Proceedings of the 27th Conference
on Decision and Control
Austin, Texas • December 1988

SPACECRAFT ALIGNMENT ESTIMATION

Gerald J. Bierman
late of Bierman & Associates
Studio City, California 91604

and

Malcolm D. Shuster
The Johns Hopkins University, Applied Physics Laboratory
Laurel, Maryland 20707

Abstract

A numerically well-behaved factorized methodology is developed for estimating spacecraft sensor alignments from prelaunch and inflight data without the need to compute the spacecraft attitude or angular velocity. Such a methodology permits the estimation of sensor alignments (or other biases) in a framework free of unknown dynamical variables. In actual mission implementation such an algorithm is usually better behaved than one which must compute sensor alignments simultaneously with the spacecraft attitude, say, by means of a Kalman filter. In particular, such a methodology is less sensitive to data dropouts of long duration and the derived measurement used in the attitude-independent algorithm usually makes data checking and editing of outliers much simpler than would be the case in the filter.

Introduction

An earlier paper [1] presented a simple but approximate algorithm for computing spacecraft misalignments. That work neglected problems of redundancy among the derived measurements as well as correlations. The present work removes these limitations and presents an exact formalism within the framework of maximum likelihood estimation.

As in the previous paper, the specific algorithm developed is for vector sensors, since these are overwhelmingly the sensors carried on a spacecraft for which inflight alignment calibration is important. The generalization of the algorithm to other sensors is straightforward, if a bit tedious. The previous algorithm, which shares many features with the present offering, has performed well in support of several NASA missions.

The Measurement Model and Definitions

A vector sensor measures the direction of some vector, for example, the position of the Sun or some distant star, the nadir, or the geomagnetic field. The measurement may be modeled as

$$\hat{\mathbf{W}}_{i,k} = A_k \, \hat{\mathbf{V}}_{i,k} + \Delta\hat{\mathbf{W}}_{i,k} \quad , \tag{1}$$

where $\hat{\mathbf{W}}_{i,k}$ is the observed vector in the spacecraft body frame and $\hat{\mathbf{V}}_{i,k}$ is the known reference vector in the frame with respect to which the attitude is to be determined. A_k is the spacecraft attitude matrix and $\Delta\hat{\mathbf{W}}_{i,k}$ is the measurement noise expressed in body coordinates. Here i is the sensor index, $i = 1, \ldots, n$, and k is the time index, $k = 1, \ldots, N$.

Experience has shown [2,3] that the measurement noise, $\Delta\hat{\mathbf{W}}_{i,k}$, is well described by a discrete white Gaussian process satisfying

$$E\{\Delta\hat{\mathbf{W}}_{i,k}\} = \mathbf{0} \quad , \tag{2}$$

$$E\{\Delta\hat{\mathbf{W}}_{i,k} \, \Delta\hat{\mathbf{W}}_{i',k'}^T\} = \delta_{ii'} \, \delta_{kk'} \, \sigma_{i,k}^2 \left(I - \left(A_k \, \hat{\mathbf{V}}_{i,k}\right) \left(A_k \, \hat{\mathbf{V}}_{i,k}\right)^T \right) \quad . \tag{3}$$

The actual observations take place not in the spacecraft body frame but in a frame defined for each sensor. Thus, the actual observation, $\hat{\mathbf{U}}_{i,k}$, is related to the body-referenced measurement according to

$$\hat{\mathbf{W}}_{i,k} = S_i \, \hat{\mathbf{U}}_{i,k} \quad , \tag{4}$$

where S_i is the alignment matrix, which is proper orthogonal. Thus, we may write equivalently in sensor coordinates

$$\Delta\mathbf{U}_{i,k} = S_i^T \, A_k \, \hat{\mathbf{V}}_{i,k} + \Delta\hat{\mathbf{U}}_{i,k} \quad , \tag{5}$$

with

$$E\{\Delta\hat{\mathbf{U}}_{i,k}\} = \mathbf{0} \quad , \tag{6}$$

$$E\{\Delta\hat{\mathbf{U}}_{i,k} \, \Delta\hat{\mathbf{U}}_{i',k'}^T\} = \delta_{ii'} \, \delta_{kk'} \, \sigma_{i,k}^2 \left(I - \left(S_i^T A_k \, \hat{\mathbf{V}}_{i,k}\right) \left(S_i^T A_k \, \hat{\mathbf{V}}_{i,k}\right)^T \right) \quad . \tag{7}$$

Considerable processing of the raw spacecraft data is necessary, of course, to compute $\hat{\mathbf{U}}_{i,k}$.

In general, a prelaunch estimate of S_i is available based on ground measurements. Thus, we write

$$S_i = M_i \, S_i^o \quad , \tag{8}$$

where S_i^o is the prelaunch value of the alignment matrix and M_i, a proper orthogonal rotation for a small rotation, is the misalignment matrix. Thus,

$$M_i = I + [\boldsymbol{\theta}_i] + O(|\boldsymbol{\theta}_i|^2) \quad , \tag{9}$$

with

$$[\boldsymbol{\theta}] \equiv \begin{bmatrix} 0 & \theta_3 & -\theta_2 \\ -\theta_3 & 0 & \theta_1 \\ \theta_2 & -\theta_1 & 0 \end{bmatrix} \quad . \tag{10}$$

The three angles, θ_1, θ_2, θ_3, we term the misalignment angles and $\boldsymbol{\theta}$, the misalignment vector. Note that in constructing S_i from S_i^o and $\boldsymbol{\theta}_i$, Eq. (9) must be replaced by a formula which is exactly orthogonal, say by interpreting $\boldsymbol{\theta}_i$ as twice the Gibbs vector characterizing the misalignment [1]. By definition, the prelaunch estimate of $\boldsymbol{\theta}_i$ is $\mathbf{0}$ and the result of the prelaunch alignment may be rewritten

$$\boldsymbol{\theta}_i^*(-) = \mathbf{0} + \Delta\boldsymbol{\theta}_i(-) \quad , \tag{11}$$

where $\Delta\boldsymbol{\theta}_i(-)$, the prelaunch alignment error, is zero mean and

$$E\{\Delta\boldsymbol{\theta}_i(-)\,\Delta\boldsymbol{\theta}_j^T(-)\} = \delta_{ij}\,P_{ii}(-) \quad . \tag{12}$$

We assume, naturally, that $\Delta\boldsymbol{\theta}_i(-)$ and $\Delta\mathbf{W}_{j,k}$ are statistically independent for all values of the indices. Note that we have used an asterisk (∗) to denote estimates in order to avoid confusion with the caret that denotes unit vectors.

The inflight sensor measurements alone are not sufficient to determine both the alignments and the spacecraft attitude. This follows from the fact that Eq. (5) is invariant under the simultaneous transformations

$$S_i \rightarrow T\,S_i \quad , \tag{13a}$$

$$A_k \rightarrow T\,A_k \quad , \tag{13b}$$

where T is proper orthogonal. Thus, using inflight data alone a common (body-referenced) misalignment is indistinguishable from an attitude error. The ambiguity is removed by the prelaunch estimate of the misalignments, which provide the necessary *a priori* condition.

The estimation of A_k , $k = 1, \ldots, N$, and $\boldsymbol{\theta}_i$, $i = 1, \ldots, n$, in a Kalman filter is straightforward. We wish to avoid that approach, however, for reasons stated in the abstract. We accomplish this by defining attitude-independent measurements which are sensitive only to the misalignments.

Attitude-Independent Estimation of Alignments

Let

$$\hat{\mathbf{W}}_{i,k}^o \equiv S_i^o\,\hat{\mathbf{U}}_{i,k} \tag{14}$$

denote the uncalibrated body-referenced measurement (based solely on the prelaunch calibration). Then we define for $i \neq j$

$$z_{ij,k} \equiv \hat{\mathbf{W}}_{i,k}^o \cdot \hat{\mathbf{W}}_{j,k}^o - \hat{\mathbf{V}}_{i,k} \cdot \hat{\mathbf{V}}_{j,k} \quad . \tag{15}$$

It is simple to show to first order that

$$z_{ij,k} = (\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o) \cdot (\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) + \Delta z_{ij,k} \quad , \tag{16}$$

where

$$\Delta z_{ij,k} = \hat{\mathbf{W}}_{i,k}^o \cdot \Delta\hat{\mathbf{W}}_{j,k} + \hat{\mathbf{W}}_{j,k}^o \cdot \Delta\hat{\mathbf{W}}_{i,k} \quad . \tag{17}$$

Thus, to first order

$$E\{\Delta z_{ij,k}\} = \mathbf{0} \quad , \tag{18}$$

$$E\{\Delta z_{ij,k}^2\} = (\sigma_{i,k}^2 + \sigma_{j,k}^2)\,|\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o|^2 \quad , \tag{19a}$$

$$E\{\Delta z_{ij,k}\,\Delta z_{i\ell,k}\} = \sigma_{i,k}^2\,(\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o) \cdot (\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{\ell,k}^o) \quad , \tag{19b}$$

$$E\{\Delta z_{ij,k}\Delta z_{\ell m,k}\} = \mathbf{0} \quad , \tag{19c}$$

where i, j, ℓ, and m above denote distinct indices.

The measurements $z_{ij,k}$, $i < j$, cannot all be distinct. If there are n sensors, each measuring a unit vector, then there are only $2n$ equivalent independent scalar measurements, while there are $n(n-1)/2$ possible $z_{ij,k}$ with $i < j$. Since three combinations of the $\hat{\mathbf{W}}_{i,k}$ determine the attitude and the $z_{ij,k}$ are by explicit construction attitude-independent, there can only be $2n-3$ statistically independent $z_{ij,k}$. Table 1 displays these numbers.

n	$2n-3$	$n(n-1)/2$
2	1	1
3	3	3
4	5	6
5	7	10
6	9	15

Table 1
Number of Independent Attitude-Independent Measurements Compared with Number of Derived Measurements

Thus, for more than three sensors the derived measurements $z_{ij,k}$ become redundant and this redundancy grows disproportionately with the number of sensors.

To determine $2n-3$ independent measurements from among the $n(n-1)/2$ possible $z_{ij,k}$ we remark that

$$2n - 3 = (n-1) + (n-2) \tag{20}$$

Thus, it is tempting to suggest that the set of measurements

$$\{z_{1j,k}\,, j = 2, \ldots, n\,;\; z_{2j,k}\,, j = 3, \ldots, n\}$$

is the desired set provided that $\hat{\mathbf{W}}_{1,k}^o$ and $\hat{\mathbf{W}}_{2,k}^o$ are not collinear. That these $2n-3$ measurements are indeed statistically independent is easily seen by arranging the noise terms as

$$\Delta z_{12,k} = \hat{\mathbf{W}}_{1,k}^o \cdot \Delta\hat{\mathbf{W}}_{2,k} + \hat{\mathbf{W}}_{2,k}^o \cdot \Delta\hat{\mathbf{W}}_{1,k} \quad ,$$

$$\Delta z_{1j,k} = \hat{\mathbf{W}}_{1,k}^o \cdot \Delta\hat{\mathbf{W}}_{j,k} + \hat{\mathbf{W}}_{j,k}^o \cdot \Delta\hat{\mathbf{W}}_{1,k} \quad , \quad j = 3, \ldots, n \quad ,$$

$$\Delta z_{2j,k} = \hat{\mathbf{W}}_{2,k}^o \cdot \Delta\hat{\mathbf{W}}_{j,k} + \hat{\mathbf{W}}_{j,k}^o \cdot \Delta\hat{\mathbf{W}}_{2,k} \quad , \quad j = 3, \ldots, n \quad .$$

The expressions in the second and third lines are clearly independent since they contain distinct components of $\Delta\hat{\mathbf{W}}_{j,k}$, $j = 3, \ldots, n$, and these are independent of the expression in the first line since it is independent of $\Delta\hat{\mathbf{W}}_{j,k}$, $j = 3, \ldots, n$.

(Note that for vector magnetometers, the sensor supplies three equivalent scalar measurements not two. The additional attitude independent measurement may be taken to be $|\mathbf{B}_k|$, the magnitude of the measured field, which is independent of the alignments but not of additive magnetometer biases, which are often significant.)

Thus, if we define

$$\mathbf{Z}_k \equiv [z_{12,k}, \ldots, z_{1n,k}, z_{23,k}, \ldots, z_{2n,k}]^T \quad , \tag{21}$$

then we may write

$$\mathbf{Z}_k = H_k \mathbf{\Theta} + \Delta\mathbf{Z}_k \quad , \tag{22}$$

where

$$\mathbf{\Theta} \equiv [\boldsymbol{\theta}_1{}^T \ldots, \boldsymbol{\theta}_n^T]^T \quad , \tag{23}$$

and $\Delta\mathbf{Z}_k$ is a white Gaussian sequence with covariance matrix $P_{\mathbf{Z},k}$ (which will be discussed in greater detail in the next section). H_k and $P_{\mathbf{Z},k}$ are obtained directly from Eqs. (16) through (19). The inflight estimate $\mathbf{\Theta}^*(+)$ together with $P_{\mathbf{\Theta\Theta}}(+)$ may be obtained by solving the normal equations:

$$\left[P_{\mathbf{\Theta\Theta}}^{-1}(-) + \sum_{k=1}^{N} H_k^T P_{\mathbf{Z},k}^{-1} H_k\right] \mathbf{\Theta}^*(+) = \sum_{k=1}^{N} H_k^T P_{\mathbf{Z},k}^{-1} \mathbf{Z}_k \quad , \tag{24}$$

$$P_{\mathbf{\Theta\Theta}}^{-1}(+) = P_{\mathbf{\Theta\Theta}}^{-1}(-) + \sum_{k=1}^{N} H_k^T P_{\mathbf{Z},k}^{-1} H_k \quad . \tag{25}$$

Factorized Methodology

The above methodology suffers from two important complications. First, the set of active sensors may be different in every frame (labeled by k) and a complicated logic may be required to determine $\hat{\mathbf{W}}_{1,k}$ and $\hat{\mathbf{W}}_{2,k}$ in each frame. Also, if, perchance, one of these two vectors is nearly collinear with one of the remaining vectors, the measurement model may suffer unduly from numerical errors. The present section presents a mathematically equivalent but numerically superior algorithm to the one just derived.

Note that

$$\begin{aligned} E\{\Delta\hat{\mathbf{W}}_{i,k}\, \Delta\hat{\mathbf{W}}_{i,k}^T\} &= \sigma_{i,k}^2 \left(I - \left(A_k \hat{\mathbf{V}}_{i,k}\right)\left(A_k \hat{\mathbf{V}}_{i,k}\right)^T\right) \\ &= \left(\sigma_{i,k}\,[A_k \hat{\mathbf{V}}_{i,k}]\right)\left(\sigma_{i,k}\,[A_k \hat{\mathbf{V}}_{i,k}]\right)^T \quad . \end{aligned} \tag{26}$$

Thus, we may write

$$\Delta\hat{\mathbf{W}}_{i,k} = \sigma_{i,k}\,[A_k \hat{\mathbf{V}}_{i,k}]\,\boldsymbol{\epsilon}_{i,k} \quad , \tag{27}$$

where

$$E\{\boldsymbol{\epsilon}_{i,k}\} = \mathbf{0} \quad , \tag{28}$$

$$E\{\boldsymbol{\epsilon}_{i,k}\, \boldsymbol{\epsilon}_{j,k}^T\} = \delta_{ij}\, I_{3\times 3} \quad . \tag{29}$$

Therefore, we may write to lowest order

$$z_{ij,k} = (\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o)\cdot(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) + B_{ij,k}^i \boldsymbol{\epsilon}_{i,k} + B_{ij,k}^j \boldsymbol{\epsilon}_{j,k} \quad , \tag{30}$$

with

$$\begin{aligned} B_{ij,k}^i &= (\hat{\mathbf{W}}_{j,k}^o)^T (\sigma_{i,k}\,[\hat{\mathbf{W}}_{i,k}^o]) \\ &= \sigma_{i,k}\,(\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o)^T \quad , \end{aligned} \tag{31a}$$

$$B_{ij,k}^j = \sigma_{j,k}\,(\hat{\mathbf{W}}_{j,k}^o \times \hat{\mathbf{W}}_{i,k}^o)^T \quad . \tag{31b}$$

Thus, we may write

$$\mathbf{Z}_k = H_k \mathbf{\Theta} + B_k \boldsymbol{\epsilon}_k \quad , \tag{32}$$

where $\boldsymbol{\epsilon}_k$ is a discrete white Gaussian process with covariance matrix equal to $I_{3n\times 3n}$ and, therefore,

$$P_{\mathbf{Z},k} = B_k B_k^T \quad . \tag{33}$$

Note also from Eqs. (16) and (31) that

$$z_{ij,k} = (\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o)\cdot(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j + \sigma_{i,k}\boldsymbol{\epsilon}_{i,k} - \sigma_{j,k}\boldsymbol{\epsilon}_{j,k}) \quad , \tag{34}$$

so that

$$B_k = H_k \begin{bmatrix} \sigma_{1,k} I_{3\times 3} & O_{3\times 3} & \cdots & O_{3\times 3} \\ O_{3\times 3} & \sigma_{2,k} I_{3\times 3} & \cdots & O_{3\times 3} \\ \vdots & \vdots & \ddots & \vdots \\ O_{3\times 3} & O_{3\times 3} & \cdots & \sigma_{n,k} I_{3\times 3} \end{bmatrix} \quad . \tag{35}$$

The above formulation holds true whether $\mathbf{Z}_k$ is chosen to be a vector of length $2n-3$ or $n(n-1)/2$. In the general (and common) case where n is a function of k, the latter case is easier to treat in practice. Since n is most often 3 and seldom greater than 6, the computational burden of choosing the longer $\mathbf{Z}_k$ is not prohibitive.

By the singular-value-decomposition (SVD) theorem [4], B_k may be factored as

$$B_k = U_k \mathcal{S}_k V_k^T \quad , \tag{36}$$

where U_k and V_k are orthogonal and $\mathcal{S}_k$ is a diagonal $m_k \times 3n$ matrix, where m_k is the dimension of $\mathbf{Z}_k$, and

$$(\mathcal{S}_k)_{11} \geq (\mathcal{S}_k)_{22} \geq \ldots \geq 0 \quad , \tag{37}$$

so that

$$P_{\mathbf{Z},k} = U_k D_k U_k^T \quad , \tag{38}$$

with

$$D_k = \mathcal{S}_k \mathcal{S}_k^T \quad , \tag{39}$$

which is a diagonal and positive semi-definite $m_k \times m_k$ matrix.

If we now define

$$\boldsymbol{\zeta}_k \equiv U_k^T \mathbf{Z}_k \quad , \tag{40}$$

$$C_k \equiv U_k^T H_k \quad , \tag{41}$$

then

$$\boldsymbol{\zeta}_k = C_k \mathbf{\Theta} + \mathcal{S}_k \boldsymbol{\epsilon}_k' \quad , \tag{42}$$

where $\boldsymbol{\epsilon}_k' \sim \mathcal{N}(\mathbf{0}_{3n}, I_{3n\times 3n})$. If $\ell_{\max,k}$ is the largest index for which $(\mathcal{S}_k)_{\ell\ell} > 0$, then the first $\ell_{\max,k}$ components of $\boldsymbol{\zeta}_k$ contribute $\ell_{\max,k}$ independent measurements of $\mathbf{\Theta}$. (The remaining components are simply the constraints on $\mathbf{Z}_k$.) This minimal set of measurements may now be combined in a batch algorithm to yield the normal equations

$$\left[P_{\mathbf{\Theta\Theta}}^{-1}(-) + \sum_{k=1}^{N} C_k'^T D_k'^{-1} C_k'\right] \mathbf{\Theta}^*(+) = \sum_{k=1}^{N} C_k'^T D_k'^{-1} \boldsymbol{\zeta}_k' \quad , \tag{43}$$

$$P_{\mathbf{\Theta\Theta}}^{-1}(+) = P_{\mathbf{\Theta\Theta}}^{-1}(-) + \sum_{k=1}^{N} C_k'^T D_k'^{-1} C_k' \quad , \tag{44}$$

where the prime denotes the appropriate truncation. Likewise, if we wish to process the measurements sequentially, the structure of the derived measurement noise is now in a form which is convenient for square-root information filters [5].

Discussion and Conclusions

The above methodology considerably enhances the algorithm of Reference 1 in that it now treats the statistical properties of the measurements properly within the framework of maximum likelihood estimation.

The estimate is not optimal over all $2n$ sensor measurements since effectively three of the measurements have been removed in order to achieve attitude independence. Thus, the results of this algorithm will differ somewhat from the results that would obtain from a complete Kalman filter treating both the attitude and the alignments. In particular, since the alignments are computed from derived measurements to which the attitude is not sensitive, it follows that the alignment estimates as computed by the present algorithm will always be statistically independent of the *uncalibrated* attitude estimate, which will not be the case of the truly optimal estimate over all $2n$ measurements. The difference in the estimates is not expected to be significant, however. Those alignments which are sensitive to the set of derived measurements will clearly not be very sensitive to the attitude errors once sufficient data has been processed. Likewise, those alignments which are not very sensitive to the derived measurements will be determined largely by the prelaunch calibration anyway. Thus, we expect little accuracy to be lost in our approximation.

Since the $z_{ij,k}$ are zero-mean and have easily calculable variances, the detection of outliers among these derived measurements is very direct. In general, these outliers will be due to an improper measurement of a vector $\hat{\mathbf{W}}_{i,k}$ or $\hat{\mathbf{W}}_{j,k}$ (due, for example, to the misidentification of a star). In this case, it is expected that most of the $z_{ij,k}$ for the given i and k or j and k will be outside the bounds. This makes the identification of outliers simple.

Very often, spacecraft carry sensors of widely different accuracies, some supplying attitude information with an accuracy of 10 arc sec while others are accurate to only 0.5 deg. If there are sufficient highly accurate sensors available most of the time, there is little to be gained by estimating the alignment of the coarse sensors simultaneously with the fine. In this case it is preferable to estimate the alignments of the fine sensors first using the attitude independent algorithm presented here and then use simple regression techniques to estimate the attitude of the coarser sensors using the computed spacecraft attitude from the fine sensors to remove the ambiguity.

Acknowledgements

The present work was supported in part under NASA contract NAS5–29268. The encouragement and support of Sanford Hinkel of NASA Goddard Space Flight Center is gratefully acknowledged.

References

[1] Shuster, M. D., D. M. Chitre, and D. P. Niebur, "In-flight Estimation of Spacecraft Attitude Sensor Accuracies and Alignments," *Journal of Guidance, Control, and Dynamics*, Vol. **5**, No. 4, July-August 1982, pp. 339-343.

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

[3] Shuster, M. D., "Maximum Likelihood Estimation of Spacecraft Attitude," to appear in *Journal of the Astronautical Sciences.*

[4] Golub, Gene H., and Charles F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1983.

[5] Bierman, Gerald J., *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York, 1977.