The Journal of the Astronautical Sciences, Vol. 36, No. 3, July-September 1988, pp. 245–258

Attitude Determination using Vector Observations and the Singular Value Decomposition[1]

F. Landis Markley[2]

Abstract

A new method for finding the attitude matrix minimizing Wahba's loss function, based on the singular value decomposition of a 3×3 matrix, is presented. Equations are given for the covariance matrix of the attitude estimate, as well as for the eigenvalues and eigenvectors of this matrix, in terms of the singular value decomposition matrices. The singular value decomposition method is compared with Shuster's implementation of Davenport's q-method, which is more efficient than the new algorithm but does not give the eigenvalues and eigenvectors of the covariance matrix. These are often useful for analysis, since the maximum eigenvalue and its eigenvector give the magnitude and direction of the largest component of the attitude uncertainty.

Introduction

In 1965, Wahba [1] posed the three-axis attitude determination problem in terms of finding the proper othogonal matrix A_{opt} that minimizes the least-squares loss function

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

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

$$\mathbf{b}_i = A_{true}\mathbf{r}_i + \boldsymbol{\epsilon}_i \quad (2)$$

for all i, where A_{true} is the true attitude matrix and the $\boldsymbol{\epsilon}_i$ are related to the measurement errors. The attitude matrix minimizing Wahba's loss function is not the minimum variance estimate.

[1]Paper No. AAS 87–490, AAS/AIAA Astrodynamics Conference, Kalispell, Montana, August 10–13, 1987.
[2]Head, Mathematical Analysis Section, Systems Development Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

We assume that we have at least one observation, so we can normalize the weights to give

$$\sum_{i=1}^{n} a_i = 1 . \tag{3}$$

Then it is easy to show that [2]

$$L(A) = 1 - \sum_{i=1}^{n} a_i \mathbf{b}_i^T A \mathbf{r}_i = 1 - \operatorname{tr}(AB^T) , \tag{4}$$

where

$$B \equiv \sum_{i=1}^{n} a_i \mathbf{b}_i \mathbf{r}_i^T . \tag{5}$$

The superscript T denotes the matrix transpose, and tr denotes the trace of a matrix. The last step in equation (4) follows from the invariance of the trace of a product of matrices under a cyclic permutation of the factors in the product.

This problem has a long history [3] and has been the basis for several operational spacecraft attitude determination algorithms, especially in the form discovered by Davenport (as reported by Keat [4]), and later modifed and extended by Shuster [5]. This paper presents a solution of Wahba's problem involving the singular value decomposition (SVD) of the matrix B [6–9]. The SVD provides both an elegant tool for theoretical analysis and a robust procedure for computing the attitude matrix minimizing Wahba's loss function. The SVD is implicit in early solutions of Wahba's problem by Farrell and Stuelpnagel [2], Wessner [10], Velman [11], and Brock [12, 13], and in papers by Davenport [14] and Björck and Bowie [15], but it has not been used for computation of the optimal attitude before. The method of Farrell and Stuelpnagel [2] uses a polar decomposition of B followed by diagonalization of its positive semidefinite factor. Other approaches compute the eigenvalues and eigenvectors of the matrix B^TB, resulting in a loss of accuracy [7]. Methods that require division by the square roots of the eigenvalues of B^TB fail when B is singular, as do some, but not all, iterative schemes for computing the optimal attitude [15–21]. It should be pointed out that the iterative methods were applied in cases where B was known to be close to an orthogonal matrix and therefore nonsingular.

Davenport's q-method [4] finds the quaternion representing the optimal rotation as the eigenvector with maximum eigenvalue of a symmetric 4×4 matrix, the elements of which are simple linear functions of the elements of B. Finding all the eigenvectors would be more expensive than performing the SVD of a 3×3 matrix, but this can be avoided by iterative computation of the maximum eigenvalue using the knowledge that it is very close to unity [5]. Tietze has proposed a method based on inverse iteration and the same knowledge about the maximum eigenvalue [22].

The derivation of the SVD algorithm is given in the next section. Then the possibility of ambiguous solutions is considered and some compact formulas for the attitude covariance in terms of the SVD are derived. Finally, the implementation of the algorithm is discussed.

The SVD Solution to Wahba's Problem

The singular value decomposition of the matrix B defined in equation (5) is given by

$$B = USV^T, \tag{6}$$

where U and V are orthogonal matrices, and

$$S = \text{diag}(s_1, s_2, s_3) \tag{7}$$

with

$$s_1 \geq s_2 \geq s_3 \geq 0 . \tag{8}$$

The notation of equation (7) means that S is a diagonal matrix with diagonal elements s_1, s_2, s_3, the singular values of B. It is convenient to define the proper orthogonal matrices

$$U_+ \equiv U[\text{diag}(1, 1, \det U)], \tag{9}$$

$$V_+ \equiv V[\text{diag}(1, 1, \det V)], \tag{10}$$

and

$$W \equiv U_+^T A V_+ = \cos\Phi I + (1 - \cos\Phi)\mathbf{e}\mathbf{e}^T - \sin\Phi[\mathbf{e}\times], \tag{11}$$

where I is the 3×3 identity matrix and $[\mathbf{e}\times]$ is defined as

$$[\mathbf{e}\times] = \begin{bmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & -e_1 \\ -e_2 & e_1 & 0 \end{bmatrix}. \tag{12}$$

This Euler axis/angle representation of W by a unit vector $\mathbf{e}$ and a rotation angle Φ is possible for any proper orthogonal matrix. The correspondence defined by equation (12) between a skew-symmetric 3×3 matrix and a 3×1 column vector will be assumed to hold for several other 3-component quantities defined below. This notation reflects the fact that the matrix product $[\mathbf{e}\times]\mathbf{v}$ is equal to the cross product $\mathbf{e} \times \mathbf{v}$ for any 3-vector $\mathbf{v}$. We also define the diagonal matrix

$$S' \equiv \text{diag}(s_1, s_2, ds_3) , \tag{13}$$

where

$$d \equiv (\det U)(\det V) = \pm 1 . \tag{14}$$

In terms of these matrices equation (6) can be written

$$B = U_+ S' V_+^T \tag{15}$$

Substituting this into equation (4) and using the cyclic invariance of the trace and equation (11) gives

$$L(A) = 1 - \text{tr}(S'W) = 1 - \text{tr}S' + (1 - \cos\Phi)[s_2 + ds_3 + (s_1 - s_2)e_2^2 + (s_1 - ds_3)e_3^2] . \tag{16}$$

In view of equation (8), we see that $L(A)$ is minimized for $\Phi = 0$, which gives $W = I$,

$$L(A_{opt}) = 1 - \text{tr}S' = 1 - s_1 - s_2 - ds_3 , \tag{17}$$

and, from equations (9)–(11),

$$A_{opt} = U_+ V_+^T = U[\text{diag}(1, 1, d)]V^T . \tag{18}$$

This minimum is unique unless $s_2 + ds_3 = 0$, in which case there is at least a one-parameter family of minimizing W matrices given by setting $e_2 = e_3 = 0$, so that

$$W = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\Phi & \sin\Phi \\ 0 & -\sin\Phi & \cos\Phi \end{bmatrix} \tag{19}$$

These include the identity matrix as the special case given by $\Phi = 0$. For specificity, we always use equation (18) for A_{opt}.

Equation (18) represents the transformation from reference to body coordinates as the product of two transformations. The matrix V_+^T transforms from the reference frame to an intermediate frame, which we shall call the S-frame, and U_+ transforms from the S-frame to the spacecraft body frame.

Uniqueness of the SVD Solution

The uniqueness of the solution is closely related to the rank of the B matrix, which is equal to the number of non-zero singular values [6]. If the rank of B is less than two, we have $s_2 = s_3 = 0$, and the attitude matrix is not unique. This is not at all surprising, since the rank of B is equal to the number of linearly independent reference vectors, and it is known that at least two linearly independent reference vectors are needed to determine the attitude uniquely.

We will now consider only B matrices of rank two or three, so $s_2 > 0$. The determinant of B is

$$\det B = \det(USV^T) = ds_1s_2s_3 . \tag{20}$$

Thus if $\det B \geq 0$, either $d = 1$ or else $d = -1$ and $s_2 > s_3 = 0$; the optimal solution is unique in either case. This is the situation if the measurement errors are zero, as the following argument shows. In the absence of errors, equations (2) and (5) give

$$B = MA_{true} , \tag{21}$$

where M is the real, positive semidefinite, symmetric matrix

$$M \equiv \sum_{i=1}^{n} a_i \mathbf{b}_i \mathbf{b}_i^T . \tag{22}$$

Since A_{true} is a proper orthogonal matrix,

$$\det B = \det M = m_1m_2m_3 \geq 0 , \tag{23}$$

where m_1, m_2, m_3 are the non-negative eigenvalues of M. Thus the optimal solution is unique in the absence of measurement errors if B has rank two or three. In the presence of errors, the determinant of B may be negative, but its magnitude will be small, on the order of the errors. Thus $s_2 + ds_3$ will be zero for positive s_2 only if $d = -1$ and

$s_2 = s_3 \approx 0$, which means that the solution is unique unless B is very close to having rank less than two.

Sensitivity Analysis

We want to find the sensitivity of the optimal attitude estimate to variations in the reference and observation vectors. These variations are not to be interpreted as time-dependent motions of the vectors. Such variations change the B matrix by δB, causing variations δU_+ and δV_+, giving

$$\delta A_{opt} = (\delta U_+)V_+^T + U_+(\delta V_+)^T, \tag{24}$$

to first order in the variations. Ignoring second-order terms, the orthogonality of A_{opt} means that the matrix

$$[\mathbf{\Theta}\times] \equiv -(\delta A_{opt})A_{opt}^T, \tag{25}$$

is skew-symmetric. The column vector $\mathbf{\Theta}$, which corresponds to $[\mathbf{\Theta}\times]$ by the convention introduced below equation (12), is a more compact representation of the attitude variation than δA_{opt}; it contains the components in the spacecraft body frame of the angular variation in the attitude. Substituting for δA_{opt} and A_{opt}^T in equation (25) gives

$$[\mathbf{\Theta}\times] = -[(\delta U_+)V_+^T + U_+(\delta V_+)^T]V_+U_+^T = U_+([\mathbf{u}\times] + [\mathbf{v}\times])U_+^T, \tag{26}$$

where the skew-symmetric matrices $[\mathbf{u}\times]$ and $[\mathbf{v}\times]$ are defined as

$$[\mathbf{u}\times] \equiv -U_+^T(\delta U_+) \tag{27}$$

and

$$[\mathbf{v}\times] \equiv -(\delta V_+)^T V_+ . \tag{28}$$

Equation (26) is equivalent to the vector equation

$$\mathbf{\Theta} = U_+(\mathbf{u} + \mathbf{v}) . \tag{29}$$

The 3×1 column vectors $\mathbf{v}$ and $\mathbf{u}$ contain the components in the S-frame of the angular variations in the transformations from the reference frame to the S-frame and from the S-frame to the body frame, respectively. We must now find $\mathbf{u} + \mathbf{v}$ in terms of δB.

$$\delta B = (\delta U_+)S'V_+^T + U_+(\delta S')V_+^T + U_+S'(\delta V_+)^T, \tag{30}$$

so using equations (27) and (28) gives

$$U_+^T(\delta B)V_+ = -[\mathbf{u}\times]S' + \delta S' - S'[\mathbf{v}\times] \tag{31}$$

and

$$[\mathbf{z}\times] \equiv [U_+^T(\delta B)V_+]^T - U_+^T(\delta B)V_+ = ([\mathbf{u}\times] + [\mathbf{v}\times])S' + S'([\mathbf{u}\times] + [\mathbf{v}\times]) . \tag{32}$$

Solving equation (32) for the components of $\mathbf{u} + \mathbf{v}$ and substituting into equation (29) gives

$$\mathbf{\Theta} = U_+D^{-1}\mathbf{z}, \tag{33}$$

where

$$D \equiv \mathrm{diag}(s_2 + ds_3, ds_3 + s_1, s_1 + s_2) = [1 - L(A_{opt})]I - S' . \qquad (34)$$

We can use equation (5) to find δB in terms of the variations $\delta \mathbf{r}_i$ and $\delta \mathbf{b}_i$ in the reference and observation vectors:

$$\delta B = \sum_{i=1}^{n} a_i[(\delta \mathbf{b}_i)\mathbf{r}_i^T + \mathbf{b}_i(\delta \mathbf{r}_i)^T] . \qquad (35)$$

Then equation (32) gives, after some algebra

$$\mathbf{z} = \sum_{i=1}^{n} a_i[(U_+^T \delta \mathbf{b}_i) \times (V_+^T \mathbf{r}_i) + (U_+^T \mathbf{b}_i) \times (V_+^T \delta \mathbf{r}_i)] . \qquad (36)$$

Equations (33), (34), and (36) give the first-order variation in the attitude $\boldsymbol{\Theta}$ as a function of the variations $\delta \mathbf{r}_i$ and $\delta \mathbf{b}_i$. They show the extreme sensitivity of the computed attitude when $s_2 + ds_3 \approx 0$; this restates quantitatively the uniqueness results of the last section.

Covariance Analysis

The covariance of the attitude estimate is a statistical measure of the estimation errors arising from errors in the reference and observation vectors. The estimation errors, which are the difference between the true and estimated attitude, should not be confused with control errors, which are the difference between the true and commanded attitude. We assume that these errors are small and linearize about the underlying error-free solution, using equations (33) and (36) with a different interpretation. For the sensitivity analysis $\mathbf{b}_i$ and $\mathbf{r}_i$ are the actual measured vectors, while for the statistical analysis $\mathbf{b}_i$ and $\mathbf{r}_i$ are the underlying error-free vectors and $\delta \mathbf{b}_i$ and $\delta \mathbf{r}_i$ are statistical errors, so $\boldsymbol{\Theta}$ is the attitude estimation error vector. The matrices U_+ and V_+ and the singular values are also error-free for the statistical analysis, since they are computed from $\mathbf{b}_i$ and $\mathbf{r}_i$; and the attitude estimate A_{opt} is equal to A_{true} if the attitude solution is unique, as we shall assume. Thus we now have

$$\mathbf{b}_i = A_{true}\mathbf{r}_i = A_{opt}\mathbf{r}_i = U_+ V_+^T \mathbf{r}_i , \qquad (37)$$

which gives, upon substitution into equation (36),

$$\mathbf{z} = U_+^T \sum_{i=1}^{n} a_i[\delta \mathbf{b}_i \times (A_{opt}\mathbf{r}_i) + \mathbf{b}_i \times (A_{opt}\delta \mathbf{r}_i)] = U_+^T \sum_{i=1}^{n} a_i[\mathbf{b}_i \times](A_{opt}\delta \mathbf{r}_i - \delta \mathbf{b}_i) . \qquad (38)$$

The attitude estimate is unbiased if the expectation value of the attitude estimation error vector, denoted by $E[\boldsymbol{\Theta}]$, vanishes. This is

$$E[\boldsymbol{\Theta}] = U_+ D^{-1} E[\mathbf{z}] = U_+ D^{-1} U_+^T \sum_{i=1}^{n} a_i[\mathbf{b}_i \times]E[A_{opt}\delta \mathbf{r}_i - \delta \mathbf{b}_i] . \qquad (39)$$

The errors in each observation vector, assuming their distribution to be axially symmetric about the underlying error-free vector, are shown in Appendix A to have the expectation values

$$E[\delta \mathbf{b}_i] = -\sigma_{bi}^2 \tau_{bi} \mathbf{b}_i \tag{40}$$

and

$$E[(\delta \mathbf{b}_i)(\delta \mathbf{b}_i)^T] = \sigma_{bi}^2 [I - (3 - 2\tau_{bi}) \mathbf{b}_i \mathbf{b}_i^T], \tag{41}$$

where σ_{bi}^2 is the variance of the observation errors in radians squared per axis, and $\tau_{bi} \geq 1$ tends to one as the errors tend to zero. The reference vector errors have the similar expectation values

$$E[\delta \mathbf{r}_i] = -\sigma_{ri}^2 \tau_{ri} \mathbf{r}_i , \tag{42}$$

and

$$E[(\delta \mathbf{r}_i)(\delta \mathbf{r}_i)^T] = \sigma_{ri}^2 [I - (3 - 2\tau_{ri}) \mathbf{r}_i \mathbf{r}_i^T]. \tag{43}$$

The statistical independence of the different errors gives

$$E[(\delta \mathbf{b}_i)(\delta \mathbf{r}_j)^T] = E[\delta \mathbf{b}_i] E[\delta \mathbf{r}_j^T], \tag{44}$$

$$E[(\delta \mathbf{b}_i)(\delta \mathbf{b}_j)^T] = E[\delta \mathbf{b}_i] E[\delta \mathbf{b}_j^T], \quad \text{for } i \neq j, \tag{45}$$

and

$$E[(\delta \mathbf{r}_i)(\delta \mathbf{r}_j)^T] = E[\delta \mathbf{r}_i] E[\delta \mathbf{r}_j^T], \quad \text{for } i \neq j. \tag{46}$$

Using equations (40) and (42) in equation (39) gives a vanishing expectation value, since both terms contain the product $[\mathbf{b}_i \times]\mathbf{b}_i = 0$, either explicitly or implicitly from equation (37). This proves that the attitude estimate is unbiased, to first order in the errors of the observation and reference vectors.

The covariance of the estimation error angle vector in the spacecraft body frame is

$$P_{body} = E[\boldsymbol{\Theta}\boldsymbol{\Theta}^T] = U_+ D^{-1} E[\mathbf{z}\mathbf{z}^T] D^{-1} U_+^T ; \tag{47}$$

this matrix is denoted $P_{\Theta\Theta}$ by Shuster, who was the first to provide a statistical analysis of this problem [5]. With equation (38),

$$E[\mathbf{z}\mathbf{z}^T] = U_+^T \sum_{i,j=1}^{n} a_i a_j [\mathbf{b}_i \times] E[(A_{opt} \delta \mathbf{r}_i - \delta \mathbf{b}_i)(A_{opt} \delta \mathbf{r}_j - \delta \mathbf{b}_j)^T] [\mathbf{b}_j \times]^T U_+ . \tag{48}$$

The only nonvanishing contributions to this expectation value are those arising from the identity matrix terms in equations (41) and (43); all other terms contain the product $[\mathbf{b}_i \times]\mathbf{b}_i = 0$, either explicitly or implicitly from equation (37). Thus the expectation value is

$$E[\mathbf{z}\mathbf{z}^T] = U_+^T \sum_{i=1}^{n} a_i^2 (\sigma_{bi}^2 + \sigma_{ri}^2) [\mathbf{b}_i \times][\mathbf{b}_i \times]^T U_+ . \tag{49}$$

Using the identity

$$[\mathbf{b}_i \times][\mathbf{b}_i \times]^T = I - \mathbf{b}_i \mathbf{b}_i^T = I - \mathbf{b}_i \mathbf{r}_i^T V_+ U_+^T \tag{50}$$

and assuming that the weights are chosen to be

$$a_i = \sigma_{tot}^2 / (\sigma_{bi}^2 + \sigma_{ri}^2) . \tag{51}$$

where

$$\sigma_{tot}^2 \equiv \left[\sum_{i=1}^{n} (\sigma_{bi}^2 + \sigma_{ri}^2)^{-1}\right]^{-1}, \tag{52}$$

we find that

$$E[\mathbf{z}\mathbf{z}^T] = \sigma_{tot}^2 \sum_{i=1}^{n} a_i(I - U_+^T \mathbf{b}_i \mathbf{r}_i^T V_+) = \sigma_{tot}^2(I - U_+^T B V_+) = \sigma_{tot}^2(I - S'). \tag{53}$$

Substituting equation (53) into equation (47) gives the final result for the covariance in the body frame

$$P_{body} = U_+ P_s U_+^T = U P_s U^T, \tag{54}$$

where the diagonal matrix

$$P_s \equiv \sigma_{tot}^2(I - S')D^{-2} \tag{55}$$

is the covariance matrix of the estimation error angle components in the S-frame. The covariance matrix of the estimation error angle components in the reference frame has the equally simple form

$$P_{ref} = A_{opt}^T P_{body} A_{opt} = V_+ P_s V_+^T = V P_s V^T. \tag{56}$$

The commutativity of the multiplication of diagonal matrices has been used several times in equations (54)–(56). We can now see the significance of the S-frame; it is a frame in which the components of the attitude estimation error vector are uncorrelated. In the case that the attitude is not uniquely determined, equation (19) represents a rotation by angle Φ about the S-frame axis with infinite covariance.

The derivation assumes the covariance matrix to be computed from the underlying error-free reference and observation vectors, which are not known. In practice, we must compute the covariance using the measured vectors containing errors; this will be indistinguishable from the true covariance if the loss function is negligible compared to the smallest diagonal element of D, which means that the attitude is well determined. Alternate forms of the covariance that are identical if they are computed from error-free reference and observation vectors give different generalizations if the loss function is not negligible. This is discussed, with an example, in Appendix B.

Equations (3), (5), and (6) give a useful inequality for the singular values:

$$s_1 + s_2 + s_3 = \text{tr } S = \text{tr}(U^T B V) = \text{tr}\left[\sum_{i=1}^{n} a_i (U^T \mathbf{b}_i)(V^T \mathbf{r}_i)^T\right]$$
$$= \sum_{i=1}^{n} a_i (V^T \mathbf{r}_i)^T (U^T \mathbf{b}_i) \leq 1. \tag{57}$$

The last step reflects the fact that the inner product of two unit vectors, such as $V^T\mathbf{r}_i$ and $U^T\mathbf{b}_i$, cannot be greater than unity. We define the optimal measurement geometry to be the arrangement of measurement and observation vectors that minimizes the trace of the covariance matrix, which is equal to the sum of the diagonal elements of P_s, for a given total measurement error σ_{tot}^2. Then the optimal geometry consistent with equation (57) is given by

$$s_1 = s_2 = s_3 = 1/3, \qquad d = 1, \qquad P_s = \sigma_{tot}^2 \,\mathrm{diag}(3/2, 3/2, 3/2)\,. \tag{58}$$

The optimal geometry for a rank-two B matrix is

$$s_1 = s_2 = 1/2, \qquad s_3 = 0, \qquad P_s = \sigma_{tot}^2 \,\mathrm{diag}(2, 2, 1)\,. \tag{59}$$

This reiterates that B is not required to have full rank to give a good attitude estimate.

Implementation

The proposed solution of Wahba's problem has the following steps:

(a) Compute B from equation (5)
(b) Find the SVD, equation (6), of B. This is the only computationally expensive step.
(c) Compute d from equation (14)
(d) Compute A_{opt} from equation (18)
(e) Compute $L(A_{opt})$ and P_s from equations (17), (34), and (55), and other statistics of interest.

This method was tested on a DEC VAX 11/780, using the LINPACK subroutine DSVDC [9] for the singular value decomposition. A listing of the FORTRAN 77 implementation is given in [23]. The iterative QR method with shifts, employed in DSVDC, converged with fewer iterations after a sign error in the shift computation was corrected; this gave the performance presented below. In principle, the QR method will converge in a single iteration if the shift used is an exact eigenvalue of the matrix B^TB [6], so DSVDC was modified to shift by the exact solution of the cubic characteristic equation of B^TB. This method failed to converge in a single iteration due to numerical inaccuracies, so it was abandoned, being no faster on the average than the unmodified DSVDC.

The tests used the following sets of reference vectors:

$$\mathbf{r}_1 = [1, 0.01, 0], \qquad \mathbf{r}_2 = [1, -0.01, 0] \qquad \text{(nearly collinear)}, \tag{60}$$

$$\mathbf{r}_1 = [1, 0, 0], \qquad \mathbf{r}_2 = [0, 1, 0] \qquad \text{(orthogonal)}, \tag{61}$$

$$\mathbf{r}_1 = [1, 0, 0], \qquad \mathbf{r}_2 = [0, 1, 0], \qquad \mathbf{r}_3 = [0, 0, 1] \qquad \text{(orthogonal)}, \tag{62}$$

$$\mathbf{r}_1 = [1, 0, 0], \qquad \mathbf{r}_2 = [-0.5, 0.8, 0], \qquad \mathbf{r}_3 = [-0.5, -0.8, 0] \qquad \text{(coplanar)}\,. \tag{63}$$

Larger sets of reference vectors were formed by repeating these vectors, modeling repeated measurements of two or three references. The observation vectors were given by equation (2) with

$$A_{true} = \begin{bmatrix} 0.352 & 0.864 & 0.360 \\ -0.864 & 0.152 & 0.480 \\ 0.360 & -0.480 & 0.800 \end{bmatrix}, \tag{64}$$

which was chosen to have all nonzero matrix elements with exact decimal representations, and is otherwise arbitrary. The components of the error vectors ϵ_i were given by random noise uniformly distributed between ± 0.01. All $\mathbf{b}_i$ and $\mathbf{r}_i$ were converted to true unit vectors and all observations were equally weighted.

For the test cases specified by equations (60)–(64) with n observations, the SVD method gave execution times, obtained by averaging over 1000 calls for each case, of:

$$t_{SVD} = [11 + 0.5(n - 2)] \text{ msec} \tag{65}$$

with only two distinct reference vectors (equations (60) and (61)),

$$t_{SVD} = [21 + 0.5(n - 3)] \text{ msec} \tag{66}$$

with a B matrix of full rank (equation (62)), and intermediate values with three coplanar reference vectors (equation (63)). For these test cases, the SVD method always converged to the optimal estimate.

Shuster's implementation of Davenport's q-method was applied to the same test cases, giving execution times of

$$t_q = [2.3 + 0.5(n - 2) + 0.2n_{iter}] \text{ msec}, \tag{67}$$

where n_{iter} is the number of Newton-Raphson iterations, usually 1 or 2, used to compute the maximum eigenvalue of the 4×4 matrix. This algorithm also converged to the optimal estimate for all the test cases. It is significantly faster than the SVD method due to the fact that it contains no iterative matrix computations, only a scalar iteration for the maximum eigenvalue. Other values of A_{true} and ϵ_i were tested with similar results, except that the SVD method showed shorter execution times when $\epsilon_i \equiv 0$ and A_{true} had zero elements, resulting in zero matrix elements in B.

Conclusions

A new method for finding the attitude matrix minimizing Wahba's loss function has been presented. This method, based on the singular value decomposition of a 3×3 matrix, requires no approximations, is numerically stable for matrices of rank two or three, and gives a convenient method for computing the covariance of the attitude estimate. It shares these advantages with Shuster's implementation of Davenport's q-method, which is faster than the new method. The singular value decomposition method has the advantage of giving the eigenvalues and eigenvectors of the covariance matrix, which are often useful for analysis, since the maximum eigenvalue and its eigenvector give the magnitude and direction of the largest component of the attitude uncertainty.

Appendix A–Unit Vector Error Model

Let $\mathbf{x}_{true}$ be an error-free unit vector and choose a coordinate system such that

$$\mathbf{x}_{true} = [0, 0, 1]^T. \tag{A1}$$

The corresponding vector containing measurement errors is

$$\mathbf{x}_{measured} = [\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta]^T \tag{A2}$$

with some probability represented by a density function $\rho(\theta, \phi)$ over the unit sphere.

The polar coordinate θ should be distinguished from the estimation error angle vector $\mathbf{\Theta}$ defined in the text. The error in $\mathbf{x}$ is

$$\delta\mathbf{x} \equiv \mathbf{x}_{measured} - \mathbf{x}_{true} = [\sin\theta\cos\phi, \sin\theta\sin\phi, -(1 - \cos\theta)]^T . \quad \text{(A3)}$$

We assume that the error distribution is axially symmetric about $\mathbf{x}_{true}$, which means that the probability distribution is independent of ϕ; this assumption is discussed by Shuster [5]. Averages over ϕ are trivial in this case, and we need the following averages over θ;

$$\sigma^2 \equiv \tfrac{1}{2}E[\sin^2\theta] , \quad \text{(A4)}$$

$$\tau \equiv E[1 - \cos\theta]/\sigma^2 , \quad \text{(A5)}$$

and

$$E[(1 - \cos\theta)^2] = E[2(1 - \cos\theta) - \sin^2\theta] = 2\sigma^2(\tau - 1) . \quad \text{(A6)}$$

With these relations, it is easy to show that

$$E[\delta\mathbf{x}] = -\sigma^2\tau\mathbf{x}_{true} \quad \text{(A7)}$$

and

$$E[(\delta\mathbf{x})(\delta\mathbf{x})^T] = \sigma^2[I - (3 - 2\tau)\mathbf{x}_{true}\mathbf{x}_{true}^T] . \quad \text{(A8)}$$

Equations (A7) and (A8) are expressed in a form that is independent of the coordinate system. The only assumption needed in this derivation was axial symmetry; nothing was assumed about the estimation error distribution in θ.

For small errors, the probability will be concentrated near $\theta = 0$, and σ^2 will tend to zero. In this limit σ^2 can be interpreted as the variance of the angular errors about the two axes perpendicular to $\mathbf{x}_{true}$ in radians squared per axis. Equation (A6) shows that $\tau \geq 1$, and small angle approximations for sine and cosine can be used to show that τ tends to unity in the limit of vanishing errors. The expectation value of $1 - \cos\theta$ vanishes only if the probability is concentrated at $\theta = 0$, so $\delta\mathbf{x}$ cannot be unbiased except in the limit of zero errors; this follows from the constraint that both $\mathbf{x}_{measured}$ and $\mathbf{x}_{true}$ are unit vectors. Equation (A7) shows that for small errors the expectation value of the error vector is proportional to the variance of the angular errors, which is to say that it is of second order in the standard deviation of the angular errors.

Appendix B–Alternate Forms of the Covariance

If the B matrix is formed from error-free reference and observation vectors, then $A_{opt} = A_{true}$, the loss function is zero, and

$$D = I - S' \quad \text{(B1)}$$

from equation (34). The matrix P_s is equal to

$$P_s = \sigma_{tot}^2 D^{-1} \quad \text{(B2)}$$

and

$$P_s = \sigma_{tot}^2 (I - S')^{-1} . \quad \text{(B3)}$$

Shuster's expression for the covariance matrix is [5]

$$P_{\theta\theta} = \sigma_{tot}^2 (I - M)^{-1}, \tag{B4}$$

where M is the matrix defined in equation (22). Equations (15), (18), and (21) give, with no errors and $A_{opt} = A_{true}$

$$P_{\theta\theta} = U_+[\sigma_{tot}^2 (I - S')^{-1}] U_+^T, \tag{B5}$$

which is identical to the covariance given by equations (54) and (B3). Thus we have four different forms of the covariance that are identical if they are computed from error-free reference and observation vectors. When computed from measured vectors containing errors, they give indistinguishable results if the loss function is negligible compared to the smallest diagonal element of D, which means that the attitude is well determined. Equation (55) gives the best estimate of the errors to be expected, though, when the loss function is not negligible, as we will show with an example.

We will compute the covariance matrix for a simple case of two vector observations with identical noise characteristics, and therefore equal weights. The reference and observation vectors can be written without loss of generality as

$$\mathbf{b}_1 = R_b[\cos\beta, \sin\beta, 0]^T, \qquad \mathbf{b}_2 = R_b[\cos\beta, -\sin\beta, 0]^T, \tag{B6}$$

$$\mathbf{r}_1 = R_r[\cos\alpha, \sin\alpha, 0]^T, \qquad \mathbf{r}_2 = R_r[\cos\alpha, -\sin\alpha, 0]^T, \tag{B7}$$

where R_b and R_r are proper orthogonal matrices and α and β are angles between zero and $\pi/2$. Substitution of these vectors into equation (5) gives equation (15) with

$$S' = \mathrm{diag}(\cos\alpha\cos\beta, \sin\alpha\sin\beta, 0), \tag{B8}$$

$$U_+ = R_b, \quad \text{and} \quad V_+ = R_r. \tag{B9}$$

The singular values s_1 and s_2 may not be in the order specified by equation (8), but this is of no importance. The minimized Wahba loss function for this example is

$$L(A_{opt}) = 1 - \cos(\alpha - \beta), \tag{B10}$$

which is zero if and only if $\alpha = \beta$. This condition is equivalent to $\mathbf{b}_1 \cdot \mathbf{b}_2 = \mathbf{r}_1 \cdot \mathbf{r}_2$ and gives an A_{opt} such that $\mathbf{b}_i = A_{opt}\mathbf{r}_i$ for $i = 1$ and 2.

The diagonal elements of the matrix $\sigma_{tot}^{-2} U_+^T P_{body} U_+$ computed using equations (55), (B2), (B3), and (B4) are collected in Table B1. Some trigonometric identities have been employed to arrive at the forms displayed. All the entries in any column of this table are equal if $\alpha = \beta$. If $\alpha \approx \beta$, $K_i \approx 1/2$ for $i = 1, \ldots, 4$ and the differences in the entries in any column are small except for values of the 11 element when $\alpha \approx \beta \approx 0$ and values of the 22 element when $\alpha \approx \beta \approx \pi/2$. Both of these cases are for nearly collinear observations. Consider the 11 element; the argument for the 22 element is similar. Even though $\alpha \approx \beta$, the values of $1/\sin\alpha$ and $1/\sin\beta$ can differ by orders of magnitude if both of these angles are close to zero. We expect the estimation errors to be amplified if the geometry of either the measured vectors or the reference vectors is poor, that is if either pair is nearly collinear. In particular, if the covariance matrix is to give a useful estimate of the estimation errors to be expected with the actual measurements, its 11 element should have terms proportional to both $\sin^{-2}\alpha$ and $\sin^{-2}\beta$. The only form of the covariance in Table B1 with this property is the form given by equation (55), which is why we prefer this form to the alternatives.

Table B1. Alternate Forms of the Covariance for Two Observations

Equation	11 Element	22 Element	33 Element
(55)	$K_1 \sin^{-2}\alpha + K_2 \sin^{-2}\beta$	$K_3 \cos^{-2}\alpha + K_4 \cos^{-2}\beta$	$1/\cos^2(\alpha - \beta)$
(B2)	$(\sin\alpha \sin\beta)^{-1}$	$(\cos\alpha \cos\beta)^{-1}$	$1/\cos(\alpha - \beta)$
(B3)	$(K_1 \sin^2\beta + K_2 \sin^2\alpha)^{-1}$	$(K_3 \cos^2\beta + K_4 \cos^2\alpha)^{-1}$	1
(B4)	$\sin^{-2}\beta$	$\cos^{-2}\beta$	1

$K_1 \equiv \frac{1}{2}(1 + \cos^2\alpha)/(1 + \cos\alpha \cos\beta)$, $K_2 \equiv \frac{1}{2}(1 + \cos^2\beta)/(1 + \cos\alpha \cos\beta)$,
$K_3 \equiv \frac{1}{2}(1 + \sin^2\alpha)/(1 + \sin\alpha \sin\beta)$, $K_4 \equiv \frac{1}{2}(1 + \sin^2\beta)/(1 + \sin\alpha \sin\beta)$.

Acknowledgments

I would like to acknowledge many useful discussions with Malcolm Shuster and with Paul Davenport on attitude estimation in general and Wahba's problem in particular. Malcolm Shuster has guided me around, or helped me out of, several pitfalls in the covariance analysis of this problem. I would also like to thank Harry Frisch and Sally Godfrey for providing me with subroutines DSVDC and QUEST, respectively, and Rose Pajerski and Kellyann Franks for assistance on the VAX. I am especially grateful to Ed Seidewitz, whose critiques of earlier versions of this manuscript have contributed immeasurably to the present form of this paper.

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] SHUSTER, MALCOLM D. "A Comment on Fast Three-Axis Attitude Determination," *Journal of the Astronautical Sciences,* Vol. XXXI, No. 4, October–December 1983, pp. 579–583.

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

[5] 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.

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

[7] GOLUB, G. H. and REINSCH, C. "Singular Value Decomposition and Least Squares Solutions," *Numerische Mathematik,* Vol. 14, 1980, pp. 403–420.

[8] LAWSON, C. L. and HANSON, R. J. *Solving Least Squares Problems,* Prentice-Hall, Englewood Cliffs, 1974.

[9] DONGARRA, J. J., BUNCH, J. R., MOLER, C. B., and STEWART, G. W. *LINPACK User's Guide,* SIAM Publications, Philadelphia, Pennsylvania, 1979.

[10] WESSNER, R. H. "A Least Squares Estimate of Spacecraft Attitude," *SIAM Review,* Vol. 8, No. 3, July 1966, p. 386.

[11] VELMAN, J. R. ibid.

[12] BROCK, J. E. ibid.

[13] BROCK, JOHN E. "Optimal Matrices Describing Linear Systems," *AIAA Journal,* Vol. 6, No. 7, July 1968, pp. 1292–1296.

[14] DAVENPORT, PAUL B. *Attitude Determination and Sensor Alignment via Weighted Least Squares Affine Transformations,* NASA X-514-71-312, August 1971.

[15] BJÖRCK, A. and BOWIE, C. "An Iterative Algorithm for Computing the Best Estimate of an Orthogonal Matrix," *SIAM Journal on Numerical Analysis,* Vol. 8, 1971, pp. 358–364.

[16] BAR-ITZHACK, I. Y. and FEGLEY, K. A. "Orthogonalization Techniques of a Direction Cosine Matrix," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 5, 1969, pp. 798–804.

[17] BAR-ITZHACK, I. Y. "Iterative Optimal Orthogonalization of the Strapdown Matrix," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 11, No. 1, January 1975, pp. 30–37.

[18] BAR-ITZHACK, I. Y., MEYER, J., and FUHRMANN, P. A. "Strapdown Matrix Orthogonalization: the Dual Iterative Algorithm," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 12, 1976, pp. 32–37.

[19] BAR-ITZHACK, I. Y. and MEYER, J. "On the Convergence of Iterative Orthogonalization Processes," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 12, 1976, pp. 146–151.

[20] MEYER, J. and BAR-ITZHACK, I. Y. "Practical Comparison of Iterative Matrix Orthogonalization Algorithms," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 13, 1977, pp. 230–235.

[21] HIGHAM, NICHOLAS J. "Computing the Polar Decomposition-with Applications," *SIAM Journal on Scientific and Statistical Computing,* Vol. 7, No. 4, October 1986, pp. 1160–1174.

[22] TIETZE, JOHN L. "Fast Three-Axis Attitude Determination Using Vector Observations and Inverse Iteration," *Journal of the Astronautical Sciences,* Vol. XXX, No. 2, April–June 1982, pp. 171–179.

[23] MARKLEY, F. LANDIS "Three-Axis Attitude Determination Using Vector Observations and the Singular Value Decomposition," AAS Paper 87-490, *Proceedings of the AAS/AIAA Astrodynamics Specialist Conference,* Kalispell, Montana, August 1987.