Matrix mutual orthogonality and parameter independence

1981 J. Phys. A: Math. Gen. 14 L429

(http://iopscience.iop.org/0305-4470/14/11/002)

J. Phys. A: Math. Gen. **14** (1981) L429–L431. Printed in Great Britain

LETTER TO THE EDITOR

Matrix mutual orthogonality and parameter independence

D J Newman

Department of Physics, University of Hong Kong, Hong Kong

Received 18 August 1981

Abstract. It is shown that uncorrelated parameters can be derived from a data matrix if $\mathrm{Tr}(\boldsymbol{\varepsilon}^2)$ is minimised rather than $\Sigma_i \varepsilon_{ii}^2$, where $\boldsymbol{\varepsilon}$ is the error matrix.

It is well known that the linear least-squares fitting procedure determines uncorrelated parameters if orthogonal basis vectors are used in fitting a data vector (Hudson 1964). The problem that concerns us here is whether an analogous matrix basis $\boldsymbol{A}_\mu$ exists in the case where we wish to fit a data matrix to a linear expression of the form $\Sigma_\mu \boldsymbol{A}_\mu \theta_\mu$ where the θ_μ are uncorrelated parameters. We begin by investigating the properties of 'mutually orthogonal' matrices and then show how these are relevant in fitting parameters to spectroscopic data.

The standard definition of mutual orthogonality of two matrices $\boldsymbol{A}$, $\boldsymbol{B}$ is $\mathrm{Tr}(\boldsymbol{A}\boldsymbol{B}^\dagger) = 0$, where the dagger indicates the adjoint or Hermitian conjugate of $\boldsymbol{B}$. This definition is appropriate in an inner product space of complex matrices (e.g. see Hoffman and Kunze 1971). We shall, however, only be interested in the space of Hermitian matrices, so the above expression may be written in the simplified form $\mathrm{Tr}(\boldsymbol{A}\boldsymbol{B}) = 0$. All the standard mathematical results for inner product spaces then apply. In particular, we can find a set of N^2 mutually orthogonal Hermitian matrices $\boldsymbol{A}_\mu$ of dimension $N \times N$. These matrices have a positive norm and can thus be normalised by requiring $\mathrm{Tr}(\boldsymbol{A}_\mu^2) = 1$. We can therefore construct matrix bases $\boldsymbol{A}_\mu$ which satisfy

$$\mathrm{Tr}(\boldsymbol{A}_\mu \boldsymbol{A}_\nu) = \delta_{\mu\nu}.$$

If $\boldsymbol{A}_\mu$ is such a basis, then so is $\boldsymbol{U}^\dagger \boldsymbol{A} \boldsymbol{U}$, where $\boldsymbol{U}$ is unitary.

Consider, now, the linear least-squares problem of fitting an expression of the form $\Sigma_\mu \boldsymbol{A}_\mu \theta_\mu$ to a (Hermitian) data matrix $\boldsymbol{Y}$. As there are at most N^2 data it is necessary to use physical arguments to truncate the basis so that the number of parameters θ_μ is $M < N^2$. Writing the deviation matrix $\boldsymbol{\varepsilon} = \boldsymbol{Y} - \Sigma_\mu \boldsymbol{A}_\mu \theta_\mu$ we can minimise the positive expression $\mathrm{Tr}(\boldsymbol{\varepsilon}^2)$ with respect to the θ_μ to obtain the fitted values

$$\hat{\theta}_\mu = \mathrm{Tr}(\boldsymbol{A}_\mu \boldsymbol{Y}).$$

This result makes it apparent that the fitted values of the parameters $\hat{\theta}_\mu$ are uncorrelated. A more formal proof can be given by constructing the variance/covariance matrix.

We frequently wish to determine parameters θ_μ given only the *eigenvalues* of $\boldsymbol{Y}$. In this case the relation between the basis sets corresponding to $\boldsymbol{Y}$ and the $\boldsymbol{A}_\mu$ is not

0305-4470/81/110429+03$01.50

known, so the fitting equation becomes

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \sum_{\mu} \boldsymbol{U}^{\dagger}\boldsymbol{A}_{\mu}\boldsymbol{U}\theta_{\mu}$$

where $\boldsymbol{U}$ is unitary, but otherwise undetermined. In this case the number of θ_{μ}, M, must be less than N. It is usual to use an iterative procedure to solve this problem, starting with guessed values of the θ_{μ} and then carrying out a linear least-squares fit (with fixed $\boldsymbol{U}$) and diagonalising the last term alternately.

Again, the final fit to the iterated basis $\boldsymbol{U}^{\dagger}\boldsymbol{A}_{\mu}\boldsymbol{U}$ will be uncorrelated if we minimise $\mathrm{Tr}(\boldsymbol{\varepsilon}^2)$. Reference to the literature (e.g. see Stedman 1971) shows, however, that this is not normally done. It is more usual to minimise $\Sigma_i(\varepsilon_{ii})^2$ which leads to *correlated* parameters θ_{μ}. It is therefore proposed that this method be avoided in future.

If there are wide deviations in the uncertainty σ_i associated with different data it is usual to minimise $\Sigma_i(\varepsilon_{ii}/\sigma_i)^2$. The appropriate generalisation of our result would be to minimise $\mathrm{Tr}(\boldsymbol{\sigma}^{-1}\boldsymbol{\varepsilon}\boldsymbol{\sigma}^{-1}\boldsymbol{\varepsilon})$ where $\boldsymbol{\sigma}$ is the diagonal matrix of the σ_i. It would then be necessary to modify the basis orthogonality equation to read

$$\mathrm{Tr}(\boldsymbol{\sigma}^{-1}\boldsymbol{A}_{\mu}\boldsymbol{\sigma}^{-1}\boldsymbol{A}_{\nu}) = 0 \qquad \text{if } \mu \neq \nu.$$

In paramagnetic ion spectroscopy it is usual to associate parameters with the tensor operators $T_q^{(k)}$ with matrix elements (Judd 1963)

$$\langle L, -M|T_q^{(k)}|L, M'\rangle = (-1)^q(2k+1)^{1/2}\langle L\|T^{(k)}\|L\rangle\begin{pmatrix} L & k & L \\ M & q & M' \end{pmatrix}$$

$$= T(L\colon k, q)_{MM'}$$

where the first equality is given by the Wigner–Eckart theorem and the second defines the matrices $\boldsymbol{T}(L\colon k, q)$. An orthogonality relation for the $3j$ symbols $\left(\begin{smallmatrix} L & k & L \\ M & q & M' \end{smallmatrix}\right)$ given by

$$\sum_{M,M'}\begin{pmatrix} L & k_1 & L \\ M & q_1 & M' \end{pmatrix}\begin{pmatrix} L & k_2 & L \\ M' & q_2 & M' \end{pmatrix} = \frac{1}{2k_1+1}\delta(k_1, k_2)\delta(q_1q_2)$$

ensures that the set of matrices $\boldsymbol{T}(L\colon k, q)$ (for a given L) is mutually orthogonal. They can be normalised to unity by multiplying by $(2k_1+1)^{1/2}$.

This result shows that expressions of the type $\Sigma_{kq}T_q^{(k)}A_{kq}$, which are used in fitting crystal field parameters A_{kq}, give uncorrelated values of the A_{kq} provided that a complete $(2L+1)^2$ matrix is used in the fit. In practice this is not always the case as the experimental data may not be complete. A second requirement to obtain uncorrelated parameters is that the data should be uncorrelated and all have the same limits of uncertainty. In cases where the data does not have uniform uncertainty the matrices $\boldsymbol{\sigma}^{1/2}\boldsymbol{T}(L\colon k, q)\boldsymbol{\sigma}^{1/2}$ could be used in the fitting, and an appropriate transformation of the fitted parameters to the 'usual' A_{kq} be made subsequently. This would not, however, result in uncorrelated values of the A_{kq}.

The above discussion refers to the fitting of coefficients of the single-electron operators $T_q^{(k)}$ to experimental data as is carried out in crystal field theory (Stedman 1971). The theory of irreducible tensor operators shows that similar considerations apply to many-electron operators. For example, the matrices of the coupled operator $(T^{(k_1)}T^{(k_2)})_Q^{(K)}$ (Judd 1963) can easily be shown to be orthogonal. In particular, the Slater parameters, which are the coefficients in the case $K = Q = 0$ and $k_1 = k_2$ (even), are uncorrelated if the data satisfy the required conditions. This case is of special interest because it is possible to diagonalise the fitting expression for two-electron

matrices from symmetry considerations alone, independent of the values of the coefficients. Then the fitting procedure reduces to a simple linear least-squares fit and the off-diagonal terms in ε_{ij} are identically zero.

In summary, we have shown that the parameters conventionally fitted to spectroscopic data will be independent if the data span a fitting space in which the operators are orthogonal, the estimated errors of all data are equal, and $\mathrm{Tr}(\boldsymbol{\varepsilon}^2)$ is minimised rather than $\Sigma_i \varepsilon_{ii}^2$.

The author is grateful to Professor B R Judd and Mr W K Sze for focusing his mind on this problem and to Dr Y H Au-Yeung for clarifying the mathematical background.

References

Hoffman K and Kunze R 1971 *Linear Algebra* 2nd edn (Princeton, NJ: Prentice-Hall)

Hudson D J 1964 *Statistics Lectures II: Maximum Likelihood and Least Squares Theory* CERN 64-18 (CERN: Geneva)

Judd B R 1963 *Operator Techniques in Atomic Spectroscopy* (New York: McGraw-Hill)

Stedman G E 1971 *Comput. Phys. Commun.* **2** 191–200