How to Estimate Attitude from Vector Observations

AAS Paper 99-427, presented at the
AAS/AIAA Astrodynamics Specialist Conference,
Girdwood, Alaska
August 16–19, 1999

F. Landis Markley
Guidance, Navigation, & Control Center
NASA Goddard Space Flight Center
Greenbelt, MD

Daniele Mortari
Aerospace Engineering School
University of Rome "*La Sapienza*"
Rome, Italy

Wahba's Problem - *SIAM Review*, July 1965

Find the orthogonal matrix A with determinant +1 that minimizes

$$L(A) \equiv \tfrac{1}{2}\sum_i a_i |\mathbf{b}_i - A\mathbf{r}_i|^2 .$$

where $\{\mathbf{b}_i\}$ are unit vectors in a body frame, $\{\mathbf{r}_i\}$ are unit vectors in a reference frame, and $\{a_i\}$ are non-negative weights. Writing

$$L(A) = \lambda_0 - \mathrm{tr}(AB^{\mathrm{T}}), \text{ with } \lambda_0 \equiv \sum_i a_i \text{ and } B \equiv \sum_i a_i \mathbf{b}_i \mathbf{r}_i^{\mathrm{T}},$$

it is clear that we can minimize $L(A)$ by maximizing $\mathrm{tr}(AB^{\mathrm{T}})$.

This is equivalent to the orthogonal Procrustes problem, which is to find the orthogonal matrix A that is closest to B in the sense of the Frobenius norm, $\|M\|_{\mathrm{F}}^2 \equiv \sum_{i,j} M_{ij}^2 = \mathrm{tr}(MM^{\mathrm{T}})$, since

$$\|A - B\|_{\mathrm{F}}^2 = \|A\|_{\mathrm{F}}^2 + \|B\|_{\mathrm{F}}^2 - 2\mathrm{tr}(AB^{\mathrm{T}}) = 3 + \|B\|_{\mathrm{F}}^2 - 2\mathrm{tr}(AB^{\mathrm{T}})$$

First Solutions - *SIAM Review*, July 1966

1: J. L. Farrell and J. C. Stuelpnagel:

B has the polar decomposition $B = WH$ where W is orthogonal and H is symmetric and positive semidefinite.

If $\det W$ is positive, then $A_{opt} = W$.

Else, diagonalize H by $H = VDV^T$, where V is orthogonal and D is diagonal with elements arranged in decreasing order.

Then $A_{opt} = WV \, \mathrm{diag}[1 \;\; 1 \;\; \det W]V^T$.

2: R. H. Wessner:

$A_{opt} = (B^T)^{-1}(B^T B)^{1/2}$, which is equivalent to $A_{opt} = B(B^T B)^{-1/2}$.

This requires 3 vectors ($\det B \neq 0$); only 2 are really needed.

3–6: J. R. Velman, J. E. Brock, R. Desjardins, and Wahba.

Singular Value Decomposition (SVD) Method - 1987

B has the Singular Value Decomposition:

$$B = U \Sigma V^{T} = U \, \mathrm{diag}[\Sigma_{11} \;\; \Sigma_{22} \;\; \Sigma_{33}] \, V^{T},$$

where U and V are orthogonal and $\Sigma_{11} \geq \Sigma_{22} \geq \Sigma_{33} \geq 0$.

The optimal attitude matrix is $A_{opt} = U \, \mathrm{diag}[1 \;\; 1 \;\; (\det U)(\det V)] \, V^{T}$.

The SVD method is completely equivalent to the Farrell and Stuelpnagel solution with $U = WV$. The difference is that SVD algorithms exist now and are among the most robust numerical algorithms. MATLAB computes A_{opt} from B in two lines of code.

QUaternion ESTimator (QUEST) - 1978

The first three rows of $(\lambda_{\max} I - K)q_{\text{opt}} = 0$ give

$$q_{\text{opt}} = \frac{1}{\sqrt{\gamma^2 + |\mathbf{x}|^2}} \begin{bmatrix} \mathbf{x} \\ \gamma \end{bmatrix}, \quad \begin{array}{ll} \text{where} & \mathbf{x} \equiv \{\text{adj}[(\lambda_{\max} + \text{tr}B)I - B - B^{\text{T}}]\}\mathbf{z}, \\ \text{and} & \gamma \equiv \det[(\lambda_{\max} + \text{tr}B)I - B - B^{\text{T}}] \end{array}$$

Find $\lambda_{\max}$ by Newton-Raphson iteration of the characteristic eqn.,

$$\det(\lambda_{\max} I - K) = (\lambda_{\max} - \text{tr}\, B)\gamma - \mathbf{z}^{\text{T}}\mathbf{x} = 0.$$

Iterate from λ_0, since $\lambda_{\max}$ is very close to λ_0 if $L(A_{\text{opt}}) = \lambda_0 - \lambda_{\max}$ is small. The analytic solution is slower and no more accurate.

QUEST would fail for 180° rotations, but the method of sequential rotations (effectively permuting q components) handles this case.

EStimator of the Optimal Quaternion (ESOQ) - 1996

Solve the characteristic equation for $\lambda_{\max}$ as in QUEST.

The adjoint of $\lambda_{\max} I - K$ can be shown analytically to obey

$$\mathrm{adj}(\lambda_{\max} I - K) = (\lambda_{\max} - \lambda_2)(\lambda_{\max} - \lambda_3)(\lambda_{\max} - \lambda_4) q_{\mathrm{opt}} q_{\mathrm{opt}}^{\mathrm{T}},$$

where λ_2, λ_3, and λ_4 are the other three eigenvalues of K.

Thus q_{opt} can be computed by normalizing any non-zero column of $\mathrm{adj}(\lambda_{\max} I - K)$. This is the "4-dimensional cross-product" of the other three columns of $\lambda_{\max} I - K$.

ESOQ2 - 1997

Substituting $q_{\text{opt}} = \begin{bmatrix} \mathbf{e}\sin(\phi/2) \\ \cos(\phi/2) \end{bmatrix}$ into $(\lambda_{\max} I - K)q_{\text{opt}} = 0$ gives

$(\lambda_{\max} - \text{tr}B)\cos(\phi/2) = \mathbf{e}^{\text{T}}\mathbf{z}\sin(\phi/2)$ and

$[(\lambda_{\max} + \text{tr}B)I - B - B^{\text{T}}]\mathbf{e}\sin(\phi/2) = \mathbf{z}\cos(\phi/2)$.

Eliminating the rotation angle ϕ gives $M\mathbf{e} = \mathbf{0}$, where

$$M \equiv (\lambda_{\max} - \text{tr}B)[(\lambda_{\max} + \text{tr}B)I - B - B^{\text{T}}] - \mathbf{z}\mathbf{z}^{T} \equiv [\mathbf{m}_1 \ \ \mathbf{m}_2 \ \ \mathbf{m}_3].$$

The rotation axis is $\mathbf{e} = \mathbf{y}/|\mathbf{y}|$, where $\mathbf{y}$ is any $\mathbf{m}_i \times \mathbf{m}_j$. Then

$$q_{\text{opt}} = \frac{1}{\sqrt{|(\lambda_{\max} - \text{tr}B)\mathbf{y}|^2 + (\mathbf{z}\cdot\mathbf{y})^2}} \begin{bmatrix} (\lambda_{\max} - \text{tr}B)\mathbf{y} \\ \mathbf{z}\cdot\mathbf{y} \end{bmatrix}.$$

Fast Optimal Attitude Matrix (FOAM) - 1993

Find λ_{max} by solving the characteristic equation

$$0 = (\lambda^2 - \|B\|_F^2)^2 - 8\lambda \det B - 4\|\text{adj}B\|_F^2.$$

This becomes an easily solved quadratic in λ^2 if $\det B = 0$, as in the case of two observations. The attitude matrix is given by

$$A_{opt} = (\kappa\lambda_{max} - \det B)^{-1}[(\kappa + \|B\|_F^2)B + \lambda_{max}\text{adj}B^T - BB^T B],$$

where $\kappa \equiv \frac{1}{2}(\lambda_{max}^2 - \|B\|_F^2)$.

For the analysis in this paper, the quaternion representation of the optimal attitude is computed.

ESOQ, ESOQ2, and FOAM avoid QUEST's sequential rotations.

Two-Observation Case

In this case $\det B = 0$, the odd terms in λ in the characteristic equation vanish, and

$$\lambda_{\max} = \sqrt{a_1^2 + a_2^2 + 2a_1 a_2[(\mathbf{b}_1 \cdot \mathbf{b}_2)(\mathbf{r}_1 \cdot \mathbf{r}_2) + |\mathbf{b}_1 \times \mathbf{b}_2||\mathbf{r}_1 \times \mathbf{r}_2|]}\ .$$

This simplifies both QUEST and FOAM; FOAM gives

$$A_{\text{opt}} = \mathbf{b}_3\mathbf{r}_3^{\text{T}} + (a_1/\lambda_{\max})[\mathbf{b}_1\mathbf{r}_1^{\text{T}} + (\mathbf{b}_1 \times \mathbf{b}_3)(\mathbf{r}_1 \times \mathbf{r}_3)^{\text{T}}]$$
$$+ (a_2/\lambda_{\max})[\mathbf{b}_2\mathbf{r}_2^{\text{T}} + (\mathbf{b}_2 \times \mathbf{b}_3)(\mathbf{r}_2 \times \mathbf{r}_3)^{\text{T}}],$$

where $\mathbf{b}_3 \equiv (\mathbf{b}_1 \times \mathbf{b}_2)/|\mathbf{b}_1 \times \mathbf{b}_2|$ and $\mathbf{r}_3 \equiv (\mathbf{r}_1 \times \mathbf{r}_2)/|\mathbf{r}_1 \times \mathbf{r}_2|$.

This goes over to the TRIAD solution for $a_1 = 0$, $a_2 = 0$, or $a_1 = a_2$.

Sequential Methods

The basic idea is to propagate B or K to time t and then update.

Filter QUEST -1989

$$B(\text{new}) = \mu \Phi_{3\times 3} B(\text{old}) + \sum_{i=k+1}^{k+n_t} a_i \mathbf{b}_i \mathbf{r}_i^{\mathrm{T}}, \text{ sum over } n_t \text{ observations at } t.$$

Recursive QUEST (REQUEST) -1996

$$K(\text{new}) = \mu \Phi_{4\times 4} K(\text{old}) \Phi_{4\times 4}^{\mathrm{T}} + \sum_{i=k+1}^{k+n_t} a_i \tilde{K}_i, \text{ where}$$

$$\tilde{K}_i \equiv \begin{bmatrix} \mathbf{b}_i \mathbf{r}_i^{\mathrm{T}} + \mathbf{r}_i \mathbf{b}_i^{\mathrm{T}} - (\mathbf{b}_i \cdot \mathbf{r}_i) I & (\mathbf{b}_i \times \mathbf{r}_i) \\ (\mathbf{b}_i \times \mathbf{r}_i)^{\mathrm{T}} & \mathbf{b}_i \cdot \mathbf{r}_i \end{bmatrix}.$$

These are mathematically equivalent. Filter QUEST requires fewer computations, but neither has been successful in practice.

Reynolds's Sequential Algorithm - 1997

There are two orthogonal quaternions that map a vector $\mathbf{r}_i$ into $\mathbf{b}_i$:

$$q_1 \equiv \frac{1}{\sqrt{2(1+\mathbf{b}_i \cdot \mathbf{r}_i)}} \begin{bmatrix} \mathbf{b}_i \times \mathbf{r}_i \\ 1+\mathbf{b}_i \cdot \mathbf{r}_i \end{bmatrix} \quad \text{and} \quad q_2 \equiv \frac{1}{\sqrt{2(1+\mathbf{b}_i \cdot \mathbf{r}_i)}} \begin{bmatrix} \mathbf{b}_i + \mathbf{r}_i \\ 0 \end{bmatrix}.$$

These span the subspace of 4D quaternion space consistent with this measurement. The projection matrix onto this 2D subspace is

$$q_1 q_1^{\mathrm{T}} + q_2 q_2^{\mathrm{T}} = \tfrac{1}{2}(I + \tilde{K}_i).$$

We update the quaternion by $q(+) = (I + \eta \tilde{K}_i) q(-) \big/ \left| (I + \eta \tilde{K}_i) q(-) \right|$, where $\eta = 1$ for perfect measurements, and $0 < \eta < 1$ for filtering.

Star Camera Attitude Determination (SCAD) - 1998

Write $\mathbf{r}_i = \bar{\mathbf{r}} + (\mathbf{r}_i - \bar{\mathbf{r}})$ with $\bar{\mathbf{r}} \equiv (\sum_i a_i \mathbf{r}_i)/(\sum_i a_i)$, and similarly for $\mathbf{b}_i$.

Then $$L(A) \equiv \left(\tfrac{1}{2}\sum_i a_i\right)\left|\bar{\mathbf{b}} - A\bar{\mathbf{r}}\right|^2 + \tfrac{1}{2}\sum_i a_i \left|(\mathbf{b}_i - \bar{\mathbf{b}}) - A(\mathbf{r}_i - \bar{\mathbf{r}})\right|^2.$$

For small-field-of-view sensors, the second term is much smaller than the first. The general quaternion minimizing the first term is

$$q(\psi) = q_1 \cos(\psi/2) + q_2 \sin(\psi/2)$$

where q_1 and q_2 are the two quaternions that map $\bar{\mathbf{r}}/|\bar{\mathbf{r}}|$ into $\bar{\mathbf{b}}/|\bar{\mathbf{b}}|$. Then an arctangent gives the ψ that minimizes the second term. Computation of $\bar{\mathbf{r}}$ and $\bar{\mathbf{b}}$ makes SCAD fairly slow.

Testing

MATLAB versions of the q method, the SVD method, QUEST, ESOQ, ESOQ2 and FOAM were tested. The q method used `eig` and the SVD method used `svd`. The other methods used 0, 1, or 2 iterations of the characteristic equation to compute λ_{max}.

Three test scenarios were simulated:

Star tracker scenario: 5 stars in narrow field-of-view star tracker, 6 arcsecond per star per axis measurement errors

Unequal weights: one measurement with 1 arcsecond measurement noise, two with 1° measurement noise

Mismodeled weights: two measurements with 0.1° measurement noise and one with 1°, all weighted equally.

Each scenario was tested with 1000 random attitude matrices.

Accuracy Results

All algorithms performed equally well in the star tracker scenario.

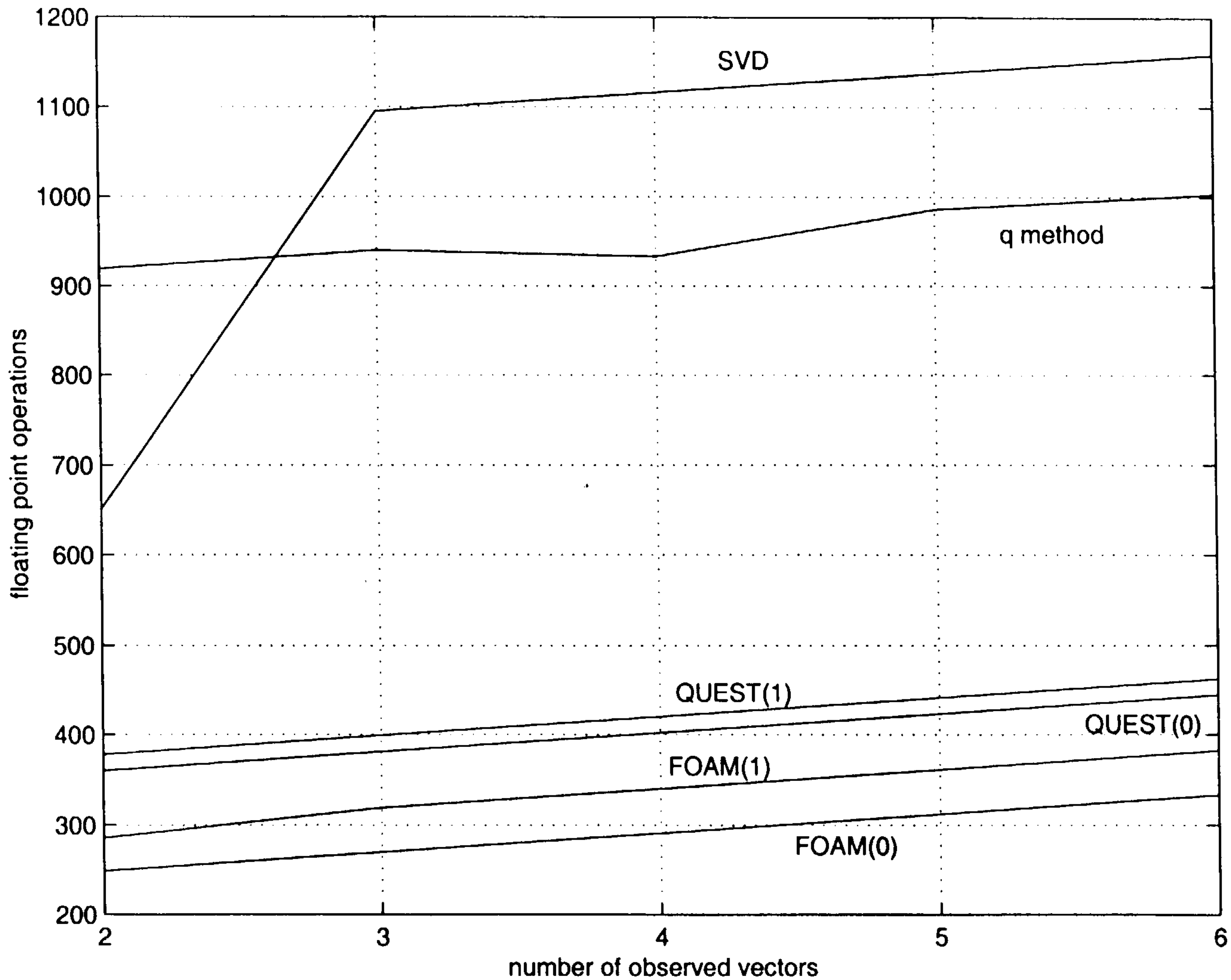
The q method, the SVD method, and FOAM performed well in the unequal weight scenario. The iterative refinement of λ_{max} failed in QUEST, ESOQ, and ESOQ2 in this scenario.

A single update of λ_{max} is required for best performance in the mismodeled weight scenario (except in the q and SVD methods).

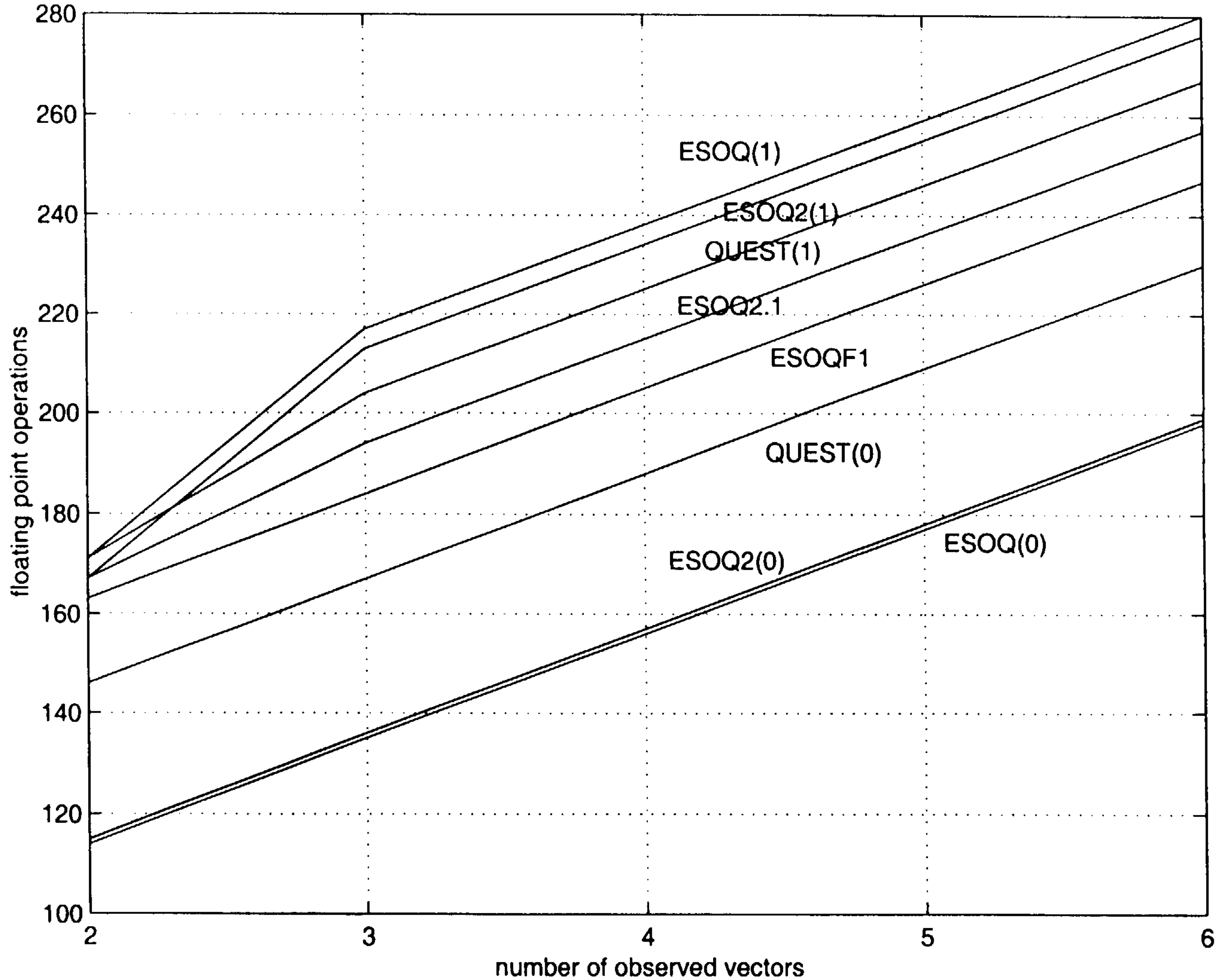
A second update of λ_{max} may improve the agreement of the estimate with the optimal (q and SVD method) attitude, but it never improves the agreement with the (simulated) true attitude.

Special first-order variants of ESOQ (ESOQF1) and ESOQ2 (ESOQ2.1) were developed to take advantage of this observation.

Timing of Slower Methods



Timing of Faster Methods



Summary

The most robust estimators minimizing Wahba's loss function are Davenport's q method and the SVD method. The q method is faster than the SVD method with three or more measurements.

The other algorithms are less robust since they solve the characteristic polynomial equation to find the maximum eigenvalue of Davenport's K matrix. They are only preferable when speed or processor power is an important consideration.

Of these, FOAM is the most robust and faster than the q method.

Robustness is only an issue for measurements with widely differing accuracies, so the fastest algorithms, QUEST, ESOQ, and ESOQ2, are well suited to star sensor applications.