} - \mathbf{h}(\mathbf{A}))^T \mathbf{W}(\mathbf{y} - \mathbf{h}(\mathbf{A})) \qquad (6)$$

subject to (5). Here $\mathbf{y}$ is an n-dimensinal vector, $\mathbf{h}$ is a vector-valued function of attitude matrix $\mathbf{A}$, and $\mathbf{W}$ is an nxn-matrix.

2.2. Attitude Determination Problem

The general iterative solutions to the above formulated problems are then applied to two attitude determination problems which are also formulated in various forms.

2.2.1. Vector Observation Problem (Wahba's Problem[3])

C.1. DCM LSQ Type Formulation

$$\text{minimize} \quad f(\mathbf{A}) = \sum_{i=1}^{n} a_i \left| \mathbf{w}_i - \mathbf{A}\mathbf{v}_i \right|^2 \qquad (7)$$

where a_i are positive constants, $\mathbf{v}_i$ are known unit vectors in the reference space (inertial frame for many applications) and $\mathbf{w}_i$ corresponding to obeservations are also known unit vectors in the spacecraft body frame.

This problem can be treated as this form, but it can be still rewritten in various forms. The first is to convert it to a matrix form;

C.2. DCM General Type Formulation[7]

$$\text{minimize} \quad f(\mathbf{A}) = -tr(\mathbf{A}\mathbf{B}^T) \qquad (8)$$

where

$$\mathbf{B} \equiv \sum_{i=1}^{n} a_i \mathbf{w}_i \mathbf{v}_i^T \qquad (9)$$

C.3. Quaternion LSQ Type Formulation

The attitude matrix $\mathbf{A}$ is related to the quaternion q by [1,2]

$$\mathbf{A} = (q_4^2 - \mathbf{q}^2)\mathbf{I}_{3\times 3} + 2\mathbf{q}\mathbf{q}^T + 2q_4[\mathbf{q}] \qquad (10)$$

where $q \equiv (q_1, q_2, q_3, q_4)^T \equiv (\mathbf{q}^T, q_4)^T$, $\mathbf{q}$ is a three-dimensional vector, and $[\,]$ is a notation to convert a three-dimensional vector to a skew symmetric matrix. Any function $f(\mathbf{A})$ can be regarded as a function of quaternion as $f(\mathbf{A}(q))$. Then Wahba's problem is also regarded as a quaternion LSQ type problem

$$\text{minimize} \quad f(q) = \sum_{i=1}^{n} a_i \left| \mathbf{w}_i - \mathbf{A}(q)\mathbf{v}_i \right|^2 \qquad (11)$$

C.4. Quaternion General Type Formulation

It is well known that Wahba's problem is also rewritten as[1,5]

$$\text{minimize} \quad f(q) = -q^T \mathbf{K} q \tag{12}$$

where $\mathbf{K}$ is a 4x4-matrix defined as follows;

$$\mathbf{K} \equiv \begin{pmatrix} \mathbf{S} - \sigma \mathbf{I} & \mathbf{Z} \\ \mathbf{Z}^T & \sigma \end{pmatrix} \tag{13}$$

$$\sigma \equiv tr\mathbf{B} = \sum_{i=1}^{n} a_i \mathbf{w}_i{}^T \mathbf{v}_i \tag{14}$$

$$\mathbf{S} \equiv \mathbf{B} + \mathbf{B}^T = \sum_{i=1}^{n} a_i (\mathbf{w}_i \mathbf{v}_i{}^T + \mathbf{v}_i \mathbf{w}_i{}^T) \tag{15}$$

$$\mathbf{Z} \equiv [\mathbf{B} - \mathbf{B}^T] = \sum_{i=1}^{n} a_i \mathbf{w}_i \times \mathbf{v}_i \tag{16}$$

In (16), tr means the trace operation for matrix, and the notation [] means a conversion from a skew matrix to a corresponding three-dimensional vector. This bracket will be introduced in detail later (see 6.1).

2.2.2. Scalar Observation Problem[8,9,10]

The general form for the scalar observation problem is;

D.1. DCM LSQ Type Formulation

$$\text{minimize} \quad f(\mathbf{A}) = \sum_{i=1}^{n} a_i (y_i - \mathbf{w}_i{}^T \mathbf{A} \mathbf{v}_i)^2 \tag{17}$$

In these two problems, explanations on a_i, $\mathbf{v}_i$, $\mathbf{w}_i$ are the same as in (7) and y_i are measrured data. The problem (17) is more general than Wahba's problem.

D.2. Quaternion LSQ Type Formulation

This can be also regarded as a quaternion problem. With the relation (10) we have a problem

$$\text{minimize } f(q) = \sum_{i=1}^{n} a_i (y_i - \mathbf{w}_i{}^T \mathbf{A}(q) \mathbf{v}_i)^2 \tag{18}$$

In the recent GPS attitude determination problem[8,9] all combinations of $\mathbf{v}_i$ and $\mathbf{w}_i$ appear in (17) and for this case we have slightly different problems

E.1. DCM LSQ Type Formulation (GPS case)

$$\text{minimize } f(\mathbf{A}) = \sum_{i=1}^{m} \sum_{j=1}^{n} (y_{ij} - \mathbf{w}_i{}^T \mathbf{A} \mathbf{v}_j)^2 \tag{19}$$

Here we assumed that all a_{ij} are independent of the indeces and omitted in the right-hand side.

E.2. DCM General Type Formulation (GPS case)

$$\text{minimize } f(\mathbf{A}) = tr(-\mathbf{A}\mathbf{B}^T + \mathbf{A}\mathbf{V}\mathbf{A}^T\mathbf{W}) \tag{20}$$

where the matrices in the right-hand side are;

$$\mathbf{B} \equiv 2\sum_{i=1}^{m} \mathbf{w}_i \sum_{j=1}^{n} y_{ij} \mathbf{v}_j{}^T \tag{21}$$

$$\mathbf{V} \equiv \sum_{j=1}^{n} \mathbf{v}_j \mathbf{v}_j{}^T \tag{22}$$

$$\mathbf{W} \equiv \sum_{i=1}^{m} \mathbf{w}_i \mathbf{w}_i{}^T \tag{23}$$

We do not formulate this case using quaternion, since no more efficient solution than the solution to (18) was obtained this time.

3. Gray's Method

From the previous works we review Gray's method[10]. We rewrite the DCM LSQ type problem B.2 or (6) as

$$\text{minimize} \quad \mathbf{h}(\mathbf{A})^T \mathbf{W} \mathbf{h}(\mathbf{A}) \tag{6)'$$

The attitude matrix $\mathbf{A}$ is locally parametrized by a three-dimensional vector Θ as;

$$\mathbf{A} = \mathbf{A}_0 e^{-[\Theta]} \tag{24}$$

where $\mathbf{A}_0$ is an approximate solution. (Notations have been blended with those of this paper.) Then any function of attitude is considered as defined in the three-dimensional space R^3. Utilizing an approximation of $e^{-[\Theta]}$ as

$$e^{-[\Theta]} \approx \mathbf{I} - [\Theta] \tag{25}$$

any attitude function can be differentiated with respect to Θ. Those derivatives are three-dimensional vectors and the Jacobian $\mathbf{J} \in R^{n\times 3}$ is calculated at $\mathbf{A}_0$ as

$$\mathbf{J} \equiv \left(\frac{\partial h_i(\mathbf{A})}{\partial \Theta} \right) \tag{26}$$

Then the Newton-Gauss update scheme becomes the pair of (24) and

$$\Theta = -(\mathbf{J}^T \mathbf{W} \mathbf{J})^{-1} \mathbf{J}^T \mathbf{W} \mathbf{h}(\mathbf{A}_0) \tag{27}$$

This method was applied to both Wahbas's problem (7) and the scalar observation problem (17). In Wahba's problem $\mathbf{h}(\mathbf{A})$ is 3n-dimensional and is a set of $\mathbf{h}_1(\mathbf{A}) = \mathbf{A}\mathbf{v}_1$, …, $\mathbf{h}_n(\mathbf{A}) = \mathbf{A}\mathbf{v}_n$. For function $\mathbf{h}_i(\mathbf{A}) = \mathbf{A}\mathbf{v}_i$ its derivative with respect to ξ was calculated as a quite simple form of

$$\frac{\partial(\mathbf{A}\mathbf{v}_i)}{\partial \Theta} = \mathbf{A}[\mathbf{v}_1] \tag{28}$$

and the Jacobian $\mathbf{J}$ was derived in the form of

$$\mathbf{J} = \begin{pmatrix} \mathbf{A}[\mathbf{v}_1] \\ \cdots \\ \mathbf{A}[\mathbf{v}_n] \end{pmatrix} \tag{29}$$

Also in the scalar observation problem D.1 or (17) the function $h_i(\mathbf{A}) = \mathbf{w}_i{}^T \mathbf{A}\mathbf{v}_i$ was differentiated as

$$\frac{\partial}{\partial \Theta} \mathbf{w}_i{}^T \mathbf{A}\mathbf{v}_i = -\mathbf{w}_i{}^T \mathbf{A}[\mathbf{v}_1] \tag{30}$$

and the Jacobian $\mathbf{J}$ was derived in the form of

$$\mathbf{J} = \begin{pmatrix} -\mathbf{w}_1{}^T \mathbf{A}[\mathbf{v}_1] \\ \cdots \\ -\mathbf{w}_n{}^T \mathbf{A}[\mathbf{v}_n] \end{pmatrix} \tag{31}$$

Notations were largely altered here, and therefore for citing purpose referring the original paper is recommended.

4. General Solutions to Quaternion Formulation[13]

In this section we deal with attitude function's minimization problems with quaternion formulation. While the method of Yoshikawa et al.[11] used the algebraic nature of the quaternion very effectively, our method does not do so. We use the quaternion only as a four-dimensional vector but with a constraint of unit length. This can be said that the totality of the attitude quaternion is just the unit sphere S^3 as a smooth manifold in the Euclidean space R^4 .

4.1. Solution to General Minimization Problem

From the well-known Kuhn-Tucker's condition[12] any solution to A.1 or (1) and (2) is a stationary point of the following Lagrangian function $f_\lambda(q)$ with a Lagrange multiplier λ:

$$f_\lambda(q) \equiv f(q) + \lambda(q^T q - 1) \tag{32}$$

or it satisfies

$$\frac{\partial f_\lambda(q)}{\partial q} = \frac{\partial f(q)}{\partial q} + 2\lambda q = 0$$

Applying the unity condition (2) we have

$$2\lambda = -q^T \frac{\partial f(q)}{\partial q}$$

Deleting λ, we have a necessary condition which any soluton must satisfy as

$$(\mathbf{I}_{4\times 4} - qq^T) \frac{\partial f(q)}{\partial q} = 0 \tag{33}$$

Since the function f is not specified, no explicit algebraic solution is applicable. According to Newton's method, solution is sought near an approximation solution q_0 as

$$q = q_0 + \delta q \tag{34}$$

where δq is a sufficiently small vector. Hereafter we write equations within the equivalency of $o(|\delta q|)$.

Then the constraint (2) is equivalently rewritten as

$$q_0{}^T \delta q = 0 \qquad (35)$$

And this is still equivalent to

$$\mathbf{I}_{q0} \delta q = \delta q \qquad (36)$$

$$\mathbf{I}_{q0} \equiv \mathbf{I}_{4\times 4} - q_0 q_0{}^T \qquad (37)$$

The matrix $\mathbf{I}_{q0}$ works as a projection to the tangential plane to S^3 at q_0.

Substituting (34) into (33), we have

$$\left\{ \mathbf{I} - (q_0 + \delta q)(q_0 + \delta q)^T \right\} \left\{ \frac{\partial f}{\partial q} + \frac{\partial^2 f}{\partial q^2} \delta q \right\} = 0 \qquad (38)$$

This is rearranged as

$$\mathbf{H} \delta q = -\mathbf{I}_{q0} \frac{\partial f}{\partial q}(q_0) \qquad (39)$$

$$\underset{(4x4)}{\mathbf{H}} \equiv \mathbf{I}_{q0} \frac{\partial^2 f}{\partial q^2}(q_0) - \underset{(4x1)}{q_0} \underset{(1x4)}{\left(\frac{\partial f}{\partial q}(q_0)\right)^T}$$

$$- q_0{}^T \frac{\partial f}{\partial q}(q_0) \mathbf{I}_{q0} \qquad (40)$$

We see that the second term of the right-hand side becomes a higher-order term when multiplied by δq and therefore it can be dropped in (40). Also taking (36) into account brings (39) and (40) to the following simple expressions.

$$\mathbf{H}^* \delta q = -\mathbf{I}_{q0} \frac{\partial f}{\partial q}(q_0) \qquad (41)$$

$$\mathbf{H}^* \equiv \mathbf{I}_{q0} \frac{\partial^2 f}{\partial q^2}(q_0) \mathbf{I}_{q0} - q_0{}^T \frac{\partial f}{\partial q}(q_0) \mathbf{I}_{q0} \qquad (42)$$

Eq. (41) can be regarded as a linear equation in a three-dimensional hyper plane (35), which is hereafter denoted as π. The 4x4-matrix $\mathbf{H}^*$ is not singular if its domain is restricted to π. Then the inverse $\mathbf{H}^{*+}$ of $\mathbf{H}^*$ in π is defined as a 4x4-projection matrix on the hyper plane π which satisfies

$$\mathbf{H}^{*+} \mathbf{H}^* = \mathbf{H}^* \mathbf{H}^{*+} = \mathbf{I}_{q0} \qquad (43)$$

One calculation procedure of $\mathbf{H}^{*+}$ can be written as

$$\mathbf{H}^{*+} = (\mathbf{H}^* + \varepsilon q_0 q_0{}^T)^{-1} - \varepsilon^{-1} q_0 q_0{}^T \qquad (44)$$

where any non-zero number, e.g., unity, is applicable to ε. Using this inverse matrix we obtain a solution formula as;

$$\delta q = -\mathbf{H}^{*+} \frac{\partial f}{\partial q}(q_0) \qquad (45)$$

Substituting (42) here yields the solution formula as

$$\delta q = -\left\{ \mathbf{I}_{q0} \frac{\partial^2 f}{\partial q^2}(q_0) \mathbf{I}_{q0} - q_0{}^T \frac{\partial f}{\partial q}(q_0) \mathbf{I}_{q0} \right\}^+ \cdot \frac{\partial f}{\partial q}(q_0) \qquad (46)$$

4.2. Solution to LSQ Type Problem

While the previous solution needed derivatives of the second-order or the Hessian matrix, the solution to this case needs only those of the first-order. If the function $\mathbf{h}(q)$ in (3) can be appropriately linearized with respect to $\delta q = q - q_0$, then the whole function is rearranged as quadratic and an iterative solution will be obtained (the Newton-Gauss method).

Firstly we deal with a linearization of a function $h(q):\ S^3 \to R^1$ and introduce a notation of the first-order derivative as $\nabla_S h(q)$. In this section we define it as a row vector or 1x4-matrix which satisfies both

$$h(q_0 + \delta q) = h(q_0) + \nabla_S h(q_0) \delta q + o(|\delta q|) \qquad (47)$$

and

$$\nabla_S h(q_0)\mathbf{I}_{q0} = \nabla_S h(q_0) \quad (48)$$

where $\mathbf{I}_{q0}$ is the same matrix as defined by (37). Eq. (48) requires that $\nabla_S h(q_0)$ is tangential to S^3 at q_0. This gradient is calculated from the conventional derivative $\partial h / \partial q$ (this is also treated as a row vector) as

$$\nabla_S h(q_0) = \frac{\partial h}{\partial q}(q_0)\mathbf{I}_{q0} \quad (49)$$

For a vector function $\mathbf{h}(q): S^3 \to R^n$, its gradient $\nabla_S \mathbf{h}(q)$ is defined as an nx4-matrix by

$$\nabla_S \mathbf{h}(q) \equiv \begin{pmatrix} \nabla_S h_1(q) \\ \cdots \\ \nabla_S h_n(q) \end{pmatrix} \quad (50)$$

Then we have

$$\mathbf{h}(q_0 + \delta q) = \mathbf{h}(q_0) + \nabla_S \mathbf{h}(q_0)\delta q + o(|\delta q|) \quad (51)$$

Substituting Eq. (51) into (3) yields

$$f(q) = \sum_{i=1}^{n}(\mathbf{y} - \mathbf{h}(q_0))^T \mathbf{W}(\mathbf{y} - \mathbf{h}(q_0))$$
$$-2\sum_{i=1}^{n}(\mathbf{y} - \mathbf{h}(q_0))\mathbf{W}\nabla_S \mathbf{h}(q_0)\delta q$$
$$+\delta q^T \left\{\sum_{i=1}^{n} \nabla_S \mathbf{h}(q_0)^T \mathbf{W}\nabla_S \mathbf{h}(q_0)\right\}\delta q$$
$$+ residual \quad (52)$$

The residual in (52) can be neglected to derive iterative solution to minimize $f(q)$ as

$$\delta q = \left\{\sum_{i=1}^{n} \nabla_S \mathbf{h}(q_0)^T \mathbf{W}\nabla_S \mathbf{h}(q_0)\right\}^+$$
$$\cdot \sum_{i=1}^{n} \nabla_S \mathbf{h}(q_0)^T \mathbf{W}(\mathbf{y} - \mathbf{h}(q_0)) \quad (53)$$

where the notation + is used to mean the same inverse as $\mathbf{H}^{*+}$ in Eq. (43).

4.3. Quaternion Update

When δq is obtained as an increment vector in the tangential plane π and substituted into (34) to obtain a new quaternion q, it will violate the condition (2). Therefore a normalization step is to be taken at every update. The following conventional normalization is applicable with a small load.

$$q = |q_0 + \delta q|^{-1}(q_0 + \delta q) \quad (54)$$

5. Attitude Determination in Quatrernion Formulation

The results of the previous section with quaternion are immediately applicable to the attitude determination problems. Since general formulae are already derived, the task here is quite straightforward.

5.1. Solution to Vector Observation Problem

5.1.1. Application of General Type Formula (46)

The problem C.4 or (12) is treated with the iterative formula (46), to which we can substitute

$$\frac{\partial f}{\partial q} = -2\mathbf{K}q \quad (55)$$

$$\frac{\partial^2 f}{\partial q^2} = -2\mathbf{K} \quad (56)$$

Then an iterative solution is obtained as

$$\delta q = -\left\{\mathbf{I}_{q0}\mathbf{K}\mathbf{I}_{q0} - (q_0{}^T\mathbf{K}q_0)\mathbf{I}_{q0}\right\}^+ \mathbf{K}q_0 \quad (57)$$

5.1.2. Application of LSQ Type Formula (53)

The problem C.3 or (11) is rearranged to (3) with the following correspondence

$$\mathbf{y} \equiv (\mathbf{w}_1{}^T, \ldots, \mathbf{w}_n{}^T)^T \quad (58)$$

$$\mathbf{h}(q) \equiv \left((\mathbf{Av}_1)^T, \ldots, (\mathbf{Av}_n)^T \right)^T \qquad (59)$$

$$\mathbf{W} \equiv diag(a_1 \mathbf{I}_{3\times3}, \ldots, a_n \mathbf{I}_{3\times3}) \qquad (60)$$

Of course $\mathbf{A}$ must be replaced by (10) for quaternion representation. Then in order to apply the iterative formula (53), we need to calculate $\nabla_S \mathbf{Av}_i$. Using (10) we have an expression for $\mathbf{Av}_i$ as

$$\mathbf{Av}_i = (q_4{}^2 - \mathbf{q}^2)\mathbf{v}_i + 2\mathbf{q}(\mathbf{q}^T \mathbf{v}_i) + 2q_4[\mathbf{q}]\mathbf{v}_i \qquad (61)$$

The conventional gradient $\nabla \mathbf{Av}_i$ with respect to q is calculated as

$$\underset{(3x4}{\nabla \mathbf{Av}_i} = \Big(\underset{(3x3)}{\mathbf{qv}_i{}^T - \mathbf{v}_i \mathbf{q}^T + \mathbf{q}^T \mathbf{v}_i \mathbf{I} - q_4 [\mathbf{v}_i]}, \quad \underset{(3x1)}{\mathbf{v}_i q_4 + [\mathbf{v}_i]\mathbf{q}} \Big) \qquad (62)$$

Then we have $\nabla_S \mathbf{Av}_i$ as

$$\nabla_S \mathbf{Av}_i = \nabla \mathbf{Av}_i \mathbf{I}_q \qquad (63)$$

All the necessary quantities are derived and the iterative formula (53) is now applicable.

5.2. Solution to Scalar Observation Problem

For this case we apply the LSQ type formula only. We have expression

$$\mathbf{w}_i{}^T \mathbf{Av}_i = (q_4{}^2 - \mathbf{q}^2)\mathbf{w}_i{}^T \mathbf{v}_i + 2(\mathbf{w}_i{}^T \mathbf{q})(\mathbf{q}^T \mathbf{v}_i) + 2q_4 \mathbf{w}_i{}^T [\mathbf{q}]\mathbf{v}_i \qquad (64)$$

Using (62) and (63) we can calculate

$$\nabla_S \mathbf{w}_i{}^T \mathbf{Av}_i = \mathbf{w}_i{}^T \nabla_S \mathbf{Av}_i \qquad (65)$$

Then the iterative solution formula (53) is applicable.

6. General Solution to DCM Formulation

The direction cosine matrix (DCM) is the most direct attitude representation and in this section we deal with minimization problems expressed by DCM. The basic idea is to use a local representation of attitude matrix by a three-dimensional vector around any fixed attitude. The procedure is the same as that adopted by Gray[10], while the development here derives an explicit expression for the gradient of any function with respect to the local variables.

We have already used the bracket operator [] to the mutual conversion between a three-dimensional vector and a skew symmetric matrix. Although the usual usage is to place a vector inside the bracket, we use this bracket also to the reverse conversion like $[\mathbf{X}]$ where $\mathbf{X}$ is a 3x3 skew symmetric matrix. Some natures of this bracket which are used in this paper are briefly summarized.

6.1. The bracket operator []

For a three-dimensional vector $\mathbf{u} = (u_1, u_2, u_3)^T$, $[\mathbf{u}]$ is defined as a skew symmetric matrix[1,2]

$$[\mathbf{u}] \equiv \begin{pmatrix} 0 & u_3 & -u_2 \\ -u_3 & 0 & u_1 \\ u_2 & -u_1 & 0 \end{pmatrix} \qquad (66)$$

With Levi-Civita symbol ε_{ijk} this is expressed by[2]

$$[\mathbf{u}]_{ij} = \sum_{k=1}^{3} \varepsilon_{ijk} u_k \qquad (67)$$

This matrix $[\mathbf{u}]$ works as a cross-product operator as

$$[\mathbf{u}]\mathbf{v} = -\mathbf{u} \times \mathbf{v} \qquad (68)$$

where $\mathbf{v}$ is any three-dimensional vector.

Using the same bracket we write $[\mathbf{X}]$ for a 3x3 skew-symmetric matrix $\mathbf{X}$ to denote a three-dimensional vector

$$[\mathbf{X}] \equiv \begin{pmatrix} X_{23} \\ X_{31} \\ X_{12} \end{pmatrix} \qquad (68)$$

which is expressed by using ε_{ijk} as

$$[\mathbf{X}]_i \equiv \frac{1}{2}\sum_{j,k=1}^{3}\varepsilon_{ijk}X_{jk} \quad (69)$$

This bracket satisfies the following natures and will be freely used hereafter

$$\big[[\mathbf{u}]\big] = \mathbf{u} \quad (70)$$

$$\big[[\mathbf{X}]\big] = \mathbf{X} \quad (71)$$

$$[\mathbf{u}\times\mathbf{v}] = \mathbf{u}\mathbf{v}^T - \mathbf{v}\mathbf{u}^T \quad (72)$$

$$\left[\mathbf{u}\mathbf{v}^T - \mathbf{v}\mathbf{u}^T\right] = \mathbf{u}\times\mathbf{v} \quad (72)'$$

$$tr\left\{\mathbf{C}\mathbf{X}^T\right\} = [\mathbf{C}-\mathbf{C}^T]^T[\mathbf{X}] \quad (73)$$

$$[\mathbf{X}\mathbf{C}-\mathbf{C}^T\mathbf{X}^T] = \left\{tr(\mathbf{C})\mathbf{I}_{3\times 3}-\mathbf{C}\right\}[\mathbf{X}] \quad (74)$$

Here $\mathbf{u}$ and $\mathbf{v}$ are three-dimensional vectors, $\mathbf{X}$ is a 3x3 skew symmetric matrix, and $\mathbf{C}$ is any 3x3-matrix. The relations (73) and (74) may be new with this bracket, but their derivation is directly possible from the definition and omitted here.

6.2. Solution to General Minimization Problem

We treat the general minimization problem expressed by the DCM, B.1 or (4). Rewrite (24)

$$\mathbf{A} = e^{[\xi]}\mathbf{A}_0 \quad (24)$$

where ξ is a three-dimensional vector. Then any attitude function can be considered as a function of ξ as $f(e^{[\xi]}\mathbf{A}_0)$. To $\mathbf{A}_0$ corresponds $\xi = \mathbf{0}$. Also it is necessary to note that this representation is not global and depends on the fixed attitude $\mathbf{A}_0$. If the function $f(\mathbf{A})$ has a local minimum at $\mathbf{A}_0$, then the derivative with respect to ξ must vanish there or

$$\left.\frac{\partial}{\partial\xi}f(e^{[\xi]}\mathbf{A}_0)\right|_{\xi=\mathbf{0}} = \mathbf{0} \quad (75)$$

Considering no constraint on $\mathbf{A}\equiv(A_{ij})$ (3x3-matrix), we define the derivative of any function $f(\mathbf{A})$ with respect to $\mathbf{A}$ as a 3x3-matrix

$$\frac{\partial f(\mathbf{A})}{\partial\mathbf{A}} \equiv \left(\frac{\partial f(\mathbf{A})}{\partial A_{ij}}\right) \quad (76)$$

Using this expression we can expand $f(\mathbf{A}_0+\delta\mathbf{A})$ as

$$f(\mathbf{A}_0+\delta\mathbf{A}) = f(\mathbf{A}_0)+\sum_{i,j=1}^{3}\frac{\partial f}{\partial A_{ij}}(\mathbf{A}_0)\delta A_{ij} + o(|\delta\mathbf{A}|) = f(\mathbf{A}_0)+tr\left\{\frac{\partial f}{\partial\mathbf{A}}(\mathbf{A}_0)\delta\mathbf{A}^T\right\} + o(|\delta\mathbf{A}|) \quad (77)$$

Substituting an evaluation

$$\mathbf{A}_0+\delta\mathbf{A}\equiv e^{[\xi]}\mathbf{A}_0 = \mathbf{A}_0+[\xi]\mathbf{A}_0+o(|\xi|) \quad (78)$$

into (77) yields

$$f(e^{[\xi]}\mathbf{A}_0) = f(\mathbf{A}_0) + tr\left\{\frac{\partial f}{\partial\mathbf{A}}(\mathbf{A}_0)[\xi]^T\right\}+o(|\xi|) \quad (79)$$

Here utilizing the formula (73), we can rewrite (79) as

$$f(e^{[\xi]}\mathbf{A}_0) = f(\mathbf{A}_0) + \left[\frac{\partial f}{\partial\mathbf{A}}(\mathbf{A}_0)\mathbf{A}_0{}^T - \mathbf{A}_0\left(\frac{\partial f}{\partial\mathbf{A}}(\mathbf{A}_0)\right)^T\right]\xi + o(|\xi|) \quad (80)$$

Then we have for the left-hand side of (75)

$$\left.\frac{\partial}{\partial \xi} f(e^{[\xi]}\mathbf{A}_0)\right|_{\xi=\mathbf{0}}$$

$$= \left[\frac{\partial f}{\partial \mathbf{A}}(\mathbf{A}_0)\mathbf{A}_0{}^T - \mathbf{A}_0\left(\frac{\partial f}{\partial \mathbf{A}}(\mathbf{A}_0)\right)^T\right] \quad (81)$$

Since the attitude $\mathbf{A}_0$ is arbitrary, dropping the index 0 is possible and it yields a general expression

$$\frac{\partial f}{\partial \xi}(\mathbf{A}) = \left[\frac{\partial f}{\partial \mathbf{A}}\mathbf{A}^T - \mathbf{A}\left(\frac{\partial f}{\partial \mathbf{A}}\right)^T\right] \quad (82)$$

This formula is so general to enable us to calculate the gradient vector with respect to ξ only from the original function itself. When we conduct an attitude analysis on ground, the program memory size is not a large concern and this kind of general formula will be very useful. Then finally we have reached a necessary condition which any local minimum must satisfy

$$\frac{\partial f}{\partial \mathbf{A}}\mathbf{A}^T - \mathbf{A}\left(\frac{\partial f}{\partial \mathbf{A}}\right)^T = 0 \quad (83)$$

The condition (83) comprises only three independent relations from its symmetry. Together with the condition that $\mathbf{A}$ is a matrix in $SO(3)$ or must satisfy (5), all nine elements of $\mathbf{A}$ are determined.

It is necessary to adopt some technique in applying the Newton-Gauss method to derive a simple iterative formula. It is because we did not succeed to obtain a simple expression for the second-order derivatives or the Hessian matrix $\partial^2 f / \partial \mathbf{A}^2$ which intrinsically contains 81 elements. We here adopt two approximation methods.

Case that $\partial f / \partial \mathbf{A}$ is nearly constant

We assume that $\partial f / \partial \mathbf{A}$ is constant or almost constant. For example, assume that $f(\mathbf{A}) = tr\mathbf{A}\mathbf{B}^T$, where $\mathbf{B}$ is constant. Then $\partial f / \partial \mathbf{A} = \mathbf{B}$, and the assumption is met. Also assume that $f(\mathbf{A}) = tr\{\mathbf{A}\mathbf{B}^T + \varepsilon\mathbf{A}^2\mathbf{B}\}$. Then $\partial f / \partial \mathbf{A} = \mathbf{B} + \varepsilon(\mathbf{B}\mathbf{A}^T + \mathbf{A}^T\mathbf{B})$ is almost constant, if ε is small. However, in case of $f(\mathbf{A}) = tr\mathbf{A}^2\mathbf{B}^T$, $\partial f / \partial \mathbf{A} = \mathbf{B}\mathbf{A}^T + \mathbf{A}^T\mathbf{B}$, the derivative can not be regarded as nearly constant.

Around an approximate point $\mathbf{A}_0$, we linearize (83) as

$$\frac{\partial f}{\partial \mathbf{A}}(\mathbf{A}_0)\left(\mathbf{A}_0 + [\xi]\mathbf{A}_0\right)^T - \left(\mathbf{A}_0 + [\xi]\mathbf{A}_0\right)\left(\frac{\partial f}{\partial \mathbf{A}}(\mathbf{A}_0)\right)^T \approx 0 \quad (84)$$

where no second-order derivatives are kept under the present assumption. Writing

$$\mathbf{C} \equiv \frac{\partial f}{\partial \mathbf{A}}(\mathbf{A}_0) \quad (85)$$

yields

$$[\xi]\mathbf{A}_0\mathbf{C}^T - \mathbf{C}\mathbf{A}_0{}^T[\xi]^T \approx \left(\mathbf{C}\mathbf{A}_0{}^T - \mathbf{A}_0\mathbf{C}^T\right) \quad (86)$$

Applying (74) to the left-hand side we have

$$\left\{tr(\mathbf{A}_0\mathbf{C}^T)\mathbf{I} - \mathbf{A}_0\mathbf{C}^T\right\}\xi \approx \left[\mathbf{C}\mathbf{A}_0{}^T - \mathbf{A}_0\mathbf{C}^T\right]$$

which is solved with respect to ξ as an iterative update formula

$$\xi = \left\{tr(\mathbf{A}_0\mathbf{C}^T)\mathbf{I} - \mathbf{A}_0\mathbf{C}^T\right\}^{-1}[\mathbf{C}\mathbf{A}_0{}^T - \mathbf{A}_0\mathbf{C}^T] \quad (87)$$

This iterative algorithm depends strongly on the above assumption, and it must be noted that a general convergence condition is not guaranteed.

Method to calculate $\partial^2 f / \partial \xi^2$ numerically

Since an explicit expression for $\partial f / \partial \xi$ is available by (82), there is a way to calculate $\partial^2 f / \partial \xi^2$ numerically. With an appropriate small number ε we can numerically calculate $\partial^2 f / \partial \xi^2$ by using two discrete evaluations for $i = 1, \ldots, 3$ as

$$\frac{\partial}{\partial \xi_i}\frac{\partial f}{\partial \xi} \approx \varepsilon^{-1}\left\{\frac{\partial f}{\partial \xi}\left(e^{[\varepsilon \mathbf{e}_i]}\mathbf{A}_0\right) - \frac{\partial f}{\partial \xi}(\mathbf{A}_0)\right\} \quad (88)$$

Here $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are three orthogonal unit vectors and ε is an appropriate non-zero number. As an example 10^{-5} is applicable to ε. Then we can apply Newton's method and obtain an update scheme

$$\xi = -\left(\frac{\partial^2 f}{\partial \xi^2}\right)^{-1}\frac{\partial f}{\partial \xi} \quad (89)$$

It should be noted that the above two methods will give an exact solution, if we have a convergence. This is because we are using an accurate expression (82) for $\partial f / \partial \xi$.

6.3. Solution to LSQ-Type Minimization Problem

The minimization problem of the least squares formulation B.2 or (6) is treated. In this case for a function $h(\mathbf{A}): S^3 \to R^1$, we define the derivative of the first-order derivative $\partial h / \partial \xi$ as a row vector, and from (82) it has an expression of

$$\frac{\partial h}{\partial \xi}(\mathbf{A}) = \underset{(1\text{x}3)}{\left[\frac{\partial h}{\partial \mathbf{A}}\mathbf{A}^T - \mathbf{A}\left(\frac{\partial h}{\partial \mathbf{A}}\right)^T\right]} \quad (90)$$

For a vector observation function $\mathbf{h}(\mathbf{A}) \equiv \left(h_1(\mathbf{A}), \ldots, h_n(\mathbf{A})\right)^T$, we define the gradient as

$$\mathbf{C} \equiv \frac{\partial \mathbf{h}}{\partial \xi}(\mathbf{A}) = \underset{(\text{nx}3)}{\begin{pmatrix} \frac{\partial h_1}{\partial \xi}(\mathbf{A}) \\ \ldots. \\ \frac{\partial h_n}{\partial \xi}(\mathbf{A}) \end{pmatrix}} \quad (91)$$

Then applying the Newton-Gauss method, we obtain an update formula for ξ as

$$\xi = (\mathbf{C}^T\mathbf{W}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{W}\left(\mathbf{y} - \mathbf{h}(\mathbf{A}_0)\right) \quad (92)$$

where $\mathbf{C}$ is evaluated at $\mathbf{A} = \mathbf{A}_0$.

6.4. Attitude Update

The attitude update is calculated by (24) or $\mathbf{A} = e^{[\xi]}\mathbf{A}_0$. Then practically it is necessary to calculate the exponential function of a skew symmetric matrix. Since it can be expected that ξ is small, one candidate for the calculation procedure is (25) or $e^{[\xi]} \approx \mathbf{I} + [\xi]$. But this formula violates the requirement that $e^{[\xi]}$ must be an orthogonal matrix or in $SO(3)$.

Therefore it is desirable to find a simple calculation formula for $e^{[\xi]}$ under the condition that the result is still in $SO(3)$.

One is to use a well-known expression which utilize the trigonometric functions[1,2]

$$e^{[\xi]} = \cos\theta\mathbf{I} + (1-\cos\theta)\mathbf{u}\mathbf{u}^T + \sin\theta[\mathbf{u}] \quad (93)$$

where θ and $\mathbf{u}$ are defined by

$$\theta \equiv |\xi| \quad (94)$$

$$\mathbf{u} \equiv \xi / |\xi| \quad (95)$$

When the iteration is repeated, $|\xi|$ will vanish and the calculation (95) will diverge. In such a case the next formula is applicable;

$$e^{[\xi]} \approx \sqrt{1-\xi^2}\,\mathbf{I} + \left(1+\sqrt{1-\xi^2}\right)^{-1}\xi\xi^T + [\xi] \quad (96)$$

Although ξ must not exceed unity, the right-hand side stays for any case in $SO(3)$ and the difference of both the sides is the order of $o\left(|\xi|^2\right)$.

7. Attitude Determination in DCM Formulation

We would like to demonstrate that the resulted general iterative formulae in the previous section are efficiently applicable to the attitude determination problems.

7.1. Solution to Vector Observation Problem

7.1.1. Application of General Type Formula

Wahba's problem is treated in a rewritten form C.2 or (8) and (9)

$$\text{minimize} \qquad f(\mathbf{A}) = -tr(\mathbf{A}\mathbf{B}^T) \qquad (8)$$

where $\mathbf{B}$ is a constant matrix defined by (9). For this case the first-order derivative $\partial f(\mathbf{A})/\partial \mathbf{A}$ becomes constant as

$$\frac{\partial f}{\partial \mathbf{A}}(\mathbf{A}) = -\mathbf{B} \qquad (97)$$

Then the necessary condition for local minimum (83) is expressed as;

$$\mathbf{A}\mathbf{B}^T - \mathbf{B}\mathbf{A}^T = 0 \qquad (98)$$

This condition is very basic for Wahba's problem and was derived by Brock[4,5]. He did an algebraic treatment of this condition to derive a simple calculation formula which we do not further enter.

Now the iterative solution formula (87) is applicable and we obtain

$$\xi = \left\{tr(\mathbf{A}_0\mathbf{B}^T)\mathbf{I} - \mathbf{A}_0\mathbf{B}^T\right\}^{-1}[\mathbf{B}\mathbf{A}_0{}^T - \mathbf{A}_0\mathbf{B}^T] \qquad (99)$$

This result is already very simple, but we try to find simpler expressions.

No Update of Gain Matrix

If the initial attitude matrix $\mathbf{A}_0$ is already near the solution, then the first factor of the right-hand side of (99), $\mathbf{G} \equiv \left\{tr(\mathbf{A}_0\mathbf{B}^{\mathrm{T}})\mathbf{I} - \mathbf{A}_0\mathbf{B}^{\mathrm{T}}\right\}^{-1}$, which we call here the *gain matrix*, will stay almost constant during the iteration. Then it is feasible to stop the calculation of this factor if it is once obtained. This procedure will serve for saving the calculation time in a real application. Program code is sill needed.

Utilization of the Original Vectors

The matrix $\mathbf{B}$ was defined by (9) using the original vectors $\mathbf{v}_i$ and $\mathbf{w}_i$. Returning to these vectors we can write

$$\mathbf{A}_0\mathbf{B}^T = \sum_{i=1}^{n} a_i \mathbf{A}_0 \mathbf{v}_i \mathbf{w}_i{}^T \qquad (100)$$

If the observation is made under a condition of low-level noises, then $\mathbf{A}_0\mathbf{v}_i$ must be near $\mathbf{w}_i$ or $\mathbf{A}_0\mathbf{v}_i \approx \mathbf{w}_i$. Then we have

$$\mathbf{A}_0\mathbf{B}^T \approx \sum_{i=1}^{n} a_i \mathbf{w}_i \mathbf{w}_i{}^T . \qquad (101)$$

$$tr(\mathbf{A}_0\mathbf{B}^T) = \sum_{i=1}^{n} a_i tr(\mathbf{A}_0\mathbf{v}_i\mathbf{w}_i{}^T)$$

$$= \sum_{i=1}^{n} a_i \mathbf{w}_i{}^T \mathbf{A}_0 \mathbf{v}_i \approx \sum_{i=1}^{n} a_i \qquad (102)$$

These two relations state that the Gain matrix $\mathbf{G}$ can be evaluated without using a priori attitude estimate as

$$\mathbf{G} \approx \mathbf{G}_0 \equiv \left\{\left(\sum_{i=1}^{n} a_i\right)\mathbf{I} - \sum_{i=1}^{n} a_i \mathbf{w}_i \mathbf{w}_i{}^T\right\}^{-1} \qquad (103)$$

Using this $\mathbf{G}_0$ we obtain one of the simplest iterative solution formula as

$$\xi = \left\{ \left(\sum_{i=1}^{n} a_i \right) \mathbf{I} - \sum_{i=1}^{n} a_i \mathbf{w}_i \mathbf{w}_i^{\ T} \right\}^{-1}$$

$$\cdot \left[\mathbf{A}_0 \mathbf{B}^T - \mathbf{B} \mathbf{A}_0^{\ T} \right] \qquad (104)$$

In the real application this formula requires the summation and inverse operation only at the beginning to obtain both the gain matrix and the matrix $\mathbf{B}$. After that during the iteration very small amount of matrix calculation is required.

7.1.2. Application of LSQ Type Formula

The iterative solution to Wahba's problem here becomes the same one as that of Gray (see 3). However, we demonstrate that our general formula is automatically applicable to this problem. And also we try to see how simpler the result can be reduced.

For applying the LSQ type solution formula to Wahba's problem, it is necessary to calculate $\partial \mathbf{Av}_i / \partial \xi$.

We demonstrate the derivation of this derivative by our present method. As is needed in (86), we first derive the derivatives of $\mathbf{Av}_i$ with respect to $\mathbf{A}$. $\mathbf{Av}_i$ is expressed as

$$\mathbf{Av}_i = \left(\mathbf{e}_1^{\ T} \mathbf{Av}_i , \mathbf{e}_2^{\ T} \mathbf{Av}_i , \mathbf{e}_3^{\ T} \mathbf{Av}_i \right)^T \qquad (105)$$

where $\mathbf{e}_1 \equiv (1,0,0)^T, \mathbf{e}_2 \equiv (0,1,0)^T, \mathbf{e}_3 \equiv (0,0,1)^T$.

Then we can calculate $\partial \mathbf{e}_k^{\ T} \mathbf{Av}_i / \partial \mathbf{A}$ as;

$$\frac{\partial}{\partial \mathbf{A}} \mathbf{e}_k^{\ T} \mathbf{Av}_i = \frac{\partial}{\partial \mathbf{A}} tr\left(\mathbf{A} (\mathbf{e}_k \mathbf{v}_i^{\ T})^T \right)$$

$$= \mathbf{e}_k \mathbf{v}_i^{\ T} \qquad (106)$$

and then $\partial \mathbf{e}_k^{\ T} \mathbf{Av}_i / \partial \xi$ as

$$\frac{\partial}{\partial \xi} \mathbf{e}_k^{\ T} \mathbf{Av}_i = \left[\mathbf{e}_k (\mathbf{Av}_i)^T - (\mathbf{Av}_i) \mathbf{e}_k^{\ T} \right]$$

$$= \mathbf{e}_k \times \mathbf{Av}_i = -\mathbf{e}_k^{\ T} \underset{(1\text{x}3)}{\left[\mathbf{Av}_i \right]} \qquad (107)$$

The result is expressed as a row vector. Then we have the same expression for $\partial \mathbf{Av}_i / \partial \xi$ as

$$\frac{\partial}{\partial \xi} \mathbf{Av}_i = -\left[\mathbf{Av}_i \right] \qquad (108)$$

Substituting the expression (108) into (91) and (92), we have an update formula

$$\xi = -\left(\sum_{i=1}^{n} a_i \left[\mathbf{A}_0 \mathbf{v}_i \right]^T \left[\mathbf{A}_0 \mathbf{v}_i \right] \right)^{-1}$$

$$\cdot \sum_{i=1}^{n} a_i \left[\mathbf{A}_0 \mathbf{v}_i \right]^T (\mathbf{w}_i - \mathbf{A}_0 \mathbf{v}_i)$$

$$= \left(\sum_{i=1}^{n} a_i \left(\mathbf{I}_{3\times 3} - \mathbf{A}_0 \mathbf{v}_i (\mathbf{A}_0 \mathbf{v}_i)^T \right) \right)^{-1}$$

$$\cdot \sum_{i=1}^{n} a_i \mathbf{w}_i \times \mathbf{A}_0 \mathbf{v}_i \qquad (109)$$

We see that this result has a very similar form as the previous result (104). We can show the two formulae are basically equivalent as follows.

Firstly $\mathbf{A}_0 \mathbf{v}_i$ in the first factor can be replaced by $\mathbf{w}_i$, since these two are near each other. And the second factor can be calculated as

$$\sum_{i=1}^{n} a_i \mathbf{w}_i \times \mathbf{A}_0 \mathbf{v}_i = \sum_{i=1}^{n} a_i \left[[\mathbf{w}_i \times \mathbf{A}_0 \mathbf{v}_i] \right]$$

$$= \sum_{i=1}^{n} a_i \left[\mathbf{w}_i (\mathbf{A}_0 \mathbf{v}_i)^T - (\mathbf{A}_0 \mathbf{v}_i) \mathbf{w}_i^{\ T} \right]$$

$$= \left[\mathbf{A}_0 \mathbf{B}^T - \mathbf{B} \mathbf{A}_0^{\ T} \right] \qquad (110)$$

Then we have again the formula (104). Therefore, one can select the simplest formula from these several variations in any real application considering its specific condition.

7.2. Solution to Scalar Observation Problem

For the scalar observation problem D.1 and E.2, our method is also automatically applicable.

7.2.1. Application of General Type Formula

The attitude determination problem E.2 or (20) is treated. We differentiate (20) first by $\mathbf{A}$ as

$$\frac{\partial}{\partial \mathbf{A}} f(\mathbf{A}) = \frac{\partial}{\partial \mathbf{A}} tr(-\mathbf{A}\mathbf{B}^T + \mathbf{A}\mathbf{V}\mathbf{A}^T\mathbf{W})$$

$$= -\mathbf{B} + 2\mathbf{W}\mathbf{A}\mathbf{V} \qquad (111)$$

Then substituting this into (82) we have $\partial f(\mathbf{A}) / \partial \xi$ as

$$\frac{\partial}{\partial \xi} f(\mathbf{A}) = \left[-\mathbf{B}\mathbf{A}^T + \mathbf{A}\mathbf{B}^T\right]$$

$$+\left[2\mathbf{W}\mathbf{A}\mathbf{V}\mathbf{A}^T - 2\mathbf{A}\mathbf{V}\mathbf{A}^T\mathbf{W}\right] \qquad (112)$$

This expression can be used in the iterative solution formula (89) with numerical differentiation (88) for $\partial^2 f(\mathbf{A}) / \partial \xi^2$.

7.2.2. Application of LSQ Type Formula

The problem D.1 or (17) is treated. What is needed in applying the formula (92) is only to differentiate the function

$$h(\mathbf{A}) \equiv \mathbf{w}_i{}^T \mathbf{A}\mathbf{v}_i \qquad (113)$$

with respect to ξ. The main process was already done to derive (108) and we have

$$\frac{\partial}{\partial \xi} \mathbf{w}_i{}^T \mathbf{A}\mathbf{v}_i = \mathbf{w}_i{}^T \frac{\partial \mathbf{A}\mathbf{v}_i}{\partial \xi} = -\mathbf{w}_i{}^T \left[\mathbf{A}\mathbf{v}_i\right]$$

$$= \underset{(1\text{x}3)}{\left(\mathbf{w}_i \times \mathbf{A}\mathbf{v}_i\right)^T} \qquad (114)$$

where the result is expressed as a row vector.

If we write the whole update formula, it becomes

$$\xi = \left\{\sum_{i=1}^n a_i \left(\mathbf{w}_i \times \mathbf{A}_0\mathbf{v}_i\right)\left(\mathbf{w}_i \times \mathbf{A}_0\mathbf{v}_i\right)^T\right\}^{-1}$$

$$\cdot \sum_{i=1}^n a_i \left(\mathbf{w}_i \times \mathbf{A}_0\mathbf{v}_i\right)\left(y_i - \mathbf{w}_i{}^T \mathbf{A}_0\mathbf{v}_i\right) \qquad (115)$$

8. Comments on Covariance Analysis

We focused our attention on solving minimization problems. If the problem arose from some statistical estimation problem, then a necessity of covariance analysis must accompany it. We briefly comment on it.

The covariance analysis is related to least squares (LSQ) type minimization problem. All the iterative solution to that type had the form

$$\delta q = \mathbf{G}\left(\mathbf{y} - \mathbf{h}(q)\right) \qquad (116)$$

for the quaternion formulation, or

$$\xi = \mathbf{G}\left(\mathbf{y} - \mathbf{h}(\mathbf{A})\right) \qquad (117)$$

for the DCM formulation.

In this form we can regard that the part $\mathbf{y} - \mathbf{h}$ is the observation noise and the left-hand side is the estimation error. Then we have an error propagation formula

$$\delta q \text{ or } \xi = \mathbf{G}\mathbf{n} \qquad (118)$$

where $\mathbf{n}$ is the observation error.

For any covariance analysis, we can assume that the statistical characteristics of the observation error are given as, e.g.,

$$\mathbf{E}\mathbf{n} \equiv \mathbf{0} \qquad (119)$$

$$\mathbf{V}\mathbf{n} \equiv \mathbf{E}\mathbf{n}\mathbf{n}^T = \mathbf{P}_0 \qquad (120)$$

where $\mathbf{E}\mathbf{n}$ and $\mathbf{V}\mathbf{n}$ are the mean value and the covariance of the noise $\mathbf{n}$, respectively.

Then we can calculate the statistical quantities for the estimation error as

$$\mathbf{E}\delta q \text{ or } \mathbf{E}\xi = \mathbf{0} \qquad (121)$$

$$\mathbf{V}\delta q \text{ or } \mathbf{V}\xi = \mathbf{G}\mathbf{P}_0\mathbf{G}^T \qquad (122)$$

If the above procedure is applied to Wahba's problem with one of the present solutions, we can derive the same formula for estimation error covariance as that derived by Shuster et al.[6].

9. Conclusion

Newton's method is very effectively applicable to various minimization problem of attitude functions. If the problem was established in the quaternion formulation or in the direction cosine matrix formulation, then it was shown that an appropriate iterative formula can be chosen.

Also it was shown that those solution formulae are successfully applicable to the existing attitude determination problems and give very efficient calculation procedures.

REFERENCES

[1]Wertz, J. R. (Ed.), *Spacecraft Attitude Determination and Control*, D. Reidel Publ. Co., Dordrecht Holland, 1978.

[2]Shuster, M. D., A Survey of Attitude Representations, *The Journal of the Astronautical Sciences*, Vol. 41, No.4, October-December 1993, pp. 439-517.

[3]Wahba, G., Problem 65-1: A Least Squares Estimate of Spacecraft Attitude, *SIAM Review*, Vol. 7, No.3, July 1965, p. 409.

[4]Brock, J. E., Solution to Wahba's Problem, *SIAM Review*, Vol. 8, No.3, July 1966, p. 386.

[5]Brock, J. E., Optimal Matrices Describing Linear Systems, *AIAA Journal*, Vol. 6, No.7, July 1968, pp. 1292-1296.

[6]Shuster M. D., and Oh, S. D., Three-Axis Attitude Determination from Vector Observations, *Journal of Guidance and Control*, Vol. 4, No.1, January-February, 1981, pp. 70-77.

[7]Markley F. L., Attitude Determination using Vector Observations and the Singular Value Decomposition, *The Journal of the Astronautical Sciences*, Vol. 36, No.3, July-September 1988, pp. 245-258.

[8]Crassidis, J. L., Markley, F. L., New Algorithm for Attitude Determination Using Global Positioning System Signals, *Journal of Guidance, Control and Dynamics*, Vol. 20, No.5, September-October, 1997, pp. 891-896.

[9]Bar-Itzhaack, I. Y., Montgomery P. Y., and Garrick, J. C., Algorithms for Attitude Determination using GPS, *AIAA Paper No. 97-3616, AIAA Guidance and Control Conference*, New Orleans, August 1997.

[10]Gray, C. W., Applications of Newton's Method to Attitude Determination, *AAS Paper No. 92-012*, 1992.

[11]Yoshikawa, S., and Yamada, K., Attitude Estimation Based on Star Sensors, *Proceedings of ISAS 6th Workshop on Astrodynamics and Flight Mechanics*, Sagamihara JAPAN, July 15-16 1996, pp.130-136.

[12]Fletcher, R., Practical Methods of Optimization, Volume 2 Constrained Optimization, John Wiley & Sons, Ltd., 1981.

[13]Horii, M., Application of Reduced-Gradient and Reduced Hessian Matrix to Attitude Determination, *MAS Working Paper No. 281*, European Space Operations Centre, ESA, 1988.