

# Unconstrained Optimal Transformation Matrix

We generalize Brock's minimization of the loss function proposed by Wahba to give a least-squares estimate of satellite attitude. We follow Brock in relaxing Wahba's requirement that the minimizing attitude matrix be orthogonal, and discuss the statistical properties of our new solution and its relation to existing attitude estimation algorithms.

## INTRODUCTION

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

$$L(A) = \frac{1}{2} \sum_{j=1}^n \|\mathbf{v}_j - A\mathbf{u}_j\|^2 \quad (1)$$

where the unit vectors  $\mathbf{u}_j$  are representations in a reference frame of the directions to  $n$  observed objects and the  $\mathbf{v}_j$  are the unit vector representations of the corresponding observations in the spacecraft body frame. The motivation for this loss function is that if the vectors are error-free and the true attitude matrix  $A_{\text{true}}$  is assumed to be the same for all the measurements, then  $\mathbf{v}_j$  is equal to  $A_{\text{true}}\mathbf{u}_j$  for all  $j$  and the loss function is equal to zero for  $A$  equal to  $A_{\text{true}}$ . It is customary to generalize this loss function by introducing an  $n \times n$  symmetric positive-definite matrix  $W$  to weight the different observations and to rewrite it as

$$L(A) = \frac{1}{2} \text{tr}[W(AU - V)^T(AU - V)] \quad (2)$$

where  $\text{tr}$  denotes the matrix trace, the superscript  $T$  denotes the transpose, and the  $3 \times n$  matrices  $U$  and  $V$  are defined as

$$U \equiv [\mathbf{u}_1 : \mathbf{u}_2 : \cdots : \mathbf{u}_n] \quad (3)$$

and

$$V \equiv [\mathbf{v}_1 : \mathbf{v}_2 : \cdots : \mathbf{v}_n]. \quad (4)$$

The problem of finding the *constrained* (by orthogonality) minimum of this loss function has become known as Wahba's problem, and attitude

determination algorithms based on it have been used for many years [2–8]. Using the assumed orthogonality of  $A$  and the invariance of the trace under transposition and cyclic permutation of its arguments, the loss function can be written

$$\begin{aligned} L(A) &= \frac{1}{2} \text{tr}[W(U^T U - U^T A^T V - V^T A U + V^T V)] \\ &= \frac{1}{2} \text{tr}[W(U^T U + V^T V)] - \text{tr}(AB^T) \end{aligned} \quad (5)$$

where  $B$  is the attitude profile matrix [6]

$$B \equiv VWU^T. \quad (6)$$

This matrix contains sufficient information to define the optimal attitude. In fact, the optimal rotation matrix  $A$  is the orthogonal matrix that is the closest to  $B$  in the Euclidean (or Frobenius) norm [8–10]. The original solutions solved for the spacecraft attitude matrix directly [2], but most practical applications have been based on Davenport's  $q$ -method [3–6], which solves for the quaternion that parameterizes the attitude matrix. Some newer methods solve for the attitude matrix directly, though [7, 8]. These methods all make use of the fact that the quadratic term in  $A$  has been eliminated from (5) by use of the orthogonality constraint; conversely, these methods all require enforcement of the orthogonality constraint to provide a unique solution.

The loss function of (2) can be further generalized to

$$L(A) = \frac{1}{2} \text{tr}[W(AU - V)^T Z(AU - V)] \quad (7)$$

where  $Z$  is a  $3 \times 3$  symmetric positive-definite matrix that weights the Cartesian components of the observations in the body frame. The quadratic term in  $A$  cannot be eliminated from this loss function, even with the assumption of orthogonality, which makes it more difficult to analyze than Wahba's loss function. Brock [11] found a simple solution to this optimization problem when the orthogonality constraint is relaxed and the matrix  $W$  is equal to the identity, so that all measurements are weighted equally. Brock's treatment does not make it clear that  $Z$  is not a matrix that assigns different weights to different measurements, so confusion on this point can easily arise. We generalize his solution to explicitly consider an arbitrary positive-definite observation weight matrix  $W$ . The statistical properties of the new solution and its relation to existing algorithms are discussed. Similar results have previously been obtained by one of the authors [10].

## SOLUTION

Our derivation employs the concept of the directional derivative of the loss function. Let  $A_0$  denote an assumed extremum of the loss function, and consider a variation of an arbitrary size  $\epsilon$  in the direction of an arbitrary (non-zero) matrix  $H$ . The

Manuscript received October 12, 1995; revised April 30, 1997.

IEEE Log No. T-AES/34/1/00204.

Part of this work was performed on a NRC/NASA Goddard Space Flight Center Associateship.

U.S. Government work not protected by U.S. copyright.

0018-9251/98/\$10.00 © 1998 IEEE

value of the loss function is

$$\begin{aligned}
L(A_0 + \varepsilon H) &= \frac{1}{2} \text{tr}\{W[(A_0 + \varepsilon H)U - V]^T \\
&\quad \times Z[(A_0 + \varepsilon H)U - V]\} \\
&= \frac{1}{2} \text{tr}\{[Z^{1/2}(A_0 U - V)W^{1/2}]^T \\
&\quad \times [Z^{1/2}(A_0 U - V)W^{1/2}]\} \\
&\quad + \varepsilon \text{tr}[W(HU)^T Z(A_0 U - V) \\
&\quad + (HU)^T Z(A_0 U - V)W] \\
&\quad + \frac{1}{2} \varepsilon^2 \text{tr}[(Z^{1/2} H U W^{1/2})^T (Z^{1/2} H U W^{1/2})] \\
&= \frac{1}{2} \|Z^{1/2}(A_0 U - V)W^{1/2}\|_F^2 \\
&\quad + \varepsilon \text{tr}[H^T Z(A_0 U - V)W U^T] \\
&\quad + \frac{1}{2} \varepsilon^2 \|Z^{1/2} H U W^{1/2}\|_F^2 \quad (8)
\end{aligned}$$

where the subscript F denotes the Frobenius norm, and we have used the fact that  $W$  and  $Z$  have symmetric positive-definite square roots, since they are defined to be symmetric and positive definite. The extremum condition is that this loss function must have zero derivative with respect to  $\varepsilon$  at  $\varepsilon = 0$  for all  $H$ , which is the case if and only if the term linear in  $\varepsilon$  vanishes. Since  $Z$  is positive definite by assumption, this term vanishes for all  $H$  if and only if

$$(A_0 U - V)W U^T = 0. \quad (9)$$

It is interesting to observe that this condition is independent of the weighting matrix  $Z$ , as was noted by Brock [11]. If the matrix  $U$  has full rank (i.e., rank three), this equation can be solved for  $A_0$ , yielding

$$A_0 = V W U^T (U W U^T)^{-1} = B (U W U^T)^{-1} \quad (10)$$

where  $B$  is the attitude profile matrix defined in (6). In this case, the coefficient of  $\varepsilon^2$  in (8) is positive for all non-zero  $H$ , so the second directional derivative of  $L$  at  $A_0$  with respect to  $\varepsilon$  is positive, and the stationary point at  $A_0$  is clearly a unique global minimum of the loss function. This estimate is not independent of  $W$ , in general.

## STATISTICAL PROPERTIES OF THE SOLUTION

We now show that this solution provides an unbiased estimate of the attitude matrix, under reasonable statistical assumptions. We consider the reference vectors to be error-free, which is generally a good approximation since the reference vectors are usually known quite precisely. We also make the common assumption of additive random measurement errors, i.e.,

$$\mathbf{v}_j = A_{\text{true}} \mathbf{u}_j + \mathbf{n}_j \quad \text{for all } j \quad (11)$$

or equivalently

$$V = A_{\text{true}} U + N \quad (12)$$

with

$$N \equiv [\mathbf{n}_1 : \mathbf{n}_2 : \cdots : \mathbf{n}_n]. \quad (13)$$

We assume that the random errors have zero mean, which means that

$$E[N] = 0 \quad (14)$$

where  $E[\cdot]$  denotes the expected value. This assumption is not restrictive, since any non-zero average errors can be computed and removed. The attitude estimate is given by

$$\begin{aligned}
A_0 &= (A_{\text{true}} U + N) W U^T (U W U^T)^{-1} \\
&= A_{\text{true}} + \delta A \quad (15)
\end{aligned}$$

with

$$\delta A \equiv N W U^T (U W U^T)^{-1}. \quad (16)$$

It is easy to see from (14) that the expected value of  $\delta A$  is zero, so taking the expectation of (15) shows that  $A_0$  is an unbiased estimate of the transformation matrix.

An estimate of the dispersion of  $A_0$  about its expected value,  $A_{\text{true}}$ , is provided by the quantity

$$\begin{aligned}
P &\equiv E[(\delta A)^T \delta A] \\
&= (U W U^T)^{-1} U W R W U^T (U W U^T)^{-1} \quad (17)
\end{aligned}$$

where

$$R \equiv E[N^T N]. \quad (18)$$

The matrix  $P$  looks like a covariance matrix since  $\delta A$  has zero mean, but the actual covariance would be a  $9 \times 9$  matrix, since  $\delta A$  has nine components. It is natural to choose the weight matrix  $W$  to be the inverse of the  $n \times n$  measurement variance matrix  $R$

$$W = R^{-1} \quad (19)$$

which gives

$$P = (U W U^T)^{-1} = (U R^{-1} U^T)^{-1}. \quad (20)$$

If we assume that  $\{\mathbf{n}_j\}$  is a white sequence, then  $R$  and  $W$  are diagonal matrices.

The optimal matrix of (10) is not exactly orthogonal, but a simple calculation gives the deviation from orthogonality as

$$\begin{aligned}
A_0^T A_0 - I &= (A_{\text{true}} + \delta A)^T (A_{\text{true}} + \delta A) - I \\
&= (A_{\text{true}})^T \delta A + (\delta A)^T A_{\text{true}} + (\delta A)^T \delta A. \quad (21)
\end{aligned}$$

It follows from (17) and the zero-mean property of  $\delta A$  that

$$E[A_0^T A_0 - I] = P \quad (22)$$

which shows that the average deviation from orthogonality is directly related to the dispersion of the estimate about the mean.

## DISCUSSION

This generalization of Brock's solution is computationally simpler than the solution of Wahba's

problem in the general case, since it involves only inversion of a  $3 \times 3$  matrix, and no iterative computations. Thus it may be preferable if small deviations from orthogonality can be tolerated. This estimate can be useful for extracting Euler angles or quaternions for display, or as an initial estimate for a more refined method.

One disadvantage of (10) is that it provides a solution only if  $U$  is of rank three, which is to say that the vectors  $\{\mathbf{u}_j\}$  span three-space. This requires as a minimum that we have at least three observations. It is well known that two observations are sufficient to completely determine an orthogonal transformation matrix, but the orthogonality constraint plays a key role in establishing this sufficiency. In the unconstrained problem, the estimate can be made unique by the somewhat *ad hoc* procedure of adding as a third observation the cross-product  $\mathbf{v}_1 \times \mathbf{v}_2$ , regarded as a measurement of a reference  $\mathbf{u}_1 \times \mathbf{u}_2$ . The added measurement is a pseudomeasurement, and not a real measurement. Thus it does not require additional sensors or sensor data processing, but it does provide a unique estimate with only two real observations.

The three-observation case has a solution if and only if the three vectors are not coplanar. If the third vector is given as a cross-product pseudomeasurement, the requirement is that the two actual measurements not be colinear. In all these cases the matrix  $U$  is invertible, so

$$\begin{aligned} A_0 &= VWU^T(UWU^T)^{-1} \\ &= VWU^T U^{-T} W^{-1} U^{-1} = VU^{-1} \end{aligned} \quad (23)$$

which is independent of the measurement weights. This reflects the fact that the matrix  $A$  has enough degrees of freedom to produce a minimum value of zero for the loss function in the three-observation case, *if the orthogonality constraint is not enforced*. The weight matrix is significant if four or more measurements are used to form the estimate, however.

In contrast to the above, Wahba's problem of finding the optimal *orthogonal*  $A$  minimizing the loss function of (2) has a unique solution if the matrix  $B$  defined in (6) has rank two or three [7]. In the two-observation case, the cross-product of the reference vectors and the cross-product of the measurement vectors appear automatically, without requiring manual insertion [8]. This optimal estimate in the two-observation case has a closed-form solution that only requires a single square root [6, 8]. This estimate always depends on the weights, except when the measurements are error-free. This solution goes over to the simpler TRIAD solution [5, 6, 12] for certain values of the weights, so the TRIAD solution is optimal in these cases [8]. For arbitrary weights, the TRIAD solution does not minimize Wahba's loss function, however, and the closed-form solution to Wahba's problem is clearly the method of choice.

## ACKNOWLEDGMENT

We would like to thank John Crassidis for his many helpful comments on this work.

**F. LANDIS MARKLEY**  
Guidance, Navigation, and Control Branch  
NASA Goddard Space Flight Center  
Code 712  
Greenbelt, MD 20771

**ITZHACK Y. BAR-ITZHACK**  
Faculty of Aerospace Engineering  
Technion-Israel Institute of Technology  
Haifa 32000, Israel

## REFERENCES

- [1] Wahba, G. (1965)  
A least squares estimate of spacecraft attitude.  
*SIAM Review*, 7, 3 (July 1965), 409.
- [2] Farrell, J. L., & Stuelpnagel, J. C.; Wessner, R. H.; Velman, J. R.; Brock, J. E. (1966)  
A least squares estimate of spacecraft attitude.  
*SIAM Review*, 8, 3 (July 1966), 384-386.
- [3] Davenport, P. B. (1965)  
A vector approach to the algebra of rotations with applications.  
NASA, X-546-65-437, Nov. 1965.
- [4] Keat, J. (1977)  
Analysis of least squares attitude determination routine DOAOP.  
Computer Sciences Corporation, CSC/TM-77/6034, Feb. 1977.
- [5] Lerner, G. M. (1978)  
Three-axis attitude determination.  
In J. R. Wertz (Ed.), *Spacecraft Attitude Determination and Control*.  
Dordrecht, Holland: D. Reidel, 1978.
- [6] Shuster, M. D., and Oh, S. D. (1981)  
Three-axis attitude determination from vector observations.  
*Journal of Guidance and Control*, 4, 1 (Jan.-Feb. 1981), 70-77.
- [7] Markley, F. L. (1988)  
Attitude determination using vector observations and the singular value decomposition.  
*Journal of the Astronautical Sciences*, 36, 3 (July-Sept. 1988), 245-258.
- [8] Markley, F. L. (1993)  
Attitude determination using vector observations: A fast optimal matrix algorithm.  
*Journal of the Astronautical Sciences*, 41, 2 (Apr.-June 1993), 261-280.
- [9] Davenport, P. B. (1971)  
Attitude determination and sensor alignment via weighted least squares affine transformations.  
AAS/AIAA Astrodynamics Specialists Conference, Fort Lauderdale, FL, Aug. 1971, Paper 71-396.
- [10] Bar-Itzhack, I. (1992)  
Polar decomposition for attitude determination from vector observations.  
Flight Mechanics/Estimation Theory Symposium, Goddard Space Flight Center, Greenbelt, MD, May 1992, NASA Conference Publication 3186, 243-257.
- [11] Brock, J. E. (1968)  
Optimal matrices describing linear systems.  
*AIAA Journal*, 6, 7 (July 1968), 1292-1296.
- [12] Black, H. D. (1964)  
A passive system for determining the attitude of a satellite.  
*AIAA Journal*, 2, 7 (July 1964), 1350-1351.