The Journal of the Astronautical Sciences, Vol. 37, No. 1, January-March 1989, pp. 79–88

Maximum Likelihood Estimation of Spacecraft Attitude

Malcolm D. Shuster[1]

Abstract

The Wahba problem, which has been the starting point for a number of attitude estimation algorithms, most notably QUEST, is shown for an appropriate choice of weights and measurement model to be equivalent to a maximum likelihood estimation problem for the attitude. The measurement model for which this is true turns out to be the same as was used in earlier covariance analyses of the QUEST algorithm. The QUEST covariance matrix now emerges in a natural way as the inverse of the Fisher information matrix for the maximum likelihood estimator. The 3×3 attitude profile matrix of the QUEST algorithm is shown to be a useful representation for both the attitude *and* the attitude covariance.

Introduction

A topic of continuing interest has been the computation of the least-squares attitude matrix, A^*, which minimizes the cost function

$$L(A) = \frac{1}{2} \sum_{k=1}^{n} a_k |\hat{\mathbf{W}}_k - A\hat{\mathbf{V}}_k|^2 \tag{1}$$

where $\hat{\mathbf{W}}_k$, $k = 1, \ldots, n$, is a set of unit vector observations in the spacecraft-fixed reference frame, and $\hat{\mathbf{V}}_k$, $k = 1, \ldots, n$, are the representations of the same unit vectors with respect to the primary reference frame (the frame to which the attitude is referred). The a_k are a set of positive weights. Provided that at least two of the observation vectors are not parallel (nor anti-parallel) a unique minimizing attitude matrix will always exist. This cost function was first proposed by G. Wahba [1] in 1965.

Since that time more than a dozen solutions have been offered for this problem. Reviews of this work have appeared in Shuster [2] and Markley [3]. The most efficient implementation to date is that of Shuster and Oh [4], which has been used in the support of several NASA spacecraft.

[1]Senior Professional Staff, Space Department, The Johns Hopkins University, Applied Physics Laboratory, Laurel, Maryland 20707-6099.

While complete covariance analyses have been carried out for this algorithm [3,4] based on assumed error models for the spacecraft sensors, no attempt has been made to regard the Wahba problem as anything more than a naive weighted-least-squares problem. It turns out, in fact, for a particular choice of the a_k, namely, the choice suggested by weighted-least-squares arguments, that the Wahba attitude may be shown to be the maximum-likelihood estimate of the attitude given the same measurement model as has been used in covariance analyses of the algorithm.

The next section of this report presents the error model for unit-vector measurements which was used in earlier covariance analyses of the Wahba problem as well as other studies. Following this, it is shown that the Wahba attitude is indeed the maximum likelihood estimate of the attitude given this measurement model. The QUEST batch algorithm for computing the Wahba attitude, which forms the computational background of the present study, is then reviewed and the connection of the QUEST covariance matrix to the Fisher information matrix of maximum likelihood estimation is demonstrated.

The Measurement Model

Line-of-sight measurement sensors most often measure the direction of a celestial body by measuring the angle of that body from the sensor boresight in two mutually orthogonal planes. Thus, if we choose the z-axis of the sensor coordinate system to be directed outward along the boresight, then the measured quantities are effectively

$$\tan\alpha = \frac{\hat{\mathbf{W}}\cdot\hat{\mathbf{x}}}{\hat{\mathbf{W}}\cdot\hat{\mathbf{z}}}, \qquad \tan\beta = \frac{\hat{\mathbf{W}}\cdot\hat{\mathbf{y}}}{\hat{\mathbf{W}}\cdot\hat{\mathbf{z}}}, \tag{2}$$

and the reconstructed unit vector in sensor coordinates is

$$\hat{\mathbf{W}} = \frac{1}{\sqrt{1+\tan^2\alpha+\tan^2\beta}}\begin{bmatrix}\tan\alpha\\ \tan\beta\\ 1\end{bmatrix}. \tag{3}$$

Fixed-head star trackers and vector sun sensors work in this way.

We could, if we wished, base our maximum likelihood estimation scheme upon the measurements described in equation (2). This would certainly be the most rigorous approach. However, it does not lead directly to the most efficient estimator. Instead we take the point of view that $\hat{\mathbf{W}}_k$ *in body coordinates* is the measurement provided by sensor k and take the unit-vector measurement to have a probability density given by

$$p_{\hat{\mathbf{W}}_k}(\hat{\mathbf{W}}_k';A) = \mathcal{N}_k \exp\left\{-\frac{1}{2\sigma_k^2}|\hat{\mathbf{W}}_k' - A\hat{\mathbf{V}}_k|^2\right\}, \tag{4}$$

which is defined over the unit sphere,

$$|\hat{\mathbf{W}}_k'| = 1\,. \tag{5}$$

$\hat{\mathbf{W}}_k'$ is the realization of the random variable $\hat{\mathbf{W}}_k$, which also satisfies equation (5). $\hat{\mathbf{V}}_k$ is not the realization of a random variable in the present formulation and, hence, need not be distinguished by a prime. $\mathcal{N}_k$ is chosen so that the total probability will be unity. Thus,

$$\mathcal{N}_k^{-1} = \int_0^{2\pi}\int_0^{\pi} \exp\left\{-\frac{1}{\sigma_k^2}(1 - \cos\theta)\right\} \sin\theta \, d\theta \, d\phi \,, \tag{6}$$

which leads to

$$\mathcal{N}_k = \left[2\pi\sigma_k^2\left(1 - e^{-2/\sigma_k^2}\right)\right]^{-1}. \tag{7}$$

The exponential term in equation (7) is negligible. For σ_k even 1 degree, $e^{-2/\sigma_k^2} \simeq 3.9 \times 10^{-2852}$.

The above distribution makes sense intuitively. Since nearly all the probability is concentrated on a very small area about the direction $A\hat{\mathbf{V}}_k$, we may approximate the sphere near that point by a tangent plane given by

$$\mathbf{W}_k = A\hat{\mathbf{V}}_k + \Delta\mathbf{W}_k, \qquad \Delta\mathbf{W}_k \cdot A\hat{\mathbf{V}}_k = 0\,. \tag{8}$$

The sensor error $\Delta\mathbf{W}_k$ is approximately Gaussian and satisfies

$$E\{\Delta\mathbf{W}_k\} = \mathbf{0}\,, \tag{9}$$

$$E\{\Delta\mathbf{W}_k\,\Delta\mathbf{W}_k^T\} = \sigma_k^2[I - (A\hat{\mathbf{V}}_k)(A\hat{\mathbf{V}}_k)^T]\,, \tag{10}$$

where the superscript T denotes the matrix transpose. This approximate but very accurate model for the mean and covariance matrix of the sensor error was the basis for the covariance analyses of the QUEST and TRIAD algorithms in [4].

Markley [3] has shown for any probability density for $\hat{\mathbf{W}}_k$ (defined on the unit sphere) which is axially symmetric about $A\hat{\mathbf{V}}_k$ that

$$E\{\Delta\hat{\mathbf{W}}_k\} = -\rho_k^2\tau_k A\hat{\mathbf{V}}_k\,, \tag{11}$$

$$E\{\Delta\hat{\mathbf{W}}_k\,\Delta\hat{\mathbf{W}}_k^T\} = \rho_k^2[I - (3 - 2\tau_k)(A\hat{\mathbf{V}}_k)(A\hat{\mathbf{V}}_k)^T]\,, \tag{12}$$

where

$$\rho_k^2 = \frac{1}{2}E\{|\hat{\mathbf{W}}_k \times (A\hat{\mathbf{V}}_k)|^2\}\,, \tag{13a}$$

$$\tau_k = \frac{1}{\rho_k^2}E\{1 - \hat{\mathbf{W}}_k \cdot A\hat{\mathbf{V}}_k\}\,. \tag{13b}$$

For the particular probability density of equation (4), the respective values are

$$\rho_k^2 = \sigma_k^2 - \sigma_k^4 + O(e^{-2/\sigma_k^2})\,, \tag{14a}$$

$$\tau_k = \frac{1}{1 - \sigma_k^2} + O(e^{-2/\sigma_k^2})\,. \tag{14b}$$

For error levels of 1 degree ($\simeq .017$ rad), $\sigma^2 \simeq 3 \times 10^{-4}$, so that from a practical standpoint, the probability density on the sphere is indistinguishable from the corresponding density on the tangent plane. Both models, in fact, differ less from each other than does each from a more realistic parameterization of the measurement probability based on equation (2), though, of course, it is always possible to find a probability density function for $\tan\alpha$ and $\tan\beta$ which is identically equivalent to the density given by equation (4).

The Maximum Likelihood Estimator

Given the measurement model described by equation (4), the maximum likelihood estimation of the attitude is straightforward. If $\mathbf{Z}_k'$, $k = 1, \ldots, n$, is a sequence of measurements and $p_{z_1,\ldots,z_n}(\mathbf{Z}_1', \ldots, \mathbf{Z}_n'; \mathbf{x})$ is the joint probability distribution of the measurements as a function of a parameter vector $\mathbf{x}$, then the maximum likelihood estimate [5, 6] of $\mathbf{x}$ is given by

$$x_{ML}^{*\prime} = \arg \max_{x} p_{z_1,\ldots,z_n}(\mathbf{Z}_1', \ldots, \mathbf{Z}_n'; \mathbf{x}), \tag{15}$$

that is, the value of $\mathbf{x}$ at which $p_{z_1,\ldots,z_n}(\mathbf{Z}_1', \ldots, \mathbf{Z}_n'; \mathbf{x})$ achieves its maximum.[2] The maximum likelihood estimate, $\mathbf{x}_{ML}^{*\prime}$, is a function of the values of the measurements, $\mathbf{Z}_1', \ldots, \mathbf{Z}_n'$. We reserve the notation $\mathbf{x}_{ML}^{*}$ without the prime for the maximum likelihood estimator, a random variable, which depends on the measurement random variables, $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$, in the same way as the maximum likelihood estimate depends on the values (realizations) of the measurements.

For the model given by equation (4)

$$p_{z_1,\ldots,z_n}(\mathbf{Z}_1', \ldots, \mathbf{Z}_n'; A) = \prod_{k=1}^{n} \frac{1}{2\pi\sigma_k^2 f_k} e^{-|\hat{\mathbf{W}}_k' - A\hat{\mathbf{V}}_k|^2/2\sigma_k^2}, \tag{16}$$

where

$$f_k \equiv 1 - e^{-2/\sigma_k^2}, \tag{17}$$

is the "truncation defect" representing the additional normalization due to the finite area of the unit sphere.

Defining now

$$J(\mathbf{x}) = -\log p_{z_1,\ldots,z_n}(\mathbf{Z}_1', \ldots, \mathbf{Z}_n'; \mathbf{x}), \tag{18}$$

the negative-log-likelihood function, it follows that

$$J(A) = \sum_{k=1}^{n} \left\{ \frac{1}{2\sigma_k^2} |\hat{\mathbf{W}}_k' - A\hat{\mathbf{V}}_k|^2 + \log \sigma_k^2 + \log 2\pi + \log f_k \right\} \tag{19}$$

will be a minimuim at the maximum likelihood estimate of the attitude. If the identification

$$a_k = \frac{1}{\sigma_k^2} \tag{20}$$

is now made, and it is noted that only the first term in the brace (the data-dependent part of the negative-log-likelihood function) above depends on the attitude, then the negative-log-likelihood function for this measurement model is equivalent to the cost function of Wahba. Thus, the Wahba attitude is the maximum likelihood estimate of the attitude corresponding to the measurement error model give by equation (4). As a consequence of this, the covariance matrix of the Wahba attitude should be calculable directly from the Fisher information matrix [5, 6]. This, as we shall see, is true.

[2]Note that an asterisk is used to designate the estimate or estimator so as not to be confused with the caret used to designate a unit vector.

The same negative-log-likelihood function, apart from the negligible term $\log f_k$, is obtained also from the "planar" probability distribution described by equations (9)–(11). This, in fact, was the distribution used by the author in earlier derivations [7] of the Wahba problem as a maximum likelihood problem.

The Quest Algorithm[3]

To solve equation (1) for the optimal attitude, we note with Davenport [8] that the Wahba cost function, which we have seen is the same as the data-dependent part of the negative-log-likelihood function, can be rewritten as

$$L(A) = \lambda_{\max}^{(o)} - g(A), \tag{21}$$

where,

$$\lambda_{\max}^{(o)} = \sum_{k=1}^{n} \frac{1}{\sigma_k^2}, \qquad g(A) = \mathrm{tr}(AB^T), \tag{22}$$

and

$$B = \sum_{k=1}^{n} \frac{1}{\sigma_k^2} \hat{\mathbf{W}}_k \hat{\mathbf{V}}_k^T. \tag{23}$$

Here tr(·) denotes the trace of a matrix. Thus, B, the attitude profile matrix[4], contains all the information on the attitude deriving from the measurements.

The function $g(A)$ is the "gain" function, which is a maximum when $L(A)$ is a minimum. The gain function is more easily expressed in terms of the quaternion $\bar{q}$,

$$\bar{q} = [q_1, q_2, q_3, q_4]^T = \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix}, \tag{24}$$

which satisfies

$$A(\bar{q}) = (q_4^2 - \mathbf{q} \cdot \mathbf{q})I + 2\mathbf{q}\mathbf{q}^T + 2q_4[\![\mathbf{q}]\!], \tag{25}$$

where

$$[\![\mathbf{q}]\!] \equiv \begin{bmatrix} 0 & q_3 & -q_2 \\ -q_3 & 0 & q_1 \\ q_2 & -q_1 & 0 \end{bmatrix}. \tag{26}$$

The double-bracket notation, $[\![\mathbf{q}]\!]$, denotes the antisymmetric matrix of equation (26) as a function of its components. Defining the quantities

$$S = B + B^T, \qquad s = \mathrm{tr}\, B, \qquad [\![\mathbf{Z}]\!] = B - B^T, \tag{27}$$

[3]The presentation here closely follows [4] (which included a modification of Davenport's derivation) except that the weights have not been normalized to have unit sum.

[4]Note that B in equation (23) is defined as a random variable since that equation is meaningful in terms of either $\hat{\mathbf{W}}_k'$ or $\hat{\mathbf{W}}_k$. In general, when an expression is defined in terms of the (unprimed) random variables, it holds also for the realizations. When the equation is defined only for the realizations, the primes will be written explicitly.

the gain function may be rewritten in terms of the quaternion

$$g(\bar{q}) \equiv g(A(\bar{q})) = \bar{q}^T K \bar{q} \, , \tag{28}$$

where

$$K = \begin{bmatrix} S - sI & \mathbf{Z} \\ \mathbf{Z}^T & s \end{bmatrix} . \tag{29}$$

The maximization of this gain function, subject to the constraint that the quaternion have unit norm, leads to an eigenvalue equation for the optimal quaternion, which is

$$K \bar{q}^*_{ML} = \lambda_{\max} \bar{q}^*_{ML} \, , \tag{30}$$

where $\lambda_{\max}$ is the largest eigenvalue of K. In terms of the Gibbs vector ($\mathbf{Y} \equiv \mathbf{q}/q_4$) the optimal attitude may be written

$$\mathbf{Y}^* = [(\lambda_{\max} + s)I - S]^{-1}\mathbf{Z} \, . \tag{31}$$

The optimal quaternion is then reconstructed as

$$\bar{q}^* = \frac{1}{\sqrt{1 + |\mathbf{Y}^*|^2}} \begin{bmatrix} \mathbf{Y}^* \\ 1 \end{bmatrix} . \tag{32}$$

Equations (30) and (31) are Davenport's results. A very good first approximation of the optimal attitude (accurate to $O(\sigma_k^4)$) may be obtained by substituting $\lambda_{\max}^{(o)}$ for $\lambda_{\max}$ since

$$\lambda_{\max} = \lambda_{\max}^{(o)}(1 + O(\sigma_k^2)) \, . \tag{33}$$

Very efficient algorithms for computing $\lambda_{\max}$ to arbitrary accuracy are presented in [4].

The Quest Covariance From The Standpoint of Maximum Likelihood Estimation

The Fisher information matrix for the parameter vector $\mathbf{x}$ is given by

$$F_{\mathbf{xx}} = E\left\{ \frac{\partial^2}{\partial \mathbf{x} \, \partial \mathbf{x}^T} J(\mathbf{x}) \right\}_{\mathbf{x}_{\text{true}}} \, , \tag{34}$$

where the expectation is with respect to the same probability from which $J(\mathbf{x})$ was derived. We use the convention that the derivative of a scalar with respect to a column vector is a column vector. Asymptotically, i.e., as the amount of data becomes infinite, the Fisher information matrix tends to the inverse of the estimate error covariance [5, 6].

$$\lim_{n \to \infty} F_{\mathbf{xx}} = P_{\mathbf{xx}}^{-1} \tag{35}$$

If the measurements are linear in the parameter vector and Gaussian, then equation (35) is true even for finite n. The Fisher information is not well defined in terms of the quaternion since the components of the quaternion are not independent. Therefore, the Fisher information for the attitude is expressed rather in terms of incremental error angles, $\boldsymbol{\theta}$, defined according to

$$A = e^{[\boldsymbol{\theta}]} A_{true} \, , \tag{36}$$

so that θ_1, θ_2, and θ_3 are the angles of the small rotation which carries A_{true} into A. The quantity $|\theta^*|$ is small with probability very close to unity and by definition

$$\theta_{true} = \mathbf{0}\,. \tag{37}$$

Then

$$J(\theta) = \lambda_{max}^{(o)} - \mathrm{tr}(e^{[\![\theta]\!]}A_{true}B^T)\,. \tag{38}$$

Substituting this expression into equation (34), expanding the exponential function in a Taylor series, and noting that only the second-order term in $[\![\theta]\!]$ can contribute, leads to

$$F_{\theta\theta} = -\frac{\partial^2}{\partial\theta\partial\theta^T}\mathrm{tr}\left[\frac{1}{2}[\![\theta]\!]^2 A_{true}B_{true}^T\right] \tag{39}$$

$$= \mathrm{tr}(A_{true}B_{true}^T)I - A_{true}B_{true}^T\,. \tag{40}$$

Noting equation (23), the above expression may be rewritten as

$$F_{\theta\theta} = \sum_{k=1}^{n}\frac{1}{\sigma_k^2}\left(I - (\hat{\mathbf{W}}_k)_{true}(\hat{\mathbf{W}}_k^T)_{true}\right), \tag{41}$$

where

$$(\hat{\mathbf{W}}_k)_{true} \equiv A_{true}\hat{\mathbf{V}}_k\,, \tag{42}$$

which is, indeed, the inverse of the covariance used in [4]. Note that in the context of this study, as in [4], the attitude matrix is not a random matrix but a totally deterministic quantity.

The Attitude Profile Matrix as Attitude Representation

The maximum likelihood estimate of the attitude will converge to the true attitude provided that the measurements are unbiased (assumed true) and that the Fisher information matrix is positive definite (true if there are at least two nonparallel measurements). It follows that the Fisher information matrix evaluated at the maximum likelihood estimate will also tend asymptotically to the inverse of the estimate error covariance matrix. In practice, one can evaluate the Fisher information matrix only at the maximum likelihood estimate of the attitude (because in practice the true value remains forever unknown to us). Thus, to lowest nonvanishing order in σ_k, we write

$$F_{\theta\theta} \simeq \mathrm{tr}(A_{ML}^{*\prime}B'^T)I - A_{ML}^{*\prime}B'^T, \tag{43}$$

which is symmetric [2] and readily solved for B' to yield

$$B' = \left[\frac{1}{2}\mathrm{tr}(F_{\theta\theta})I - F_{\theta\theta}\right]A_{ML}^{*\prime}\,. \tag{44}$$

Thus, B' is readily calculated from $F_{\theta\theta}$ and $A_{ML}^{*\prime}$ and *vice versa*. We may also reinterpret equation (44) in terms of random B and A^*. (We would not wish to reinterpret equation (43) in terms of random variables, however, since this would mean equating a non-random quantity with an expression containing random variables.)

From the above discussion it is clear that the attitude profile matrix as given by equation (44) at best provides only an approximate representation of the spacecraft attitude covariance. However, it is accurate to lowest order in σ_k^2, which is usually a very small quantity, as observed earlier. Can it still be an exact representation of the spacecraft attitude?

To prove that the answer to this question is affirmative, we note from the form of the gain matrix as given by equation (22) that if $A_{ML}^{*\prime}$ is the maximum likelihood estimate of the attitude matrix given the attitude profile matrix B', and C_1 and C_2 are any two proper orthogonal matrices, then $C_1 A_{ML}^{*\prime} C_2^T$ is the maximum likelihood estimate of the attitude matrix given the attitude profile matrix $C_1 B' C_2^T$. Consider now an approximate attitude profile matrix, B, constructed according to equation (44) from a positive definite matrix F and a proper orthogonal matrix A. We must now show that the maximum likelihood estimate of the attitude matrix which derives from this attitude profile matrix is *exactly* A.

From the transformational properties of B it is sufficient to show that the attitude matrix constructed from $B'' = C_1 B C_2^T$ with $C_1 = I$ and $C_2 = A$ is the identity matrix. But the attitude profile matrix in this case is simply

$$B'' = \frac{1}{2}\,\mathrm{tr}(F)I - F, \tag{45}$$

which is symmetric. Hence, $\mathbf{Z}''$ computed from B'' and equation (27) must vanish. Calculating also S'' and s'' and substituting these into equation (28) leads to the gain function

$$\begin{aligned} g(\bar{q}) &= \bar{q}^T K'' \bar{q} \\ &= \frac{1}{2}\,\mathrm{tr}(F) - 2\mathbf{q}^T F \mathbf{q}, \end{aligned} \tag{46}$$

where $\mathbf{q}$ is the spatial part of the quaternion. Since F is positive definite, $g(\bar{q})$ will be a maximum when $\mathbf{q} = \mathbf{0}$, corresponding to $A = I$. QED

Thus, B as given by

$$B = \left[\frac{1}{2}\,\mathrm{tr}(F)I - F\right]A. \tag{47}$$

is an *exact* representation of the attitude and an approximate (but clearly very good) representation of the attitude error covariance. Note also that the nine elements of B constitute a minimum-dimensional representation of the spacecraft attitude (three independent parameters) and the spacecraft attitude covariance (six independent parameters).

Discussion and Conclusions

The Wahba problem has been shown to be equivalent to a maximum likelihood estimation problem for a simple but realistic probabilistic model for vector measurements. This measurement model is, in fact, the one used in earlier covariance studies of the QUEST algorithm. Thus, the Wahba problem and its solution, the QUEST algorithm, rather than being only an interesting (and practical) but arbitrarily posed optimization problem, now falls squarely within the realm of maximum likelihood estimation, so that it can be compared now with other estimators on a more fundamental level.

The restriction that the reference vectors, $\hat{\mathbf{V}}_k$, be non-random is not necessary to the discussion, but it does simplify the exposition greatly. In missions which have flown to date, such an approximation has been more than justified, since the reference vectors have been known with far greater accuracy than the observations. For some currently scheduled missions, however, where the accuracy of the attitude measurements approaches that of the reference vectors, some adjustment of the attitude estimation methodology is necessary. The capacity to account for uncertainties in the reference vectors was, in fact, already built into the QUEST algorithm [4].

If we allow $\hat{\mathbf{V}}_k$ to be distributed about some value $(\hat{\mathbf{V}}_k)_{true}$ with a Gaussian destribution given by

$$\hat{\mathbf{V}}_k = (\hat{\mathbf{V}}_k)_{true} + \Delta\hat{\mathbf{V}}_k \tag{48}$$

with

$$E\{\Delta\hat{\mathbf{V}}_k\} = \mathbf{0}, \tag{49}$$

$$E\{\Delta\hat{\mathbf{V}}_k\, \Delta\hat{\mathbf{V}}_k^T\} = \sigma_{Vk}^2[I - (\hat{\mathbf{V}}_k)_{true}(\hat{\mathbf{V}}_k)_{true}^T], \tag{50}$$

then we must also modify equations (8) through (10) to be

$$\begin{aligned}\hat{\mathbf{W}}_k &= (A\hat{\mathbf{V}}_k)_{true} + \Delta\hat{\mathbf{W}}_k, \\ &= A_{true}(\hat{\mathbf{V}}_k)_{true} + \Delta\hat{\mathbf{W}}_k\end{aligned} \tag{51}$$

with

$$E\{\Delta\hat{\mathbf{W}}_k\} = \mathbf{0}, \tag{52}$$

$$E\{\Delta\hat{\mathbf{W}}_k\, \Delta\hat{\mathbf{W}}_k^T\} = \sigma_{Wk}^2[I - (A\hat{\mathbf{V}}_k)_{true}(A\hat{\mathbf{V}}_k)_{true}^T]. \tag{53}$$

With this somewhat more detailed model it turns out that the negative-log-likelihood function remains unchanged except that we must now make the identification

$$\sigma_k^2 = \sigma_{Wk}^2 + \sigma_{Vk}^2. \tag{54}$$

This result, stated in different terms, is essentially at the bottom of the work presented in [4]. Likewise, we can posit a probability density function of the form of equation (4) with σ_k^2 so defined. Thus, all the results of this work, from equation (15) until the end, remain unchanged even when $\hat{\mathbf{V}}_k$ is allowed to be random.

The demonstration that the attitude profile matrix is an exact representation of the attitude as well as an approximate (but very good) representation of the attitude covariance will have important consequences. The Wahba problem is static, i.e., the attitude has a single value for all the measurements and is, therefore, independent of time. However, the extension to dynamic but deterministic dynamical systems is straightforward. In this context the QUEST algorithm becomes a filter, which can be shown, in fact, to be a new mechanization of the Kalman filter for attitude. The development of a QUEST filter which treats uncertain dynamic systems is also possible but extremely cumbersome unless certain simplifying approximations are made. The relationship between a sequentialized dynamic QUEST algorithm and the Kalman filter will be explored in a future work [9].

Without entering further into the discussion of a filter implementation of the QUEST algorithm here, we note that equation (44) provides the means for including an *a priori* estimate of the attitude in the QUEST algorithm. Namely, if $A^{*\prime}(-)$ is the *a priori* maximum likelihood estimate of the attitude and $P_{\theta\theta}(-)$ the corresponding *a priori* attitude-estimate-error covariance, then the appropriate attitude profile matrix taking both the *a priori* estimate and the current data into account is simply

$$B' = \left[\frac{1}{2}\,\mathrm{tr}(P_{\theta\theta}^{-1}(-))I - P_{\theta\theta}^{-1}(-)\right]A^{*\prime}(-) + \sum_{k=1}^{n}\frac{1}{\sigma_k^2}\hat{\mathbf{W}}_k\hat{\mathbf{V}}_k^T . \qquad (55)$$

This result is of prime importance in the development of a QUEST filter [9].

Acknowledgment

The author is grateful to F. Landis Markley and J. Courtney Ray for interesting discussions and helpful criticisms and to Keith L. Musser for a critical reading of the manuscript.

References

[1] WAHBA, G. "A Least-Squares Estimate of Satellite Attitude," Problem 65–1, *SIAM Review*, Vol. 7, No. 3, July 1965, p. 409.

[2] SHUSTER, M. D. "A Comment on Fast Three-Axis Attitude Determination Using Vector Observations and Inverse Iteration," *Journal of the Astronautical Sciences*, Vol. 31, No. 4, October-December 1983, pp. 579–584.

[3] MARKLEY, F. L. "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.

[4] SHUSTER, M. D. and OH, S. D. "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance, Control and Dynamics*, Vol. 4, No. 1, January-February 1981, pp. 70–77.

[5] RAO, C. R. *Linear Spectral Inference and its Applications*. John Wiley and Sons, New York, 1973.

[6] SORENSON, H. W. *Parameter Estimation*, Marcel Dekker, New York, 1980.

[7] SHUSTER, M. D. "Lectures on Spacecraft Attitude Estimation," 1983 et seq., (unpublished).

[8] DAVENPORT, P. B., unpublished.

[9] SHUSTER, M. D. "A Simple Kalman Filter and Smoother for Spacecraft Attitude" *Journal of the Astronautical Sciences*, Vol. 37, No. 1, January-March 1989, pp. 89–106.