CONF-800607--49

THE SIGNIFICANCE OF LAGRANGE MULTIPLIERS IN CROSS-SECTION ADJUSTMENT*

J. J. Wagschal† and Y. Yeivin†

Oak Ridge National Laboratory
Oak Ridge, Tennessee 37830

*Research sponsored by Reactor Research and Technology Division under Union Carbide Corporation contract W-7405-eng-26 with the U.S. Department of Energy.

†On Sabbatical leave from the Racah Institute of Physics, The Hebrew University, Jerusalem, Israel.

THE SIGNIFICANCE OF LAGRANGE MULTIPLIERS IN CROSS-SECTION ADJUSTMENT

J.J. Wagschal, Y. Yeivin (Hebrew Univ., Jerusalem and ORNL)

This note offers what we believe is the natural derivation of the explicit prescriptions incorporated in least-squares adjustment codes. We do not pretend to present any really new results, except perhaps for underlining the central role of the Lagrange multipliers, when adjustment of cross-sections by integral data is treated for what it truly is, namely a conditional-minimum problem. The evaluation of the Lagrange multipliers necessitates the inversion only of a "small" matrix, the order of which is the number of integral data by which the cross sections are adjusted. The complete solution of the adjustment problem, i.e. the adjusted differential and integral parameters and their respective uncertainty (variance-covariance) matrices, is then given in terms of the Lagrange multipliers by simple expressions, involving no additional matrix inversions.

Let us first review what adjustment is all about, which we mainly do in order to introduce a convenient notation. Consider an evaluated (energy-point or multigroup) cross-section library $\sigma\equiv(\sigma_i)$, $i=1,2,\ldots,N$, with the corresponding uncertainty matrix $C_\sigma\equiv\mathrm{cov}(\sigma_i,\sigma_j)\equiv E[(\delta\sigma)(\delta\sigma)^\dagger]$; and a set of experimental integral data (responses) $r\equiv(r_i)$, $i=1,2,\ldots,n$, with its uncertainty matrix $C_r\equiv\mathrm{cov}(r_i,r_j)$. Consider also the corresponding set of the calculated responses, $\bar{r}=\bar{r}(\sigma)$, and their sensitivities to the cross sections, i.e. the matrix of the $n\times N$ derivatives $S\equiv(\partial\bar{r}_i/\partial\sigma_j)$. Note that since $\bar{r}(\sigma+\delta\sigma)=\bar{r}+\delta\bar{r}=\bar{r}+S\delta\sigma$, the uncertainty matrix corresponding to the calculated responses is

$$C_{\bar{r}}=E[(\delta\bar{r})(\delta\bar{r})^\dagger]=E[(S\delta\sigma)(S\delta\sigma)^\dagger]=S\,E[(\delta\sigma)(\delta\sigma)^\dagger]S^\dagger=SC_\sigma S^\dagger, \quad (1)$$

which just reflects the propagation of the σ uncertainties through the calculation of $\bar{r}$. Denote the deviations of the calculated responses from their respective experimental values by $d\equiv(d_i)$: $d=\bar{r}-r$, and note that the deviations' uncertainty matrix is simply

$C_d = C_r + C_{\bar{r}}$. Let the (unknown) adjusted cross sections be σ', and the adjusted responses $r' = \bar{r}(\sigma')$. Then we further denote the actual adjustments by $x = \sigma' - \sigma$ and $y = r' - r$. We now assume that the approximation $r' \approx \bar{r} + S(\sigma' - \sigma)$ is valid, so that, in our notation,

$$y = d + Sx \ . \tag{2}$$

Then, the adjustment proposition is that the optimal adjustments are those values of x and of y which minimize the quadratic form

$$Q = x^{\dagger} C_{\sigma}^{-1} x + y^{\dagger} C_r^{-1} y \ , \tag{3}$$

subject to condition (2).

Now, a straightforward, seemingly natural course to solving this minimum problem is to eliminate y from Eq. (3), by means of Eq. (2), and then solve the resulting unconditional-minimum problem. In other words, the problem is reduced to finding the cross-section adjustments x which minimize

$$Q' = x^{\dagger}(C_{\sigma}^{-1} + S^{\dagger} C_r^{-1} S)x + 2x^{\dagger} S^{\dagger} C_r^{-1} d \ . \tag{4}$$

The solution of this minimum problem satisfies the N equations

$$\frac{\partial Q'}{\partial x} = 2[(C_{\sigma}^{-1} + S^{\dagger} C_r^{-1} S)x + S^{\dagger} C_r^{-1} d] = 0 \ , \tag{5}$$

so that

$$x = -(C_{\sigma}^{-1} + S^{\dagger} C_r^{-1} S)^{-1} S^{\dagger} C_r^{-1} d \ . \tag{6}$$

This expression is rather awkward, as it involves the inversion of a "small" (n×n) and two "large" (N×N) matrices. However, it is made to be shown that

$$(C_\sigma^{-1} + S^\dagger C_r^{-1} S)^{-1} S^\dagger C_r^{-1} = C_\sigma S^\dagger (C_r + S C_\sigma S^\dagger)^{-1} = C_\sigma S^\dagger C_d^{-1}, \tag{7}$$

where the expression on the right-hand side involves the inversion only of one small matrix, C_d. Thus, Eq.(6) actually reduces to the relatively simple expression

$$x = - C_\sigma S^\dagger C_d^{-1} d \; . \tag{8}$$

The response adjustments are now also given by a simple expression,

$$y = d + Sx = (I - S C_\sigma S^\dagger C_d^{-1}) d = [I + (C_r - C_d) C_d^{-1}] d = C_r C_d^{-1} d \; . \tag{9}$$

While the solution of the adjustment problem, as outlined above, is certainly legitimate, it is obviously not that "straightforward". At least from a methodical point of view, the more natural course is to treat adjustment as a conditional-minimum problem, which is what it really is, and to solve it by the method of Lagrange multipliers. To amplify this argument, we note that to evaluate both x and y, as given by Eqs.(8) and (9), we only have to evaluate $C_d^{-1}d$. The question then arises as to a more fundamental significance, if any, of this vector. It turns out that the components of this very vector are the Lagrange multipliers in the conditional-minimum formulation of the adjustment problem. Certainly, such insight could have not been gained without resort to this method of solution

In the conditional-minimum formulation the form

$$R(x,y) = Q(x,y) + 2\lambda^\dagger (Sx - y) \; , \tag{10}$$

where Q is the form of Eq.(3) and 2λ is the n-dimensional vector of Lagrange multipliers, is to be (unconditionally) minimized, i.e. x and y satisfy the equations

$$\frac{\partial R}{\partial x} = 2(C_\sigma^{-1} x + S^\dagger \lambda) = 0, \; \frac{\partial R}{\partial y} = 2(C_r^{-1} y - \lambda) = 0 \; . \tag{11}$$

Thus indeed

$$x = -C_{\sigma}S^{\dagger}\lambda \ , \quad y = C_{r}\lambda \ , \tag{12}$$

and, by Eq.(2),

$$d = y - Sx = (C_r + SC_{\sigma}S^{\dagger})\lambda = C_d\lambda \ , \quad \lambda = C_d^{-1}d \ . \tag{13}$$

Geometrically speaking, the last equation means that λ and d are the covariant and contravariant components, respectively, of one and the same vector with respect to the "metric" C_d. In the hypothetical situation of uncorrelated deviations with standard normal distributions (i.e. with $C_d = I$) we would indeed have $\lambda = d$. Incidentally, Eq.(13) also means that $C_{\lambda} = C_d^{-1}$.

Let us now derive the uncertainty matrices corresponding to the adjusted parameters. First, from Eq.(12) it immediately follows that

$$C_x = C_{\sigma}S^{\dagger}C_{\lambda}SC_{\sigma} \ , \quad C_y = C_r C_{\lambda} C_r \ . \tag{14}$$

Then, since for instance $r' = r + y$,

$$C_{r'} = C_r + C_y + 2E[(\delta y)(\delta r)^{\dagger}] \ . \tag{15}$$

However, as $\delta y = C_r \delta\lambda = C_r C_{\lambda}\delta(\overline{r} - r)$, and because σ (and therefore $\overline{r}$) and r are uncorrelated as implied by Eq.(3), we conclude that

$$E[(\delta y)(\delta r)^{\dagger}] = -C_r C_{\lambda} C_r = -C_y \ . \tag{16}$$

A similar derivation may, of course, be carried out for $C_{\sigma'}$, and the two final expressions are

$$C_{\sigma'} = C_{\sigma} - C_x \ , \quad C_{r'} = C_r - C_y \ . \tag{17}$$

To sum up, we have demonstrated (a) that the vector of Lagrange multipliers, λ, and its associated matrix, C_λ, are indeed the essence of the complete solution of the adjustment problem; and (b) that λ is just the geometrical counterpart of the vector of deviations **d**, with the latter's uncertainty matrix as "metric".