Acta Cryst. (1977). A33, 212–215

The Generalized Inverse and Inverse Structure

BY A. L. MACKAY

Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, England

(*Received* 13 *July* 1976; *accepted* 16 *September* 1976)

When a structure in one space is projected or mapped or otherwise described in another space or 'language', then the transformation is usually irreversible. In the case of linear transformations a generalized inverse matrix exists even if the transformation matrix is rectangular or singular. This inverse represents the best that can be done by way of reverse transformation. Typical crystallographic applications of this inverse are developed. Some are in practical computing, where failures due to singularity of a matrix can be avoided. Others are in the theoretical treatment of redundant axes, such as are usual in the description of hexagonal crystals using four indices and four axes. The idea of the inverse of a set of crystallographic axes, normally the reciprocal axes, can be extended to the concept of the inverse of any set of coordinates.

Introduction

Suppose that a system described in space A is transformed, projected or mapped into space B. In general information will be lost and the transformation will not be reversible unless external information is introduced. Nevertheless, a knowledge of the structure as described in space B imposes severe limitations on the structure in space A from which it could have been derived.

This situation is extremely common and occurs whenever a structure of any type is represented, or translated, or described in a different 'language'. We will, however, confine ourselves here to crystallographic situations.

The case in which space A is a crystal structure (real space), space B is its weighted reciprocal lattice (showing intensities) and the operation, one of taking the square of the Fourier transform, is familiar. The Patterson function then represents in space A (real space) all the information available from space B (Fourier amplitude space). It is not the same as the original real-space structure.

If the transformation, projection or mapping is linear and represented by a matrix, then this matrix, even though it may have no ordinary inverse, has a generalized inverse, which represents the best that can be done by way of reverse transformation.

The generalized inverse of a matrix can be applied in numerous crystallographic calculations and to developments in structural theory.

The generalized inverse of a matrix

For every finite matrix P with n rows and m columns of real or complex elements, there is a unique matrix Q, satisfying the following relations (where P^* is the complex transpose of P and capital letters denote arrays):

$$PQP = P$$
$$QPQ = Q$$
$$(PQ)^* = PQ$$
$$(QP)^* = QP\,.$$

A matrix which satisfies all four conditions is called the Moore–Penrose inverse of P, or, less clearly, the generalized or pseudo inverse of P, since other inverses satisfying only some of the conditions can also be so called. Every matrix, albeit singular or rectangular, has an inverse of this type, so that calculation can be continued even if there is no ordinary inverse. The set of linear equations $PX = H$ [or, giving the dimensions of the arrays necessary in setting up a programme, $P(n,m)X(m,1) = H(n,1)$] thus has solutions in all cases which are, quoting well-known results, as follows (Q, the generalized inverse of P, is denoted by P^+):

(*a*) $PX = H$ has the solution $X = P^{-1}H$ if P has an ordinary inverse, that is, is square and non-singular. In these circumstances the generalized inverse is the same as the ordinary inverse.

(*b*) $P(m,n)X(n,1) = H(m,1)$ and $m > n$. Here there are more equations than unknowns and X may be over-determined. If the surplus $(m-n)$ equations are linear combinations of the others, the solution $X = P^+H$ gives the exact answer, and if the equations are inconsistent with each other, being observational equations, but are sufficient in number, then this expression $X = P^+H$ gives the least-squares answer where $\|H - PX\|^2$, the square of the Euclidean norm, is minimized.

(*c*) If X is under-determined by $PX = H$, and P is singular, or there are less equations than unknowns and the rank r of P is less than n, then the generalized inverse still gives a unique solution, namely that for which $\|X\|^2$ is minimized (Wilkinson & Reinsch, 1971). The full solution is $X = P^+H + [I - P^+P]Z$ where Z is an arbitrary vector.

This corresponds to the case in which an attempt is made to restore solid structure from a projection. For example, the transformation:

$$(X') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (X)$$

projects on to the x_1,x_2 plane but when the inverse transformation to three dimensions is attempted, the x_3 coordinate remains unknown or arbitrary. Here $P^+=P$

and

$$X=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}(X')+\left(\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}\right)Z.$$

Method of calculation

The simplest algorithm for calculating the Moore–Penrose inverse has been given by, among others, Gupta (1971). If Q_K is the Kth approximation to the generalized inverse of the matrix P, then an improved approximation Q_{K+1} can be obtained from $Q_{K+1}=Q_K(2I-PQ_K)$. The trace of PQ_K is a monotonically increasing function of K and converges to the rank of P. Iteration is thus stopped when this trace is sufficiently close to an integral value.

The starting matrix Q_1 is taken as a positive constant α times P^*. Properly, α should be less than $2/\lambda_{max}$ where λ_{max} is the largest eigenvalue of PP^* or P^*P, but it is easier simply to take a small value of α and, if this is not small enough for convergence, to reduce it and to try again. This method is well-known for ordinary inverses (Booth, 1955).

Simple programs have been written in BASIC taking advantage of the concise matrix statements. PQ_K may use a square array of order m, but it is not necessary to employ such an array and, if $n\ll m$, a great saving in storage can be realized by generating only one column of PQ_K at a time.

The Nottingham Algorithms Group have produced an implementation (NAG Library, *F01BHF*) of the singular value decomposition programme described by Golub & Reinsch (Wilkinson & Reinsch, 1971) from which generalized inverses can be obtained to the highest standards.

Comparison of two molecules

A typical application is in the comparison of two similar molecules, each of N atoms, which may, for example, occur in the same crystal in non-symmetry-related positions. It is desired to bring one as nearly as possible into coincidence with the other by rotation and translation and to compare the residual discrepancies of position between corresponding atoms. The steps are as follows:

(*a*) Reduce each molecule to orthonormal coordinates $X_A(N,3)$ and $X_B(N,3)$ with respect to origins at the centres of gravity.

(*b*) The two sets of coordinates are then to be related to each other by a general linear transformation matrix $A(3,3)$ fitted by least squares. Thus $X_BA=X_A$ is to be solved for A with $||X_BA-X_A||^2$ minimized, which is the sum of the squares of the discrepancies in position. The solution is then $A=X_B^+X_A$, when the generalized inverse of X_B is taken and the actual discrepancies in coordinates are (X_BA-X_A).

(*c*) The transformation matrix A can then be analysed if necessary, by calculating its eigenvalues, into rotational and dilational parts. Since its dimensions are only 3×3 this is conveniently done by solving the characteristic equation, the coefficients of which are the trace, the second invariant and the determinant of the matrix A. This latter procedure is convenient for comparing coordination polyhedra, as encountered in crystals, with the idealized geometrical figures, from which they might be considered to be distorted.

(*d*) Fletterick & Wyckoff (1975) have incorporated a vector **b** representing the change of centre of gravity into the transformation matrix A, which then has twelve components:

$$(X_{B1},X_{B2},X_{B3})\begin{pmatrix}a_{11}&a_{12}&a_{13}&b_1\\a_{21}&a_{22}&a_{23}&b_2\\a_{31}&a_{32}&a_{33}&b_3\end{pmatrix}=(X_{A1},X_{A2},X_{A3},1).$$

This is an alternative to shifting each molecule to its centre of gravity.

The inverse structure of a molecule

Suppose it is required to compare a range of molecules against a standard and to know the transformation matrix for the best fit at any stage. That is, A is required for a number of sets of coordinates X_B which are each matched against a standard set X_A. Since $A=X_A^+X_B$ the calculation of the generalized inverse of X_A need be performed once only and A can be found by simple multiplication for each set of X_B.

Further, X_A^+ can thus be regarded as a kind of inverse structure existing independently. Properly, it consists of three points in N-dimensional space but could be represented by N points in three dimensions.

This concept is already used in the relationship between real and reciprocal lattice vectors but seems capable of generalization.

The sets of points for which the best orientation matrix is to be determined repeatedly may be either atoms in real space or sets of reciprocal lattice points in transform space.

The reciprocal lattice

The matrices of the components of the real and reciprocal lattice vectors with respect to orthonormal axes are inverse to each other.

$$\text{If } A=\begin{pmatrix}a_x&a_y&a_z\\b_x&b_y&b_z\\c_x&c_y&c_z\end{pmatrix} \quad \text{and } A_R^T=\begin{pmatrix}a_x^*&a_y^*&a_z^*\\b_x^*&b_y^*&b_z^*\\c_x^*&c_y^*&c_z^*\end{pmatrix}$$

then $AA_R=I=A_RA$.

Consequently, the reciprocal unit cell can be found by taking the inverse of the real unit cell, a simplex of

four points (the origin and the ends of the three vectors) defining each. Clearly this can be done also in higher-dimensional spaces. We have also $AA^T=G$, the metric matrix, and $A_RA_R^T=G^{-1}$, the metric matrix for reciprocal space.

Redundant crystallographic axes

Suppose that we apply this procedure to the hexagonal axes $XYUZ$ where there is a redundant axis, but take the generalized inverse, since a rectangular matrix has no ordinary inverse:

$$A=\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2}\sqrt{3} & 0 \\ -\frac{1}{2} & \frac{1}{2}\sqrt{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ gives } A^+=\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \\ 0 & 1/\sqrt{3} & -1/\sqrt{3} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

so that $A^+A=I(3,3)$ and the product AA^+ is

$$\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In each case the trace, and thus the rank, of the product is 3. The elements of AA^+ are the scalar products $\mathbf{a}_i.\mathbf{a}_j^*$ and are proportional to the cosines of the angles between the four real and the four reciprocal axes. The constants of proportionality are found from AA^T and A^+A^{+T} which give $|a_i|^2$ and $|a_i^*|^2$ on their diagonals. Comparing A and A^+ shows that the real and the reciprocal axes, if plotted in the same three-dimensional space, are parallel to each other.

The above gives the system of reciprocal axes discussed by Frank (1965) where a point with indices *hkil* (representing the normal to the planes (*hkil*), which make intercepts $a_1/h, a_2/k, a_3/i, a_4/l$ on the four unit-cell axes) has the position vector $h\mathbf{a}_1^*+k\mathbf{a}_2^*+i\mathbf{a}_3^*+l\mathbf{a}_4^*$ in the reciprocal space. Such is not the case for the conventional Miller–Bravais system where the redundant index i is ignored (Nicholas & Segall, 1970), a redundant reciprocal axis is not defined, and the first two reciprocal axes are 60° and not 120° apart. In both cases the redundancy condition is $h+k+i=0$ in reciprocal space and $\mathbf{a}_1+\mathbf{a}_2+\mathbf{a}_3=0$ in real space.

The tetrahedron

For some purposes requiring emphasis on the four threefold axes it might be convenient to describe cubic or tetrahedral structures with respect to four equivalent axes (along the [111] directions). Rogers & Klyne (1972) have described systems of tetrahedral coordinates used by themselves and by others and have suggested a number of applications. Suppose the orthonormal coordinates of the ends of the axes are:

$$A=\begin{pmatrix} 1 & 1 & 1 \\ \bar{1} & \bar{1} & 1 \\ 1 & \bar{1} & \bar{1} \\ \bar{1} & 1 & \bar{1} \end{pmatrix}$$

then the generalized inverse is

$$A^+=\begin{pmatrix} \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{pmatrix}.$$

The four reciprocal axes are just parallel to the four real axes. The indices of the planes are then subject to the rule that $h+k+i+l=0$. Calculations are carried out in the normal way. For example, what is the spacing of the plane $(3\bar{1}\bar{1}\bar{1})$ (an octahedral plane)?

$$\mathbf{r}^*=3\mathbf{a}_1^*-\mathbf{a}_2^*-\mathbf{a}_3^*-\mathbf{a}_4^*$$

and converting to vector components along orthonormal unit vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$,

$$\begin{aligned} \mathbf{r}^*= & \tfrac{3}{4}\mathbf{x}+\tfrac{3}{4}\mathbf{y}+\tfrac{3}{4}\mathbf{z} \\ & +\tfrac{3}{4}\mathbf{x}+\tfrac{1}{4}\mathbf{y}-\tfrac{1}{4}\mathbf{z} \\ & -\tfrac{3}{4}\mathbf{x}+\tfrac{1}{4}\mathbf{y}+\tfrac{1}{4}\mathbf{z} \\ & +\tfrac{1}{4}\mathbf{x}-\tfrac{1}{4}\mathbf{y}+\tfrac{1}{4}\mathbf{z} \\ = & \mathbf{x}+\mathbf{y}+\mathbf{z}, \text{ thus } (\mathbf{r}^*.\mathbf{r}^*)^{1/2}=3^{1/2} \\ & \text{and } d=|\mathbf{r}^*|^{-1}=3^{-1/2}. \end{aligned}$$

The icosahedron

The same procedure can be applied to the six fivefold axes in real space which join the origin to the vertices of a regular icosahedron. Just as a cell with hexagonal symmetry is described using four axes, one redundant, to preserve the symmetry, so an icosahedron should be described using six equivalent axes. In reciprocal space its transform will have the same properties.

In this notation the coordinates of a point are $(x_1,x_2,x_3\ldots x_6)$ and $\mathbf{r}=x_1\mathbf{a}_1+x_2\mathbf{a}_2+\ldots+x_6\mathbf{a}_6$. Correspondingly, if a plane makes intercepts (a_i/h_i) on the six axes, it has the indices $(h_1,\ldots,h_6)$ and the vector from the origin of reciprocal space to the point representing the set of planes of spacing d is $\mathbf{r}^*=h_1\mathbf{a}_1^*+\ldots+h_6\mathbf{a}_6^*$.

The twelve vertices of a regular icosahedron thus have coordinates $(\pm1,0,0,0,0,0)$, $(0,\pm1,0,0,0,0)$ *etc.*, with respect to the icosahedral axes. The indices of a plane which is one of the faces of an icosahedron are $(1,1,1,s,s,s)$ [where $s=0{\cdot}236068=1/(2\tau+1)$ and $\tau=\frac{1}{2}(1+\sqrt{5})$], and thus the reciprocal lattice point representing the planes is at

$$\mathbf{r}^*=1\mathbf{a}_1^*+1\mathbf{a}_2^*+1\mathbf{a}_3^*+s\mathbf{a}_4^*+s\mathbf{a}_5^*+s\mathbf{a}_6^*.$$

The actual spacing of this set of planes for unit axial lengths is $\tau^{3/2}5^{-1/4}3^{-1/2}=0{\cdot}794654$.

Since the icosahedron is not the unit cell of a lattice, the coefficients are not necessarily integral. In this way, by inversion, we obtain an inverse structure which has the same point-group symmetry as the original structure with respect to the origin chosen.

Sets of reciprocal base vectors occur in other contexts. For example, Weinhold (1975, 1976) has recently shown that the relationships between thermodynamic quantities can be represented as a set of vectors in N-dimensional space for extensive variables (such as entropy, S; volume, V; mole numbers, N_i) each of which is associated with a corresponding intensive variable (T, P, K_i). The latter can be represented as vectors reciprocal to the first set. The rank of the Gram determinant gives the dimensionality of the space and thus the number of degrees of freedom of the system.

Real-space refinements of coordinates

The convenience of the generalized inverse in the calculation of molecular vibrations (Gellai & Jancsó, 1972), since there are $3N$ Cartesian coordinates but only $3N-6$ normal modes, is well recognized. Similarly, the $3N$ orthonormal coordinates of a molecule can be refined to fit an arbitrary number of internal coordinates (bond distances, bond angles and torsion angles), so that any other internal dimensions may be calculated. This has been referred to as Carnot's problem (Mackay, 1974). A general program (*MOLECQ*) to do this has been written in BASIC and is designed for time-sharing execution.

Experience shows that, naturally enough, mathematical behaviour parallels physical behaviour. Suppose that for a regular helix of ten atoms the nine distances, eight bond angles and seven torsion angles (24 parameters in all), which are the minimum nearest-neighbour relationships to define the structure, are given. It is required to refine the 30 coordinates ($x_i y_i z_i$) until a set giving exact correspondence with the internal parameters is obtained. The problem is non-linear and the refinement to minimize the discrepancies between the internal parameters for the current configuration and those required proceeds in cycles. The changes to be applied to the 30 coordinates must be calculated from the 24 discrepancies. The matrix for this set of linear equations

$$\Delta Q_i = \frac{\partial Q_i}{\partial x_1}\Delta x_1 + \frac{\partial Q_i}{\partial x_2}\Delta x_2 \ldots \frac{\partial Q_i}{\partial x_{24}} \quad \text{(for } i = 1 \text{ to } 30\text{)}$$

(ΔQ_i represent the discrepancies between current and desired parameters appropriately weighted) is, of course, singular but, if the generalized inverse is taken instead of the ordinary inverse, refinement proceeds smoothly and all the (xyz) coordinates converge to a final value.

The conventional procedure is to fix six of the coordinates so that only 24 have to be found from 24 equations. It is difficult to choose the six fixed values elegantly. If one end of the helix is fixed then the atoms next to it must be refined into position before meaningful corrections can be applied to those much further away. Convergence is thus slow. If the generalized inverse is used, the molecule 'floats' with six degrees of freedom and convergence of several atoms can take place simultaneously.

Since we are concerned only with internal relationships then the 'floating' is of no consequence. However, it may be necessary to refine the shape of a molecule towards fitting an incomplete range of internal parameters, but at the same time making it fit a set of external coordinates, such as the peaks in an electron density distribution. In this case a number of positional equations are added to the above set of internal equations:

$$\Delta x_i = 1 \,.\, \Delta x_i \,.$$

The weights with which these positional discrepancy equations are added to those concerning discrepancies of internal parameters must be judged by trial. Adding such terms has the effect of making the matrix non-singular, adding weight to the diagonal terms.

Refinement of structure factors

The previous section has dealt with the refinement of atomic positions to minimize discrepancies between current and desired real-space parameters. The generalized inverse can also be used in reciprocal-space refinement of calculated structure factors to match those observed.

Using the generalized inverse instead of the ordinary inverse may be particularly useful in space groups such as $P2_1$ where there is a screw axis along y which leaves an arbitrary y parameter to be fixed. Considering the fixing of the y parameter of a single atom in a large molecule as a mechanical restraint applied while a large internally sprung assembly is being shaken into position convergence can be seen by physical analogy to be poor. If no constraint on y is applied then the matrix will be singular but the generalized inverse will allow the molecule to float uniformly.

An example of the refinement of a structure in $I4_1$ has been tried and the results will be reported separately.

References

BOOTH, A. D. (1955). *Numerical Methods*, p. 80. London: Butterworth.

FLETTERICK, R. J. & WYCKOFF, H. W. (1975). *Acta Cryst.* A**31**, 698–700.

FRANK, F. C. (1965). *Acta Cryst.* **18**, 862–866.

GELLAI, B. & JANCSÓ, G. (1972). *J. Mol. Struct.* **12**, 478–481.

GUPTA, N. N. (1971). *IEEE Trans. Systems, Man and Cybernetics*, January 1971, pp. 89–90.

MACKAY, A. L. (1974). *Acta Cryst.* A**30**, 440–447.

NICHOLAS, J. F. & SEGALL, R. L. (1970). *Acta Cryst.* A**26**, 522–527.

ROGERS, D. & KLYNE, W. (1972). *Tetrahedron Lett.* **15**, 1441–1444.

WEINHOLD, F. (1976). *Phys. Today*, March 1976, pp. 23–30.

WEINHOLD, F. (1975). *J. Chem. Phys.*, **63**, 2479, 2484, 2488, 2496.

WILKINSON, J. M. & REINSCH, C. (1971). *Handbook for Automatic Computation. II. Linear Algebra*, pp. 134–151. Berlin: Springer.