ATTITUDE DETERMINATION USING VECTOR OBSERVATIONS: A FAST OPTIMAL MATRIX ALGORITHM

F. Landis Markley[1]

Abstract

The attitude matrix minimizing Wahba's loss function is computed directly by a method that is competitive with the fastest known algorithm for finding this optimal estimate. The method also provides an estimate of the attitude error covariance matrix. Analysis of the special case of two vector observations identifies those cases for which the TRIAD or algebraic method minimizes Wahba's loss function.

Introduction

In 1965, Wahba posed the problem of finding the proper orthogonal matrix A that minimizes the non-negative loss function [1]

$$L(A) \equiv \tfrac{1}{2} \sum_{i=1}^{n} a_i \, |\mathbf{b}_i - A\mathbf{r}_i|^2,$$ (1)

where the unit vectors $\mathbf{r}_i$ are representations in a reference frame of the directions to some observed objects, the $\mathbf{b}_i$ are the unit vector representations of the corresponding observations in the spacecraft body frame, the a_i are positive weights, and n is the number of observations. The motivation for this loss function is that if the vectors are error-free and the true attitude matrix A_{true} is assumed to be the same for all the measurements, then $\mathbf{b}_i$ is equal to $A_{true}\mathbf{r}_i$ for all i and the loss function is equal to zero for A equal to A_{true}.

Attitude determination algorithms based on minimizing this loss function have been used for many years [2-9]. The original solutions to Wahba's problem solved for the spacecraft attitude matrix directly [2-5], but most practical applications have been based on Davenport's q-method [6-8], which solves for the quaternion representing the attitude matrix. In this paper, we present a new method that solves for the attitude matrix directly, as well as the covariance matrix, and which is competitive with the well known QUEST algorithm [9] in speed. Analysis of the special case of two observations serves to relate this method to the TRIAD or algebraic method [8, 9].

Statement of the problem

Simple matrix manipulations transform the loss function into

$$L(A) = \lambda_0 - \text{tr}(AB^T),$$ (2)

where

$$\lambda_0 \equiv \sum_{i=1}^{n} a_i \,,$$ (3)

$$B \equiv \sum_{i=1}^{n} a_i \, \mathbf{b}_i \mathbf{r}_i^T \,,$$ (4)

[1]Assistant Head, Guidance and Control Branch, Code 712, Goddard Space Flight Center, Greenbelt, MD 20771

tr denotes the trace, and the superscript T denotes the matrix transpose. Thus Wahba's problem is equivalent to the problem of finding the proper orthogonal matrix A that maximizes the trace of the matrix product AB^T. The weights are often chosen so that $\lambda_0 = 1$, but this is not always the most convenient choice, as will be discussed below.

This optimization problem has an interesting relation to a matrix norm. The Euclidean norm (also known as the Schur, Frobenius, or Hilbert-Schmidt norm) is defined for a general real matrix M by [10, 11]

$$\| M \|^2 \equiv \Sigma M_{ij}^2 = \text{tr}(MM^T), \tag{5}$$

where the sum is over all the matrix elements. The assumed orthogonality of A and properties of the trace give

$$\| A - B \|^2 = \text{tr}[(A - B)(A - B)^T] = \text{tr}\,I - 2\text{tr}(AB^T) + \| B \|^2, \tag{6}$$

where I is the 3×3 identity matrix. The orthogonal matrix A that maximizes $\text{tr}(AB^T)$ minimizes this norm, so Wahba's problem is also equivalent to the problem of finding the proper orthogonal matrix A that is closest to B in the Euclidean norm [12].

The matrix B can be shown to have the decomposition [13]

$$B = U_+ \, \text{diag}[S_1, S_2, S_3] V_+{}^T \tag{7}$$

where U_+ and V_+ are proper orthogonal matrices; diag[...] denotes a matrix with the indicated elements on the main diagonal and zeros elsewhere; and S_1, S_2, and $|S_3|$, the singular values of B, obey the inequalities

$$S_1 \geq S_2 \geq |S_3|. \tag{8}$$

The optimal attitude estimate is given in terms of these matrices by [13]

$$A_{opt} = U_+ V_+{}^T. \tag{9}$$

Equation (7) differs from the singular value decomposition (SVD) [10, 11] in that U_+ and V_+ are required to have positive determinant. In reference [13], S_3 was denoted by ds_3, where $d = \pm 1$ and $s_3 \geq 0$.

The SVD provides a robust method for computing the matrices U_+ and V_+, and thus the optimal attitude estimate, but it is not very efficient [13]. The purpose of this paper is to present a more efficient method to estimate the attitude.

Computation of the attitude matrix

Noting that the adjoint of the transpose of B and the product BB^TB can be written as

$$\text{adj}\, B^T = U_+ \, \text{diag}[S_2S_3, S_3S_1, S_1S_2] \, V_+{}^T, \tag{10}$$

and

$$BB^TB = U_+ \, \text{diag}[S_1^3, S_2^3, S_3^3] \, V_+{}^T, \tag{11}$$

it is a matter of simple algebra to see that

$$A_{opt} = [(\kappa + \|B\|^2)B + \lambda \operatorname{adj} B^T - BB^TB]/\zeta, \tag{12}$$

where

$$\|B\|^2 = S_1{}^2 + S_2{}^2 + S_3{}^2, \tag{13}$$

and the other scalar coefficients are defined by

$$\kappa \equiv S_2 S_3 + S_3 S_1 + S_1 S_2, \tag{14}$$

$$\lambda \equiv S_1 + S_2 + S_3, \tag{15}$$

and

$$\zeta \equiv (S_2 + S_3)(S_3 + S_1)(S_1 + S_2). \tag{16}$$

The matrices in equation (12) can be computed without performing the singular value decomposition, but this equation is an improvement over equation (9) only because the scalar coefficients κ, λ, and ζ can also be computed without the SVD, as we will show below.

Iterative computation of the scalar coefficients

We first find expressions for the other scalar coefficients in terms of λ. A little algebra shows that

$$\kappa = \tfrac{1}{2}(\lambda^2 - \|B\|^2) \tag{17}$$

and

$$\zeta = \kappa\lambda - \det B. \tag{18}$$

Let $A(\lambda)$ denote the expression for the attitude matrix given by equations (12), (17), and (18) as a function of λ and B. This is equal to A_{opt} if λ is given by equation (15). Equations (7), (9), and (15) give

$$\lambda = \operatorname{tr}(A_{opt} B^T), \tag{19}$$

so λ can be computed as a solution of the equation

$$\lambda = \operatorname{tr}[A(\lambda)B^T] = \operatorname{tr}[(\kappa + \|B\|^2)BB^T + \lambda (\det B)I - (BB^T)^2]/\zeta. \tag{20}$$

Substitution of equations (17), (18), and the identity

$$\|B\|^4 - \operatorname{tr}[(BB^T)^2] = 2 \|\operatorname{adj} B\|^2 \tag{21}$$

lets us write this as

$$0 = Q(\lambda) \equiv \kappa^2 - 2\lambda \det B - \|\operatorname{adj} B\|^2. \tag{22}$$

Since κ is a quadratic function of λ, $Q(\lambda)$ is a quartic polynomial. It can be shown to be the same quartic that is used in QUEST, up to an irrelevant factor of one-fourth. Substitution of equation (7) into equation (22) gives the four roots of the quartic in terms of S_1, S_2, and S_3. We must use equation (17) for κ rather than equation (14) in this substitution, which gives

$$4Q(\lambda) = (\lambda - S_1 - S_2 - S_3)(\lambda - S_1 + S_2 + S_3)(\lambda + S_1 - S_2 + S_3)(\lambda + S_1 + S_2 - S_3). \tag{23}$$

The roots of this equation are all real, and they are the four eigenvalues of the K matrix in the q-method, as is well known [7, 9]. Equations (8) and (15) show that we require the maximum root, and that this root is distinct unless $S_2 + S_3 = 0$. When $S_2 + S_3 = 0$, the attitude solution is not unique, as is discussed in reference [13]; in the method introduced in this paper, this results in $\zeta = 0$ and all the elements of A_{opt} having the indefinite form $0/0$.

We now note from equations (2) and (19) that

$$L(A_{opt}) = \lambda_0 - \lambda \geq 0. \tag{24}$$

For small measurement errrors, the loss function should be close to zero, so the maximum root of equation (22) should be close to λ_0 [9]. Thus we can find λ by Newton's method, starting with this value. This defines a sequence of estimates of λ by

$$\lambda_i = \lambda_{i-1} - Q(\lambda_{i-1})/Q'(\lambda_{i-1}), \quad i = 1, 2, \ldots. \tag{25}$$

Substitution of equation (23) shows that this sequence would be monotonically decreasing with infinite-precision arithmetic, but a computation with finite-precision arithmetic eventually finds a $\lambda_i \geq \lambda_{i-1}$. At this point, the iterations are terminated and λ_{i-1} is taken to be the desired root to full computer precision. This iteration converges extremely rapidly in practice, except in the case that the maximum root of $Q(\lambda)$ is not unique. In that case the derivative in the denominator of equation (25) goes to zero as the root is approached, so the computation is terminated and a warning is issued that the attitude is indeterminate. Halley's method [14] would give convergence in fewer iterations than Newton's method, but would require more computations per iteration, so it was not investigated further.

It is important to carry out the computation of λ to full machine precision, since otherwise the computed attitude matrix will not be orthogonal. Straightforward matrix computation gives

$$A(\lambda)A^T(\lambda) = I - Q(\lambda)(\lambda^2 I - BB^T)/\zeta^2. \tag{26}$$

This shows the orthogonality of the computed attitude matrix if λ is a root of $Q(\lambda)$, and estimates the departure from orthogonality otherwise.

Analytic computation of the scalar coefficients

The scalar coefficients can also be computed as functions of the largest singular value S_1 of B by

$$\kappa = S_1(S_2 + S_3) + S_2 S_3 = S_1(S_2 + S_3) + S_1^{-1}\det B, \tag{27}$$

$$\lambda = S_1 + (S_2 + S_3), \tag{28}$$

and

$$\zeta = (\kappa + S_1^2)(S_2 + S_3), \tag{29}$$

where

$$S_2 + S_3 = \{S_1^{-2}[\|\operatorname{adj} B\|^2 - (S_1^{-1}\det B)^2] + 2S_1^{-1}\det B\}^{1/2}. \tag{30}$$

This form is chosen to avoid near-cancellations in near-singular cases. The largest singular

540

value is found as the positive square root of the largest root of the cubic characteristic equation of the matrix BB^T [7]:

$$0 = (S_1^2)^3 - \text{tr}(BB^T)(S_1^2)^2 + \text{tr}[\text{adj}(BB^T)]S_1^2 - \det(BB^T)$$

$$= (S_1^2)^3 - \|B\|^2(S_1^2)^2 + \|\text{adj}\, B\|^2 S_1^2 - (\det B)^2. \tag{31}$$

The largest root of this equation is given by [7, 15]

$$S_1^2 = \tfrac{1}{3}\{\|B\|^2 + 2\alpha^{1/2}\cos[\tfrac{1}{3}\cos^{-1}(\alpha^{-3/2}\beta)]\}, \tag{32}$$

where

$$\alpha \equiv \|B\|^4 - 3\|\text{adj}\, B\|^2, \tag{33}$$

and

$$\beta \equiv \|B\|^6 - (9/2)\|B\|^2\|\text{adj}\, B\|^2 + (27/2)(\det B)^2. \tag{34}$$

Equation (7) can be used to show that $\alpha \geq 0$, with equality if and only if $S_1 = S_2 = |S_3|$, in which case $\beta = 0$ also. Thus we have a complete analytic solution of Wahba's problem.

Computation of the covariance matrix

The quality of the attitude estimate is best expressed in terms of the covariance of the three-component column vector ϕ of attitude error angles in the spacecraft body frame. This parameterization gives the following relation between the estimated and true attitude matrices A and A_{true}:

$$A = \{\exp[(-\phi)\times]\}A_{true} = \{I - [\phi\times] + \tfrac{1}{2}[\phi\times]^2 + \ldots\}A_{true}, \tag{35}$$

where the matrix $[u\times]$ is defined for a general three-component column vector u as

$$[u\times] \equiv \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix}. \tag{36}$$

This notation reflects the equality of the matrix product $[u\times]v$ and the cross product $u \times v$.

Shuster [16] has recast the Wahba problem as a maximum likelihood estimation problem [17], which leads to a very convenient method for computing the covariance matrix. Asymptotically, as the amount of data becomes infinite, the covariance matrix tends to the inverse of the Fisher information matrix F, which is the expected value of the Hessian of the negative-log-likelihood function J;

$$F_{jk} \equiv E[\partial^2 J/\partial\phi_j\partial\phi_k]. \tag{37}$$

The distribution of the components of the ith measurement error vector perpendicular to the true vector are assumed to be Gaussian and axially symmetric about the true vector with variance σ_i^2 per axis. Then the negative-log-likelihood function for this problem is [13, 16]

$$J = \tfrac{1}{2}\sum_{i=1}^{n} \sigma_i^{-2}|b_i - Ar_i|^2 + \ldots, \tag{38}$$

where the omitted terms are independent of attitude. For any positive λ_0 and with

$$\sigma_{tot}^2 \equiv (\sum_{i=1}^{n} \sigma_i^{-2})^{-1},$$ (39)

the weights

$$a_i = \lambda_0 \sigma_{tot}^2 / \sigma_i^2$$ (40)

are positive and satisfy equation (3). With this choice

$$J = \lambda_0^{-1} \sigma_{tot}^{-2} L(A) + \ldots ,$$ (41)

which means that the solution to Wahba's problem is a maximum-likelihood estimate, since it minimizes the negative-log-likelihood function. Substituting equation (35) into equation (2) and using the identity

$$[\mathbf{u} \times][\mathbf{v} \times] = -(\mathbf{v}^T\mathbf{u})I + \mathbf{v}\mathbf{u}^T$$ (42)

gives, to second order,

$$L(A) = \lambda_0 - \text{tr}(A_{true} B^T) + \text{tr}\{[\phi \times]A_{true} B^T\} + \tfrac{1}{2} \text{tr}\{[(\phi^T\phi)I - \phi\phi^T]A_{true} B^T\}$$

$$= \lambda_0 - \text{tr}(A_{true} B^T) + \text{tr}\{[\phi \times]A_{true} B^T\} + \tfrac{1}{2} \phi^T[\text{tr}(A_{true} B^T)I - A_{true} B^T]\phi.$$ (43)

Inserting this into equation (41) and then equation (37) gives the Fisher information matrix

$$F = \lambda_0^{-1} \sigma_{tot}^{-2} [\text{tr}(A_{true} B^T)I - \tfrac{1}{2} (A_{true} B^T + BA_{true}^T)],$$ (44)

and, by matrix inversion, the covariance matrix

$$P = \lambda_0 \sigma_{tot}^2 [\text{tr}(A_{true} B^T)I - \tfrac{1}{2} (A_{true} B^T + BA_{true}^T)]^{-1}.$$ (45)

The true attitude matrix is not known in a real attitude estimation problem, of course, so A_{opt} must be used in place of A_{true} in computing the covariance. Making this replacement in equation (46) gives, with equation (19) and the symmetry of the matrix product $A_{opt} B^T$, which follows from equations (7) and (9),

$$P = \lambda_0 \sigma_{tot}^2 (\lambda I - A_{opt} B^T)^{-1} = \lambda_0 \sigma_{tot}^2 \, \text{adj}(\lambda I - A_{opt} B^T)/\text{det}(\lambda I - A_{opt} B^T).$$ (46)

Equation (46) is one of the forms for the covariance matrix given in Appendix B of [13], which is also the result obtained in [18], simplified to the case that only the attitude is estimated. The computation of the matrix inverse can be avoided as follows [19]. Equations (7), (9), and (15) show that

$$\lambda I - A_{opt} B^T = U_+ \, \text{diag}[S_2 + S_3, S_3 + S_1, S_1 + S_2] U_+^T.$$ (47)

The determinant of this matrix is given by equation (16) as

$$\text{det} (\lambda I - A_{opt} B^T) = \zeta,$$ (48)

and its adjoint is

$$\text{adj}(\lambda I - A_{opt} B^T) = \kappa I + BB^T,$$ (49)

yielding the desired manifestly symmetric result

$$P = \lambda_0 \sigma_{tot}^2 (\kappa I + BB^T)/\zeta.$$ (50)

542

We see that the covariance matrix is infinite when $\zeta = 0$, which agrees with the conditions for indeterminacy of the attitude solution discussed above. In the case of near-indeterminacy, the singular values are approximately $S_1 \approx \lambda$, $S_2 \approx S_3 \approx 0$ [13], which gives the covariance

$$P \approx \lambda_0 \sigma_{tot}^2 U_+ \, \text{diag}[\lambda^2/\zeta, \, \lambda^{-1}, \, \lambda^{-1}] U_+^T. \tag{51}$$

A good criterion for terminating the iterative solution for λ by equation (25) is

$$Q'(\lambda) = 2\zeta < \tfrac{1}{2} \lambda_0^3 \sigma_{tot}^2, \tag{52}$$

since equation (51) predicts attitude estimation error standard deviations larger than $2\lambda/\lambda_0$ radians when this inequality is satisfied. This error can only be small if $\lambda \ll \lambda_0$, in which case the attitude estimate is poor because the loss function is large.

Normalization of the weights

The results above are valid for any positive value of the parameter λ_0, but only two choices are useful:

$$\lambda_0 = 1 \quad \text{(normalized weights)} \tag{53}$$

or

$$\lambda_0 = \sigma_{tot}^{-2} \quad \text{(unnormalized weights).} \tag{54}$$

Past treatments of this problem have generally used normalized weights, which give a B matrix with elements of order unity. This is convenient in computations using fixed-point arithmetic, but floating-point arithmetic is an option on virtually all present-day computers. The normalized form may also be useful if the measurement weights are arbitrarily assigned.

The unnormalized form is more natural if the weights are computed in terms of measurement variances, as in equation (40), since the unnormalized weights are just equal to the inverse variances. The unnormalized form also simplifies the computation of the covariance, as shown by equation (50), but this form can potentially lead to numerical problems. The elements of B are of order σ_{tot}^{-2} if the weights are not normalized, which means that $\| \text{adj} \, B \|^2$ is of order σ_{tot}^{-8}. Since σ_{tot} can be of order 10^{-6} for highly accurate sensors, $\| \text{adj} \, B \|^2$ can be of order 10^{48}, leading to exponent overflow in floating-point representations that do not provide an adequate exponent range. This is not a problem with double-precision arithmetic in conformity with ANSI/IEEE Standard 754-1985 for binary floating-point arithmetic [20], since this standard mandates eleven bits for the exponent, allowing representation of numbers as large as 10^{308}. The Standard Apple Numerical Environment [21] and VAX G_FLOATING [22] double-precision arithmetic employ eleven-bit exponents, but VAX D_FLOATING double-precision arithmetic allots only eight bits for the exponent. This is the same as in IEEE-standard single-precision arithmetic, and allows representation of numbers only as large as 10^{38}. Single-precision arithmetic would lead to exponent overflow problems for measurement variances σ_{tot}^2 less than about 10^{-9}, but double-precision arithmetic is certainly preferred in such cases.

Algorithm test - accuracy

Two forms of the new algorithm, the form with the iterative solution for λ (FOAM — Fast Optimal Attitude Matrix), and the form with the analytic solution for S_1 (SOMA — Slower Optimal Matrix Algorithm), were compared with the SVD method [13] for minimizing Wahba's loss function. The three methods were implemented in double-precision FORTRAN and executed on a DEC VAX 8830 computer. FOAM and SOMA were implemented in G_FLOATING arithmetic with unnormalized weights. FOAM was also implemented with normalized weights in both G_FLOATING and D_FLOATING arithmetic, while the SVD method was implemented with unnormalized weights in D_FLOATING arithmetic.

Four sets of reference vectors were used for the tests:

$$\mathbf{r}_1 = [1, 0, 0]^T, \mathbf{r}_2 = [0, 1, 0]^T, \mathbf{r}_3 = [0, 0, 1]^T, \tag{55}$$

$$\mathbf{r}_1 = [0.6, 0.8, 0]^T, \mathbf{r}_2 = [0.8, -0.6, 0]^T, \tag{56}$$

$$\mathbf{r}_1 = [1, 0, 0]^T, \mathbf{r}_2 = [1, 0.01, 0]^T, \mathbf{r}_3 = [1, 0, 0.01]^T, \tag{57}$$

and

$$\mathbf{r}_1 = [1, 0, 0]^T, \mathbf{r}_2 = [0.96, 0.28, 0]^T, \mathbf{r}_3 = [0.96, 0, 0.28]^T. \tag{58}$$

Set (55) models three sensors with orthogonal boresights along the spacecraft body axes, while set (56) models two sensors with orthogonal boresights not along the body axes. Reference vector set (57) is intended to model three star measurements in a single star sensor with a small field-of-view. Set (58) models one sensor with its boresight along the body x- axis and two sensors with boresights 16.26 degrees off this axis. The observation vectors were computed as

$$\mathbf{b}_i = A_{true}\,\mathbf{r}_i + \mathbf{n}_i, \tag{59}$$

where

$$A_{true} = \begin{bmatrix} 0.352 & 0.864 & 0.360 \\ -0.864 & 0.152 & 0.480 \\ 0.360 & -0.480 & 0.800 \end{bmatrix}, \tag{60}$$

which has all non-zero matrix elements with exact decimal representations and is otherwise arbitrary, and $\mathbf{n}_i$ is a vector of measurement errors. The tests were run both with $\mathbf{n}_i = 0$ and with measurement errors simulated by zero-mean Gaussian white noise on the components of $\mathbf{n}_i$. All the methods normalize the input observation and reference vectors; some efficiencies in the normalization process were found and applied to the three algorithms.

The results of the accuracy tests are presented in Table 1. The reference vector sets are labeled REF. The standard deviations (in radians) in the table were used to compute the measurement weights and also the level of measurement errors in the tests where these were simulated. Only two measurements were used in the tests in which only two standard deviations are given. The quantities presented in the table are the estimation error in radians

(computed with simulated measurement errors),

$$EST = \sin^{-1}(2^{-3/2} \| A_{opt} A_{true}^T - A_{true} A_{opt}^T \|), \tag{61}$$

the maximum computation error for all FOAM and SOMA variants (computed with $n_i = 0$),

$$COMP = \| A_{opt} - A_{true} \|, \tag{62}$$

and the maximum orthogonality error for FOAM and SOMA,

$$ORTH = \| A_{opt} A_{opt}^T - I \|. \tag{63}$$

The estimation error was the same for all methods, to the accuracy of the computation errors. As expected, the very robust SVD method gives the smallest maximum orthogonality error (2.16×10^{-16}) and computation errors $(4.72 \times 10^{-17}$ for cases 1 - 4, 1.63×10^{-10} for case 5, 3.74×10^{-15} for cases 6 - 11, and 2.10×10^{-9} for case 12). No significant differences were seen between FOAM and SOMA or between normalized and unnormalized weights. D_FLOATING arithmetic was about one decimal digit more precise than G_FLOATING arithmetic, as expected [22]; but this is not significant, since the computation errors are much less than the estimation errors in all cases with realistic noise. It is clear that cases with widely differing measurement accuracies furnish the greatest computational challenges.

Algorithm test - speed

The above methods were compared with Shuster's QUEST (QUaternion ESTimation) algorithm [9] for computational speed, since QUEST is the fastest previously known algorithm for solving Wahba's problem. In addition to the reference and observation vectors

Table 1
Accuracy Test Results. See text for explanation

CASE	REF	σ_1	σ_2	σ_3	EST	COMP	ORTH
1	(55)	10^{-6}	10^{-6}	10^{-6}	1.38×10^{-6}	4.61×10^{-16}	1.12×10^{-15}
2	(55)	10^{-6}	10^{-6}	—	2.02×10^{-6}	3.05×10^{-16}	6.11×10^{-16}
3	(55)	.01	.01	.01	1.39×10^{-2}	5.27×10^{-16}	1.01×10^{-15}
4	(55)	.01	.01	—	2.05×10^{-2}	3.05×10^{-16}	1.12×10^{-15}
5	(56)	10^{-6}	.01	—	1.12×10^{-2}	7.83×10^{-9}	2.73×10^{-8}
6	(57)	10^{-6}	10^{-6}	10^{-6}	2.51×10^{-5}	4.66×10^{-12}	8.94×10^{-12}
7	(57)	10^{-6}	10^{-6}	—	3.18×10^{-5}	7.84×10^{-12}	1.54×10^{-11}
8	(57)	.01	.01	.01	0.186	4.04×10^{-12}	7.50×10^{-12}
9	(57)	.01	.01	—	8.82×10^{-2}	5.70×10^{-12}	1.12×10^{-11}
10	(58)	10^{-6}	.01	.01	1.72×10^{-2}	1.49×10^{-7}	2.97×10^{-7}
11	(58)	10^{-6}	.01	—	3.33×10^{-2}	1.45×10^{-7}	2.87×10^{-7}
12	(58)	.01	10^{-6}	—	3.48×10^{-2}	3.01×10^{-7}	6.00×10^{-7}

and the measurement standard deviations, QUEST requires the input of five control parameters, which were taken as $QUIBBL = 0.1$, $FIBBL = 10^{-5}$, $QUACC = 10^{-8}$, $NEWT = 10$, and $IMETH = 1$. The measured CPU times were effectively the same for normalized and unnormalized weights. They consist of a part that is independent of the number of observations processed and a part proportional to the number of observations:

$$t_{QUEST} = 0.24 + 0.09\,n \quad \text{msec.} \tag{64}$$

$$t_{FOAM} = 0.27 + 0.07\,n \quad \text{msec,} \tag{65}$$

$$t_{SOMA} = 0.36 + 0.07\,n \quad \text{msec,} \tag{66}$$

$$t_{SVD} = (3 \pm 1) + 0.07\,n \quad \text{msec.} \tag{67}$$

The greater n-dependent time in QUEST as compared to the other algorithms is due to the method used to compute the covariance matrix in QUEST. The computation of λ generally requires one or two iterations in QUEST and two to six iterations in FOAM, due to the need to iterate to convergence in the latter method, which accounts for the greater n-independent time in FOAM. The transcendental function calls in SOMA account for its longer running time compared to FOAM, which is definitely preferable to SOMA since it is faster and no less accurate. The range of times for the SVD method is related to the rank and conditioning of the B matrix. This method is significantly slower than all the other methods tested, as has been noted previously; but the SVD method may still find applications in nearly singular estimation problems. The exact CPU times will vary from case to case, and the time required for either FOAM or QUEST appears to be quite modest in comparison with other computations performed in spacecraft attitude determination.

It should be pointed out that FOAM computes the attitude matrix directly, while QUEST computes an attitude quaternion. If an attitude matrix is required from QUEST, an additional step is required to compute it from the quaternion. This requires only multiplications and additions, though, and no transcendental function evaluations. If it is desired to compute a quaternion from FOAM, the standard method for extracting it from the attitude matrix can be used [23]. This requires the evaluation of one square root, but FOAM is faster than QUEST even with this addition. The principal advantage of FOAM over QUEST in practice is that it requires no control parameter input; its only inputs are the number of observations, the reference and observation vectors, and the measurement standard deviations.

Two-observation case

In the special case of two observations, the rank of B is at most two, so $\det B = 0$, which gives with equation (22)

$$\kappa = \| \mathrm{adj}\, B \|, \tag{68}$$

$$\lambda = (2\kappa + \| B \|^2)^{1/2}, \tag{69}$$

and

$$\zeta = \kappa\lambda. \tag{70}$$

Both κ and λ must be positive in order for λ to be the largest root of $Q(\lambda)$. The explicit form for B as a function of the reference and observation vectors then yields

$$\text{adj } B^T = a_1 a_2 (\mathbf{b}_1 \times \mathbf{b}_2)(\mathbf{r}_1 \times \mathbf{r}_2)^T, \tag{71}$$

$$\kappa = a_1 a_2 |\mathbf{b}_1 \times \mathbf{b}_2| \, |\mathbf{r}_1 \times \mathbf{r}_2|, \tag{72}$$

and

$$\lambda = \{a_1^2 + 2 a_1 a_2 [\, |\mathbf{b}_1 \times \mathbf{b}_2| \, |\mathbf{r}_1 \times \mathbf{r}_2| + (\mathbf{b}_1^T \mathbf{b}_2)(\mathbf{r}_1^T \mathbf{r}_2)] + a_2^2 \}^{1/2}. \tag{73}$$

The attitude is indeterminate if either the two reference vectors or the two observation vectors are parallel or antiparallel. Thus we will assume that both θ_r, the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$, and θ_b, the angle between $\mathbf{b}_1$ and $\mathbf{b}_2$, are strictly greater than zero and strictly less than pi. Now set $\lambda_0 = a_1 + a_2 = 1$ for the remainder of the discussion in this section, define

$$\varepsilon \equiv (\theta_b - \theta_r)/2, \tag{74}$$

and note that $|\varepsilon| < \pi/2$. This allows the expression for λ to be written more compactly as

$$\lambda = (1 - 4 a_1 a_2 \sin^2 \varepsilon)^{1/2}. \tag{75}$$

These expressions for λ in the two-observation case are equivalent to equation (72) in [9].

It is convenient to write the optimal attitude estimate in terms of the orthonormal triads:

$$\mathbf{r}_+ \equiv (\mathbf{r}_2 + \mathbf{r}_1)/[2\cos(\theta_r/2)], \tag{76a}$$

$$\mathbf{r}_- \equiv (\mathbf{r}_2 - \mathbf{r}_1)/[2\sin(\theta_r/2)], \tag{76b}$$

$$\mathbf{r}_+ \times \mathbf{r}_- = (\mathbf{r}_1 \times \mathbf{r}_2)/|\mathbf{r}_1 \times \mathbf{r}_2|, \tag{76c}$$

and

$$\mathbf{b}_+ \equiv (\mathbf{b}_2 + \mathbf{b}_1)/[2\cos(\theta_b/2)], \tag{77a}$$

$$\mathbf{b}_- \equiv (\mathbf{b}_2 - \mathbf{b}_1)/[2\sin(\theta_b/2)], \tag{77b}$$

$$\mathbf{b}_+ \times \mathbf{b}_- = (\mathbf{b}_1 \times \mathbf{b}_2)/|\mathbf{b}_1 \times \mathbf{b}_2|. \tag{77c}$$

Other orthogonal triads can be defined, but these preserve the maximum symmetry between the two measurements. The optimal attitude matrix expressed in terms of these triads is

$$A_{opt} = (1 - 4 a_1 a_2 \sin^2 \varepsilon)^{-1/2} [\cos\varepsilon\,(\mathbf{b}_+ \mathbf{r}_+^T + \mathbf{b}_- \mathbf{r}_-^T) + (a_1 - a_2)\sin\varepsilon\,(\mathbf{b}_+ \mathbf{r}_-^T - \mathbf{b}_- \mathbf{r}_+^T)]$$
$$+ (\mathbf{b}_+ \times \mathbf{b}_-)(\mathbf{r}_+ \times \mathbf{r}_-)^T. \tag{78}$$

It is interesting to note that a factor of $a_1 a_2$ in the denominator of equation (12) has cancelled an identical factor in the numerator. Thus the attitude estimate has a well-defined limit as either a_1 or a_2 tends to zero, even though Wahba's loss function does not have a unique minimum in either limit. Another interesting property of the two-observation case is that the optimal estimate is independent of the weights when $\varepsilon = 0$. Equations (24) and (75) with $\lambda_0 = 1$ show that the optimized loss function is zero if any of a_1, a_2, or ε is zero.

We now investigate the conditions under which this optimal attitude estimate can be obtained by a generalization of the simpler TRIAD or algebraic method [8, 9]. This is a well-known algorithm for computing an attitude matrix from two vector observations by forming orthonormal triads from the reference and observation vectors. One of vectors in the reference triad is the normalized cross product of the two reference vectors, and the other two are orthonormal linear combinations of the two reference vectors. The most general form for the reference triad that we will consider is:

$$\mathbf{r}_I \equiv \cos\psi_r \mathbf{r}_+ - \sin\psi_r \mathbf{r}_- = [\sin(\psi_r + \theta_r/2)\mathbf{r}_1 - \sin(\psi_r - \theta_r/2)\mathbf{r}_2]/\sin\theta_r, \tag{79a}$$

$$\mathbf{r}_{II} \equiv \cos\psi_r \mathbf{r}_- + \sin\psi_r \mathbf{r}_+ = [\cos(\psi_r - \theta_r/2)\mathbf{r}_2 - \cos(\psi_r + \theta_r/2)\mathbf{r}_1]/\sin\theta_r, \tag{79b}$$

$$\mathbf{r}_I \times \mathbf{r}_{II} = \mathbf{r}_+ \times \mathbf{r}_-, \tag{79c}$$

where ψ_r is some rotation angle in the plane spanned by $\mathbf{r}_1$ and $\mathbf{r}_2$. The observation triad is

$$\mathbf{b}_I \equiv \cos\psi_b \mathbf{b}_+ - \sin\psi_b \mathbf{b}_- = [\sin(\psi_b + \theta_b/2)\mathbf{b}_1 - \sin(\psi_b - \theta_b/2)\mathbf{b}_2]/\sin\theta_b, \tag{80a}$$

$$\mathbf{b}_{II} \equiv \cos\psi_b \mathbf{b}_- + \sin\psi_b \mathbf{b}_+ = [\cos(\psi_b - \theta_b/2)\mathbf{b}_2 - \cos(\psi_b + \theta_b/2)\mathbf{b}_1]/\sin\theta_b, \tag{80b}$$

$$\mathbf{b}_I \times \mathbf{b}_{II} = \mathbf{b}_+ \times \mathbf{b}_-, \tag{80c}$$

similarly. The angles ψ_r and ψ_b are chosen to give more or less weight to the two vector measurements. The choice $\psi_r = \psi_b = 0$, for example, gives equal weight to the two measurements. The choice $\psi_r = \theta_r/2$ and $\psi_b = \theta_b/2$ gives

$$\mathbf{r}_I = \mathbf{r}_1, \tag{81a}$$

$$\mathbf{r}_{II} = (\mathbf{r}_2 - \cos\theta_r \mathbf{r}_1)/\sin\theta_r, \tag{81b}$$

and similar relations for $\mathbf{b}_I$ and $\mathbf{b}_{II}$, with maximum weight on the first measurement. The choice $\psi_r = -\theta_r/2$ and $\psi_b = -\theta_b/2$, on the other hand, gives

$$\mathbf{r}_I = \mathbf{r}_2, \tag{82a}$$

$$\mathbf{r}_{II} = -(\mathbf{r}_1 - \cos\theta_r \mathbf{r}_2)/\sin\theta_r, \tag{82b}$$

and similarly for $\mathbf{b}_I$ and $\mathbf{b}_{II}$, with maximum weight on the second measurement. The key point is that ψ_r is some function of θ_r and the measurement weights, and ψ_b is *the same* function of θ_b and the weights. Note that this does not imply that $\psi_r = \psi_b$ except in the case that $\varepsilon = 0$. Often, the TRIAD method is understood to mean only the special cases of equations (81) or (82), rather than the generalized method specified by equations (79) and (80).

The TRIAD attitude estimate is given by

$$A_{\text{TRIAD}} = [\mathbf{b}_I \vdots \mathbf{b}_{II} \vdots \mathbf{b}_I \times \mathbf{b}_{II}] [\mathbf{r}_I \vdots \mathbf{r}_{II} \vdots \mathbf{r}_I \times \mathbf{r}_{II}]^T = \mathbf{b}_I \mathbf{r}_I^T + \mathbf{b}_{II} \mathbf{r}_{II}^T + (\mathbf{b}_I \times \mathbf{b}_{II})(\mathbf{r}_I \times \mathbf{r}_{II})^T$$

$$= \cos(\psi_b - \psi_r)(\mathbf{b}_+ \mathbf{r}_+^T + \mathbf{b}_- \mathbf{r}_-^T) + \sin(\psi_b - \psi_r)(\mathbf{b}_+ \mathbf{r}_-^T - \mathbf{b}_- \mathbf{r}_+^T)$$

$$+ (\mathbf{b}_+ \times \mathbf{b}_-)(\mathbf{r}_+ \times \mathbf{r}_-)^T. \tag{83}$$

We now attempt to find angles ψ_r and ψ_b such that the TRIAD solution gives the optimal

548

attitude estimate of equation (78). We immediately find such angles in four special cases:

1) If $\varepsilon = 0$, then $\psi_r = \psi_b$ automatically, all TRIAD solutions are the same, and they all agree with the optimal estimate, which is independent of the weights in the loss function.

2) If $a_1 = a_2 = 1/2$, the TRIAD solution with $\psi_r = \psi_b = 0$ and with vector triads given by equations (76) and (77) gives the optimal estimate.

3) If $a_1 = 1, a_2 = 0$, the TRIAD solution with $\psi_r = \theta_r/2$, $\psi_b = \theta_b/2$ and with triads as in equations (81) gives the optimal estimate.

4) If $a_1 = 0, a_2 = 1$, the TRIAD solution with $\psi_r = -\theta_r/2$, $\psi_b = -\theta_b/2$ and with triads as in equations (82) gives the optimal estimate.

We will now show that the TRIAD solution does not minimize Wahba's loss function except in these four special cases. Comparing equations (78) and (83) gives the following necessary condition for agreement of the TRIAD and optimal attitude estimates:

$$\tan(\psi_b - \psi_r) = (a_1 - a_2)\tan\varepsilon. \tag{84}$$

Set $\theta_r = \theta_0$, some arbitrarily chosen angle, and denote the corresponding value of ψ_r by ψ_0, which is also a function of the observation weights. Then

$$\tan(\psi_b - \psi_0) = (a_1 - a_2)\tan[(\theta_b - \theta_0)/2] \equiv (a_1 - a_2)\tau_b. \tag{85}$$

This equation must hold for any θ_r, with ψ_0 and θ_0 regarded as fixed parameters, since ψ_b is required to be a function of θ_b and the weights only, and not of θ_r. Now setting $\theta_b = \theta_0$ in equation (84) gives $\psi_b = \psi_0$ and

$$\tan(\psi_r - \psi_0) = (a_1 - a_2)\tan[(\theta_r - \theta_0)/2] \equiv (a_1 - a_2)\tau_r, \tag{86}$$

which must hold for any θ_b. In fact, equation (86) could have been written directly in analogy with equation (85), since ψ_r is required to be the same function of θ_r and the measurement weights as ψ_b is of θ_b and the weights. Now combining equations (85) and (86) with some elementary trigonometry gives

$$\tan(\psi_b - \psi_r) = \tan[(\psi_b - \psi_0) - (\psi_r - \psi_0)] = (a_1 - a_2)(\tau_b - \tau_r)/[1 + (a_1 - a_2)^2\tau_b\tau_r]$$

$$= (a_1 - a_2)\tan\varepsilon(1 + \tau_b\tau_r)/[1 + (1 - 4a_1a_2)\tau_b\tau_r]. \tag{87}$$

Equating the right sides of equations (84) and (87) gives, after some cancellations, the necessary condition

$$4a_1a_2\tau_b\tau_r(a_1 - a_2)\tan\varepsilon = 0, \tag{88}$$

which is satisfied in the four special cases discussed above. It is also satisfied if either τ_b or τ_r is zero, but these conditions cannot be satisfied in general since θ_0 is an arbitrarily chosen angle. Thus the TRIAD method cannot find the optimal attitude minimizing Wahba's loss function in the general case, but only in the special cases $\varepsilon = 0$, $a_1 = 0$, $a_2 = 0$, and $a_1 = a_2$.

Conclusions

A new algorithm for minimizing Wahba's loss function has been found, which solves for the optimal attitude matrix directly, without the intermediate computation of a quaternion or other parameterization of the attitude. The attitude quaternion can be computed from the attitude matrix, if desired; and the new method with iterative solution of the scalar coefficients in the attitude matrix is at least as fast as existing methods even with this additional computation. The scalar coefficients used in computing the optimal attitude matrix are also used to compute the covariance of the attitude error angles. Since the attitude matrix is inherently nonsingular, there are no problems with special cases like 180 degree rotations, and no special procedures are needed to deal with such cases. The principal practical advantage of the new method over existing fast optimal attitude estimators is that it requires no control parameter input; its only inputs are the number of observations, the reference and observation vectors, and the measurement standard deviations.

A closed-form solution for the optimal attitude matrix is presented for the special case of two observations. This solution is compared with the estimate produced by the well-known non-optimal method based on orthonormal triads formed from the observation and reference vectors. When the angle between the two reference vectors is equal to the angle between the two observation vectors, all triad choices give the optimal estimate, which is independent of the weights in the loss function. Except for this case, the optimal and triad-based attitude estimates agree only when the two vector measurements are given equal weights in the loss function or when the weight given to one vector measurement is negligible compared to the weight given to the other.

Acknowledgements

References

[1] WAHBA, GRACE "A Least Squares Estimate of Spacecraft Attitude," *SIAM Review*, Vol. 7, No. 3, July 1965, p. 409.

[2] FARRELL, J. L. and STUELPNAGEL, J. C. "A Least Squares Estimate of Spacecraft Attitude," *SIAM Review*, Vol. 8, No. 3, July 1966, pp. 384-386.

[3] WESSNER, R. H. ibid.

[4] VELMAN, J. R. ibid.

[5] BROCK, J. E. ibid.

[6] DAVENPORT, PAUL B. "A Vector Approach to the Algebra of Rotations with Applications," NASA X-546-65-437, November 1965.

[7] KEAT, J. "Analysis of Least Squares Attitude Determination Routine DOAOP," Computer Sciences Corporation, CSC/TM-77/6034, February 1977.

[8] LERNER, GERALD M. "Three-Axis Attitude Determination," in *Spacecraft Attitude Determination and Control*, James R. Wertz, ed., D. Reidel, Dordrecht, Holland, 1978.

[9] 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.

[10] GOLUB, GENE H. and VAN LOAN, CHARLES F. *Matrix Computations*, the Johns Hopkins University Press, Baltimore, MD, 1983

[11] HORN, ROGER A. and JOHNSON, CHARLES R. *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985

[12] DAVENPORT, PAUL B. "Attitude Determination and Sensor Alignment via Weighted Least Squares Affine Transformations," AAS Paper No. 71-396, AAS/AIAA Astrodynamics Specialists Conference, Ft. Lauderdale, FL, August 1971.

[13] MARKLEY, F. LANDIS "Attitude Determination Using Vector Observations and the Singular Value Decomposition," *Journal of the Astronautical Sciences*, Vol. 36, No. 3, July-September 1988, pp. 245-258

[14] DANBY, J. M. A. *Fundamentals of Celestial Mechanics*, 2nd edition, Willman-Bell, Inc., Richmond, VA, 1988

[15] BEYER, WILLIAM H. *CRC Standard Mathematical Tables, 2nd Edition*, CRC Press, Inc., Boca Raton, FL, 1987

[16] SHUSTER, MALCOLM D. "Maximum Likelihood Estimation of Spacecraft Attitude," *Journal of the Astronautical Sciences*, Vol. 37, No. 1, January-March 1989, pp. 79-88

[17] SORENSON, HAROLD W. *Parameter Estimation: Principles and Problems*, Marcel Dekker, New York, NY, 1980

[18] MARKLEY, F. LANDIS "Attitude Determination and Parameter Estimation Using Vector Observations: Theory," *Journal of the Astronautical Sciences*, Vol. 37, No. 1, January-March 1989, pp.41-58

[19] MARKLEY, F. LANDIS "Attitude Determination and Parameter Estimation Using Vector Observations: Application," *Journal of the Astronautical Sciences*, Vol. 39, No. 3, July-September 1991, pp. 367-381

[20] *IEEE Standard for Binary Floating-Point Arithmetic*, ANSI/IEEE Standard 754-1985, The Institute of Electrical and Electronics Engineers, Inc., New York, NY, 1985

[21] *Apple Numerics Manual, 2nd Edition*, Addison-Wesley, Reading, MA, 1988

[22] *VAX-11 FORTRAN Language Reference Manual*, Digital Equipment Corporation, Maynard, MA, 1982

[23] SHEPPERD, STANLEY W. "Quaternion from Rotation Matrix," *Journal of Guidance and Control*, Vol. 1, No. 3, May-June 1978, pp. 223-224