SIAM Review, Vol. 8, No. 3 (Jul., 1966), 384-386.

Problem 65-1, *A Least Squares Estimate of Satellite Attitude*, by GRACE WAHBA (IBM—Federal Systems Division).

Given two sets of n points $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$, and $\{\mathbf{v}_1^*, \mathbf{v}_2^*, \cdots, \mathbf{v}_n^*\}$, where $n \geqq 2$, find the rotation matrix M (i.e., the orthogonal matrix with determinant $+1$) which brings the first set into the best least squares coincidence with the second. That is, find M which minimizes

$$\sum_{j=1}^{n} \| \mathbf{v}_j^* - M\mathbf{v}_j \|^2.$$

This problem has arisen in the estimation of the attitude of a satellite by using direction cosines $\{\mathbf{v}_k^*\}$ of objects as observed in a satellite fixed frame of reference and direction cosines $\{\mathbf{v}_k\}$ of the same objects in a known frame of reference. M is then a least squares estimate of the rotation matrix which carries the known frame of reference into the satellite fixed frame of reference.

Solution by J. L. FARRELL and J. C. STUELPNAGEL (Westinghouse Defense and Space Center).

Let k denote the dimension of the column vectors $\mathbf{v}_1, \cdots, \mathbf{v}_n, \mathbf{v}_1^*, \cdots, \mathbf{v}_n^*$ and let V and V^* denote the two $k \times n$ matrices obtained by juxtaposing $\mathbf{v}_1, \cdots, \mathbf{v}_n$ and $\mathbf{v}_1^*, \cdots, \mathbf{v}_n^*$, respectively.

For any orthogonal matrix M, define $Q(M)$ as the sum of squares to be minimized, so

$$Q(M) = \sum_{j=1}^{n} \| \mathbf{v}_j^* - M\mathbf{v}_j \|^2 = \operatorname{tr} (V^* - MV)^T(V^* - MV),$$

where tr denotes the trace function and a superscript T denotes transposition.

$Q(M)$ may be rewritten as

$$Q(M) = \operatorname{tr} (V^{*T} - V^TM^T)(V^* - MV) = \operatorname{tr} V^{*T}V^* + \operatorname{tr} V^TV - 2 \operatorname{tr} V^TM^TV^*.$$

Since the first two terms are independent of M, $Q(M)$ is minimized by maximizing $F(M) = \operatorname{tr} V^TM^TV^*$, which may be written as

$$F(M) = \operatorname{tr} M^TV^*V^T.$$

It is a well-known fact that an arbitrary real square matrix A can be written as a product UP, where U is orthogonal and P is symmetric and positive semidefinite. Furthermore, if A is nonsingular, U is uniquely defined and P is positive definite. If A is singular, U is not unique, but it may be taken to have determinant $+1$. (The corresponding statement of the first result above for complex A may be found in [1, §2.8] and the result for real A follows from it.)

Applying this result to $A = V^*V^T$, we have $F(M) = \operatorname{tr} M^TUP$. Since P is symmetric, there is an orthogonal matrix N such that NPN^T is a diagonal matrix D, whose diagonal elements $d_1, \cdots, d_k$ are arranged in decreasing order. All d_i are nonnegative, since P is positive semidefinite. Now, letting $X = NM^T \cdot UN^T$, we obtain

$$F(M) = \operatorname{tr} M^TUN^TDN = \operatorname{tr} NM^TUN^TD = \operatorname{tr} XD = \sum_{i=1}^{k} d_ix_{ii}.$$

Since $F(M)$ is a linear function of the nonnegative numbers $d_1, \cdots, d_k$, its maximum is attained when the diagonal elements of X attain their maximum values. Because X is an orthogonal matrix, all elements of X are between -1 and 1, so $F(M)$ is maximized when $x_{ii} = 1, x_{ij} = 0, i \neq j$.

Because $\det M$ is required to be $+1$, $\det X = \det (NM^TUN^T) = (\det N)^2 \cdot \det M \det U = \det U$. If $\det U = -1$, then it is required that $\det X = -1$, and it is not hard to see that

$$X = \begin{pmatrix} I_{k-1} & 0 \\ 0 & -1 \end{pmatrix}$$

is a solution (since $d_1 \geqq d_2 \geqq \cdots \geqq d_k$). Letting X_0 be the matrix which maximizes $F(M)$ ($X_0 = I$ or $X_0 = \begin{pmatrix} I_{k-1} & 0 \\ 0 & -1 \end{pmatrix}$, according as $\det U = +1$ or -1), $X_0 = NM_0^TUN^T$, or $M_0 = UN^TX_0^TN$ is a rotation matrix which minimizes the sum of squares $Q(M)$. If V^*V^T is nonsingular, it is the unique rotation matrix which does so.

REFERENCE

[1] M. MARCUS, *Basic Theorems in Matrix Theory*, Nat. Bureau of Stds., Appl. Math. Series No. 57, 1960.

R. H. WESSNER (Hughes Aircraft Company) in his solution points out that if $\det A \neq 0$, then $V^*V^T = A = UP$,

$$U = (A^T)^{-1}(A^TA)^{1/2}, \quad P = (A^TA)^{1/2},$$

where $(A^TA)^{1/2}$ is the symmetric square root of A^TA with positive eigenvalues, and, hence, for $\det A > 0$,

$$M_0 = (VV^{*T})^{-1}(VV^{*T}V^*V^T)^{1/2}.$$

J. R. VELMAN (Hughes Aircraft Company) in his solution demonstrates that in the case $\det A < 0$, $M_0 = U(I - 2G)$ where G is any one-dimensional projection satisfying $GE_1 = G$, where E_1 is the eigenspace of the smallest eigenvalue of P, hence Farrel and Stuelpnagle's solution in this case is unique if the smallest eigenvalue of P has multiplicity one.

J. E. BROCK (U. S. Naval Postgraduate School) solved the problem for $\det V^*V^T \geqq 0$ by differentiating

$$\alpha = -\operatorname{tr}[V^TM^{-1}V^* + V^{*T}MV]$$

with respect to each of the 9 elements of M and setting the results equal to 0. The resulting equations turn out to be

$$M^TAM^T = A^T,$$

which implies that M^TA is symmetric, $(M^TA)(M^TA) = A^TA$, M^TA is any symmetric square root $(A^TA)^{1/2}$ of A^TA, and $M = (A^T)^{-1}(A^TA)^{1/2}$. He then gives an example in which the actual residual sum of squares is minimized by taking the positive definite symmetric square root.

Also solved by R. DESJARDINS (Goddard Space Flight Center) and the proposer.