Rotational Superposition and least squares: SVD or Quaternions? Reply to the preceding comment by G. Kneller

Evangelos A. Coutsias[1], Chaok Seok[2], and Ken A. Dill[3]
[1]Department of Mathematics and Statistics,
University of New Mexico,
Albuquerque, New Mexico 87131.
[2]School of Chemistry,
Seoul National University,
Seoul, 151-742, Republic of Korea.
[3]Department of Pharmaceutical Chemistry,
University of California in San Francisco,
San Francisco, California 94107.

June 16, 2005

1 Introduction

The Optimal Rotational Superposition problem is an old one. For the case of the least squares residual (RMSD)

$$E = \frac{1}{N}\sum_{i=1}^{N} w_i|\mathbf{x}_i - \mathbf{y}_i|^2 = \frac{1}{N}||\sqrt{W}(\mathcal{X} - \mathcal{Y})||_F^2 \tag{1}$$

two approaches, using the SVD of the matrix $\mathcal{R} := \mathcal{X}\mathcal{Y}^T$ ([1], formula (5)) and the spectral analysis of the quaternion matrix $\mathcal{F}$ ([1], formula (10)) have been available for some time together with several generalizations [2]. Here we follow the notation of Coutsias et al. [1]; $\mathcal{X}$ and $\mathcal{Y}$ are the model and target sets respectively and the subscript F denotes the Frobenius norm. For simplicity we omit the positive weights w_i. Given the importance of the problem, there did appear to be a need of a deeper understanding of these two seemingly different mathematical answers to the same mathematical question. The analysis in [1] elucidates how the two approaches inform

each other, and we feel that there is nothing misleading in our proof of the algebraic identity between the two methods, especially given the extensive discussion of chirality and its encoding in the spectrum of $\mathcal{F}$ and the singular values and vectors of $\mathcal{R}$. Of course, there were problems with the original lengthy derivation by Kabsch [3], which used Lagrange multipliers to enforce orthogonality for the optimal rotation matrix, resulting in transformations of ambiguous chirality. This problem however lies only with Kabsch's original method of proof, not with the SVD approach per se. The commonly accepted proof [4] is based on the connection between the trace of the matrix $\mathcal{Y}^T\mathcal{X}$ and the properties of the singular vectors of $\mathcal{R}$ and is entirely unambiguous, as discussed at length in [1], p. 1851.

In this regard, there is no new information in Section 2 of the comment by Kneller [5], only a restatement of the analysis of the arrangement of the spectrum in Coutsias et al.[1] given for his slightly different representation of the residual in terms of the matrix $\mathcal{M}$ ([5], formula (10)). Finally, better bounds can be found for the spectrum for the general case of non-collinear vectors using Gershgorin's theorem [4], while the bounds quoted by Kneller ([5], formula (23)) are problematic and they are corrected below.

2 The SVD and Quaternion approach yield identical results

The main point in Coutsias et al. [1] was that the Procrustes [6] (or, for the chemists, Kabsch's [3]) method and the Quaternion method [7, 8, 9] are mathematically equivalent problems, i.e. that they contain identical information and, when properly understood and applied, lead to identical answers to any question regarding least-squares rotational superposition, either by proper rotations or by rotation–reflections. The key insight for the Kabsch singular vectors is that the optimal rotations align or antialign them according to the sign of the Kabsch matrix determinant, $\det\mathcal{R}$. There may be circumstances (not in the design of protein-like molecules to be certain!) where a chiral reversal might be allowed. It is clear from the analysis in [1] that the arrangement of the eigenvalues of the Quaternion matrix contains equivalent information with that found from analysis of the chirality of the singular vectors. We demonstrated that the sign of the eigenvalue of the largest magnitude is the same as the sign of the Kabsch matrix determinant. Therefore, when both methods are correctly understood one can accomplish the same things with either method with no extra work or special case checks. Of course, computing the determinant of the Kabsch matrix adds

insignificant overhead (the cost of a cross and a dot product) compared to the work to set up the matrices in either method and the work to find the SVD of a 3x3 matrix or to that of finding the spectrum of a 4x4 traceless matrix. We can contrast the two methods as follows

1. **The Kabsch method:**
   Form matrix $\mathcal{R} = \mathcal{X}\mathcal{Y}^T$ and perform the SVD, $\mathcal{R} = \mathcal{V}\Sigma\mathcal{W}^T$.
   Form the optimal rotation matrix $\mathcal{W}\mathcal{X}\mathcal{V}^T$ where matrix $\mathcal{X}$ is the identity except that its entry $\chi_{33} = sgn \det \mathcal{R}$.
   If $det\mathcal{R} = 0$ then there is at least one zero singular value and the sets are chirality-indifferent.

   In comparing sizes of residuals between pure rotations or rotation-reflections, it is clear that the smallest value occurs when subtracting the sum of all singular values, which results from the alignment of all pairs of singular vectors. That can be accomplished without a reflection if the two sets of singular vectors have the same chirality so that $det\mathcal{R} > 0$. Otherwise a reversal is required and the best that can be done by a pure rotation is to align the pairs of right and left singular vectors for the two largest singular values, while antialigning that of the smallest.

2. **The Quaternion method:**
   Construct the matrix $\mathcal{F}$ ([1], form. (10)) and compute its eigenvalues, $\mu_1 \geq \mu_2 \geq \mu_3 \geq \mu_4$, and eigenvectors, q_i , $i = 1, 2, 3, 4$.
   If the eigenvalue of largest magnitude is positive (μ_1), corresponding eigenvector q_1 gives matrix $U(q_1)$ for optimal rotation superposition.
   Conversely, if the eigenvalue of largest magnitude is negative (μ_4), then $U(q_1)$ provides only rotational optimality, while the matrix $-U(q_4)$, i.e. a rotation based on the most negative eigenvalue followed by a reflection, provides a better, but chirality-reversing, optimum.
   If the smallest and largest eigenvalue have equal magnitudes the sets are chirality-indifferent, e.g. planar or linear.

In our view, these are equivalent statements.

3 Remarks

3.1 Chirality and Degeneracy

The ambiguity in Kabsch's original paper [3] was noted by Nyburg and Yuen and addressed by Kabsch in a subsequent publication [10]. The proper

treatment of the chirality question in terms of the right and left singular vectors of the correlation matrix $\mathcal{R}$ has been in standard textbooks for some time [4].

The fact that the rotation operator as derived from the quaternion eigenvectors is proper is well understood [11], as is the distinction between proper and enantiomeric superpositions. It is pointed out in [1] that a slight advantage of the quaternion method is its efficient coding of the degenerate case, since the invariant subspace of a degenerate eigenvalue is spanned by linear combinations of its eigenvectors, while the SVD based method does not readily provide such a representation since the relation between the quaternion and its rotation matrix is nonlinear. The simple expression for the invariant subspace did appear in [12], without contrast to the SVD picture. Apart from this case, which is discussed in [1] the two methods (SVD and quaternion) do indeed produce mathematically equivalent descriptions of the problem, and contrasting them does help gain a clearer understanding of rotational superposition: the singularity of the correlation matrix $\mathcal{R}$ reveals degeneracy, and the sign of its determinant gives the chirality of optimal superposition directly and simply, without the need of computing spectra. Here we must mention that Dr. Kneller's independent derivation of the quaternion RMSD, given in his elegant 1991 article [12] was missed by the 1999 review by Flower [2], as well as in the overview given in [1]. In any case, the argument regarding chirality and quaternions is an old one [11], and there seems to be no need to repeat it here.

3.2 Spectral estimates

Sharp bounds for the eigenvalues of $\mathcal{F}$ (or for $\mathcal{M}$) can be found easily, by applying standard Gershgorin [4] estimates to the quaternion matrix $\mathcal{F}$. Thus, the (real) eigenvalues of $\mathcal{F}$ are found on the union of the intervals

$$I_i := \left[F_{ii} - \sum_{j \neq i} |F_{ij}|, F_{ii} + \sum_{j \neq i} |F_{ij}| \right] \ , \ i,j = 1,2,3,4 \ . \tag{2}$$

Taking advantage of further results along the lines of (2) and the traceless character of $\mathcal{F}$ one can refine these estimates, and speed up the calculation of the spectrum of $\mathcal{F}$. These estimates are optimal but their geometrical content is obscure. One can proceed directly from the definition of the residual (1) and apply the triangle inequality to arrive at bounds with clearer geometrical content. These bounds, however, are in general not sharp except for perfectly alignable or collinear sets. In the preceding comment [5], an

estimate based on the triangle inequality is claimed, however the estimate (21) for the spectrum of the matrix $\mathcal{M}$ is one-sided and its conversion to the estimate (23) for the spectrum of $\mathcal{F}$ is incorrect. A careful application of the triangle inequality gives:

$$0 \leq \sum_a w_a(|\mathbf{x_a}|-|\mathbf{y_a}|)^2 \leq NE = \sum_a w_a(\mathbf{x_a}-\mathbf{y_a})^2 \leq \sum_a w_a(|\mathbf{x_a}|+|\mathbf{y_a}|)^2 \quad (3)$$

from which follows the symmetric estimate for the eigenvalues of $\mathcal{F}$:

$$-\sum_a w_a|\mathbf{x_a}||\mathbf{y_a}| \leq \mu_i \leq \sum_a w_a|\mathbf{x_a}||\mathbf{y_a}| \quad (4)$$

These bounds are realized, for example, for two collinear sets of (generally unequal) vectors. The spectrum of $\mathcal{F}$ obeys symmetric bounds e.g. in the case of double degeneracy, when the singular values of $\mathcal{R}$ are $\sigma_1 > \sigma_2 = \sigma_3 = 0$ since then $\mu_1 = \mu_2 = \sqrt{\sigma_1} = -\mu_3 = -\mu_4$.

References

[1] Coutsias, E. A., Seok, C. and Dill, K. A., J. Comput. Chem. **25**, 1849, (2004).

[2] Flower, D. L., J. Comput. Graph. Model. **17**, 238, (1999).

[3] Kabsch, W. Acta Cryst. 1976, A32, 922;

[4] Horn, R. A.; Johnson, C. R. Matrix Analysis, Cambridge Univ. Press, (1985).

[5] Kneller, G. , this issue, (2005).

[6] Schöneman, P. H. Psychometrika, **31(1)**, 1, (1966).

[7] Diamond, R. , Acta Cryst. **A44**, 211, (1988).

[8] Kearsley, S. K. , Acta Cryst. **A45**, 208, (1989).

[9] Diamond, R. , Acta Cryst. **A45**, 657, (1989).

[10] Kabsch, W. Acta Cryst. **A34**, 827, (1978).

[11] Diamond, R. , Acta Cryst. **A46**, 423, (1990).

[12] Kneller, G. , Mol. Sim. **7**, 113, (1991).