Journal of NUCLEAR SCIENCE and TECHNOLOGY, Supplement 2, p. 896–899 (August 2002)

A Note on How to Iterate Adjustment

Jehudah J WAGSCHAL* and Yehuda YEIVIN
Racah Institute of Physics, The Hebrew University of Jerusalem, Givat Ram Campus, IL-91904 Jerusalem, Israel

We study the elementary problem of adjusting a given parameter library by one measured response so that calculating the response with the adjusted library reproduces the measured value. We formulate the problem of the least-squares adjustment for non-linear responses, and offer a reasoned iteration scheme to solve it. A numerical example illustrates the success of the proposed procedure. Our scheme is identical to that suggested by Perey, and thus adds insight, offers justification, and also serves as a formal proof of the latter.

KEYWORDS: ***least-squares adjustment, the linearity assumption, iterated solution***

I. Background and Motivation

Within the adjustment-expert community, and as part of a rather extensive collection of unwritten laws (the so-called *oral law*), it has been generally known that, in applications for which the linearity assumption is valid, there is no point in re-adjusting the just adjusted parameters by the simultaneously adjusted responses. Starting from a given parameter library p and a given set of measured responses r, and adjusting the former by the latter, we achieve an adjusted library p' and the corresponding set of adjusted responses r'. If we then *carefully* apply the adjustment prescriptions to these adjusted quantities, we realize that the re-adjusted library p'' is identical to p', as indeed is r'' to r'. A formal demonstration of the above was presented at the latest Symposium on Reactor Dosimetry.[1)]

Without loss of generality, we may assume that the input p and r data are uncorrelated, i.e. that their cross-covariance matrix C_{rp} vanishes. Then

$$p' = p - C_p S^\dagger C_d^{-1} d \ , \quad r' = r + C_r C_d^{-1} d \ , \qquad (1)$$

where S is the sensitivity matrix (of the responses with respect to the parameters); the components of the vector d are the so-called residuals, which are in fact the *deviations* of the calculated responses from their corresponding measured ones: $d = \bar{r}(p) - r$; the subscripted C's are the (given) uncertainty matrices respectively associated with p and r and

$$C_d = C_{\bar{r}} + C_r = S C_p S^\dagger + C_r \qquad (2)$$

obviously is the uncertainty matrix associated with d.

It is, however, self-evident that p' and r' *are* correlated. Their cross-covariance matrix, since that of the given r and p vanishes, in fact is

$$C_{p'r'} = C_p S^\dagger C_d^{-1} C_r \ . \qquad (3)$$

An elementary, yet detailed, derivation of this equation, and indeed that of the entire conventional-adjustment *formulaire* is presented in Ref. 1. There it is demonstrated that using the prescriptions appropriate for the correlated input p' and r', the re-adjusted parameters and responses are identical to the respective input quantities, which proves our argument.

When responses are strictly linear in the parameters, then conventional adjustment is strictly correct. However, except for such cases, the validity of the linearity assumption, and practically the effectiveness of the conventional adjustment procedure, depend on how close the representative parameter-space point of the adjusted library, p', is to the original-library point p. And this may only be *a posteriori* determined from the actual results of the particular application. A common practice is to check whether $\bar{r}(p') \approx r'$. The true check however would seemingly be to examine to what extent $S' = S(p')$ is about equal to $S = S(p)$. In any event, the latter equality guarantees the former.

And so, the question naturally arises of what one could do when the linearity assumption fails, namely when the results of a particular application are such that the recalculated responses $\bar{r}(p')$ are different from the adjusted responses r'. In such circumstances it would seem quite natural to re-adjust the just adjusted parameters by the simultaneously adjusted responses. A common argument to justify such a procedure, and to repeat it, i.e. to iterate the adjustment procedure, is that the process eventually converges.

This iterated solution, however, misses the point in that it is not, in fact, the solution of the original adjustment problem. Recall that conventional (i.e. linearized) adjustment is supposed to find the library p' and the response set r' that, for uncorrelated p and r, minimize the quadratic form

$$Q(p', r') = (p' - p)^\dagger C_p^{-1} (p' - p) + (r' - r)^\dagger C_r^{-1} (r' - r) \ , \qquad (4)$$

subject to the constraint

$$S(p' - p) + d = r' - r \ . \qquad (5)$$

Well, this "natural" iterated solution *does not* generally minimize the form of Eq. (4). In fact, we propose another

* Corresponding author, Tel. +972-2-6584521, Fax. +972-2-6584461, E-mail: wagschal@vms.huji.ac.il

iterative procedure of which the converged solution indeed minimizes this form.

In the next section we analyze an elementary problem, in order to discuss the essence of our arguments, and make our point as clear and lucid as possible.

II. A Basic Problem: to Reproduce One Response

We now propose to consider a very elementary adjustment problem, in order to elaborate on the salient points of the general problem. We will then discuss the inadequacy of the iterative solution that we have described. And finally we will examine a "new" procedure that we consider to be a proper one for solving the adjustment problem when the linearity assumption is unjustified.

Let us start from the very beginning. The problem we consider is that of adjusting a given parameter library p by just one measured response r, so that the adjusted library p' reproduce the given response.

The measured response determines a manifold in the parameter space, the locus of all points (libraries) p' that reproduce the response: $\bar{r}(p')=r$. The point p representing the given library is obviously not on the manifold (since it was implicitly implied that $\bar{r}(p)\neq r$). Now, the fundamental proposition of least-squares adjustment is that the optimal p' is the manifold point that minimizes $|p'-p|^2$. It is the manifold point nearest to the given p. Such a configuration, for the two-parameter example discussed below, is depicted in Figure 1. The thick curve is the manifold determined by the measured response, and on which the point p' is the curve-point nearest to the given-parameter point p.

In any event, if the (calculated) response $\bar{r}(p)$ is linear in the parameters, or if the optimal manifold point p' is close enough to p, we would be justified in approximating

$$\bar{r}(p')\approx\bar{r}(p)+(\partial\bar{r}/\partial p)(p'-p) \quad , \tag{6}$$

where the second term on the right-hand side is the scalar product of the parameter-space gradient (evaluated at p), a *row* vector, by the column $p'-p$. Now, the calculated response at p', by definition, is r, so that Eq. (6) expresses a linear relation satisfied by the components of p'. Hence, letting p' vary under this constraint, there obtains a linear manifold that is the tangent hyper-plane to the manifold $\bar{r}(p')=r$, at its point nearest to p.

We denote the actual adjustments of the parameters by $x_n=p'_n-p_n$; and the sensitivity profile of our response by the *row* vector $s_n=\partial\bar{r}/\partial p_n$. Then Eq. (6) reduces to

$$sx+d=0 \quad . \tag{7}$$

The problem at hand is elementary: since x evidently is parallel to the vector $s^\dagger$, i.e. $x=\beta s^\dagger$, then obviously

$$sx+d=\beta ss^\dagger+d=0 \quad\Rightarrow\quad \beta=-\frac{d}{ss^\dagger} \quad\Rightarrow\quad x=-\frac{d}{ss^\dagger}s^\dagger \,. \tag{8}$$

III. A Numerical Example

We now illustrate the foregoing arguments by means of a manifestly non-linear example. Let us then consider a response that is a function of just two parameters, a and b:

$$\bar{r}(a,b)=ae^{-b} \quad . \tag{9}$$

Let the given parameter "library" be $(a,b)=(½,½)$, and the given measured value of the response $r=1$. The curve determined by the given response is $b=\ell n\,a$. The curve point nearest to p is the point (1,0), since the normal to the curve at the latter point passes through the former. Indeed,

$$\frac{\partial\bar{r}}{\partial a}=e^{-b} \;,\; \frac{\partial\bar{r}}{\partial b}=-ae^{-b} \quad . \tag{10}$$

Therefore the normal at (1,0) is the vector (1,–1), and this line is $b=1-a$, which obviously passes through (½,½). All this is shown in **Fig. 1**.

Thus, any worthy adjustment procedure should take us from the library (½,½) to one as close as possible to (1,0). We also note that $\chi^2 = min\ x^2 = ½$. Let us then see what happens if we apply the conventional procedure, i.e. the prescription of Eq. (8), to the problem at hand. First, $\bar{r}(p)=½\,e^{-½}=0.303265$. Then $s(p) = (1,-½)\,e^{-½}$. So that we easily find

$$p'=\begin{pmatrix}1.418976\\0.040512\end{pmatrix} \quad . \tag{11}$$

This, not surprisingly, is rather far from the desirable (1,0). Further, $\bar{r}(p')=1.362639$, manifestly different from the given r = 1. And, incidentally, $x^2 = 1.055646$.

We now repeat the procedure, and adjust p' by r = 1. Referring to Eq. (8) with $d\to d'=\bar{r}(p')-1=0.362639$, and $s\to s'=(0.960298,-1.362640)$, we come by

$$p''=\begin{pmatrix}1.293662\\0.218330\end{pmatrix} \quad . \tag{12}$$

Now, $\bar{r}(p'')=1.039923$, much closer to r = 1 than $\bar{r}(p')$, but the new x^2 is the still unsatisfactory 0.709237.

The third iteration, with $s''=(0.80386,-1.039923)$ and $d''=0.039923$, yields

$$p'''=\begin{pmatrix}1.275086\\0.242362\end{pmatrix} \quad . \tag{13}$$

We thus find that $\bar{r}(p''')=1.000652$, which practically reproduces r = 1, and the iteration process comes to its end. However, the current x^2, 0.667136, is only slightly lower than its former value. We therefore realize that p''' indeed lies on the curve determined by the given measured response, but away from the point nearest to the given p. The above three iterated values of the parameter pairs are shown in Fig. 1 along the dashed line marked "natural."

Thus the *seemingly* natural iteration procedure fails the criterion for acceptability. We will now consider another, more successful, scheme.

IV. An Alternative Iteration Scheme

We now examine a different iteration procedure, and demonstrate that it does converge to the parameter point (1,0), as a *deserving* adjustment procedure should. The primary idea of the proposed procedure derives from the elementary observation that the normal to the curve determined by the measured response ($b = \ell n\ a$ in our case), at the point p' nearest to the given-parameter point

p, passes through the latter point. In other words, the desired p' is the foot of the perpendicular to the curve dropped from p. Thus, if s' is the sensitivity profile, the gradient, of the (calculated) response, evaluated at p', then $x = p' - p = \beta(s')^{\dagger}$. The problem, of course, is to find what the value of β is.

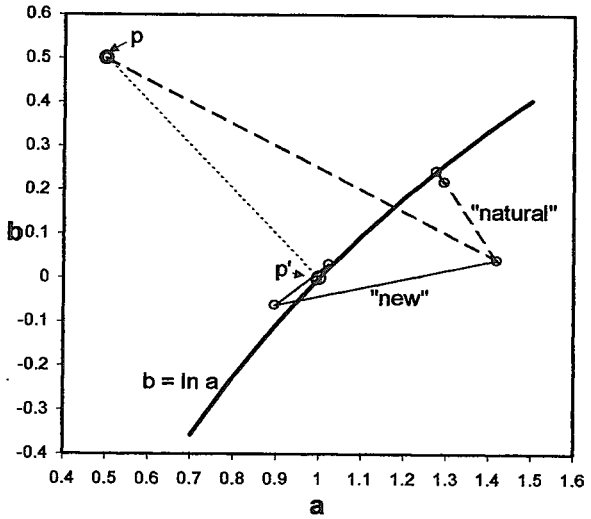

Fig. 1 The parameter (a,b) plane.
The given "library" p, the curve determined by the given measured response r ($b = \ell n\, a$), and the curve point p' nearest to p. The dashed line follows the seemingly natural iteration scheme, and the thin line follows the proposed new scheme.

To find a relation of the given data that involves β, we expand the response about p':

$$\bar{r}(p) = \bar{r}(p') + \left.\frac{\partial \bar{r}}{\partial p}\right|_{p'} (p - p') + \varepsilon \ , \qquad (14)$$

where ε represents the remainder in the Taylor expansion, which obviously does not vanish in our case of the non-linear response. We further recall that $\bar{r}(p) - \bar{r}(p') = \bar{r}(p) - r = d$. Thus the last equation may also be expressed as

$$s'x + d - \varepsilon = 0 \ . \qquad (15)$$

And as we substitute $x = \beta(s')^{\dagger}$, we find that

$$\beta = \frac{\varepsilon - d}{s'(s')^{\dagger}} \ . \qquad (16)$$

This is fine, except that we know neither s' nor ε. Of course, *if* the linearity assumption holds, then s' can be replaced by s, the sensitivity profile evaluated at p, ε vanishes, and Eq. (16), as expected, reduces to Eq. (8).

In order to proceed, we first refer to the first conventional iteration. We will denote its result, given in Eq. (11), by p_0. We will henceforth reserve the notation p' for our target point (1,0). In any event, we start by approximating $s' \approx s(p_0) = s_0$. Using this approximation, we then refer to Eq. (14) to obtain a reasonable guess for ε, or rather for $\varepsilon - d$:

$$\varepsilon - d \approx r - \bar{r}(p_0) + s_0(p_0 - p) \ . \qquad (17)$$

And so our next guess for the adjusted parameters, by Eqs. (16) and (17), is now

$$p_1 = p + \frac{r - \bar{r}(p_0) + s_0(p_0 - p)}{s_0 s_0^{\dagger}} s_0^{\dagger} \ . \qquad (18)$$

The numerical values of all the quantities are now known, and the result p_1 of the first step of the proposed procedure turns out to be

$$p_1 = \begin{pmatrix} 0.895978 \\ -0.061913 \end{pmatrix} \ . \qquad (19)$$

This point, and the successive iterations are shown in Fig. 1, connected by the line marked "new." Anyhow, $\bar{r}(p_1) = 0.953204$, which is about as close to r = 1 as $\bar{r}(p'')$ was. And then the corresponding x^2 is 0.472545. This is even lower than our target 0.5, but since p_1 is not on the curve determined by r, and with $\bar{r}(p_1) < 1$ lies above the curve, it only means that it is nearer to p than p' is. We note another positive indication of progress in the right direction. Ideally, we should reach a gradient parallel to $s' = (1,-1)$. The angle between our starting s and s' is 18.4°, that between s_0 and s' is 9.8°, and now we find that that the angle between s_1 and s' is already 3.14°.

To figure out the next "new" iteration, we refer to Eq. (18), in which we replace p_0 and s_0 by p_1 and s_1, so that:

$$p_2 = p + \frac{r - \bar{r}_1 + s_1(p_1 - p)}{s_1 s_1^{\dagger}} s_1^{\dagger} \ . \qquad (20)$$

Note that each iteration refers to the very given, original, parameters, using the preceding iteration just to improve the approximation of s' (by s_{n-1}) and of $\varepsilon - d$ (by means of p_{n-1}). Substituting the known numerical values, we obtain

$$p_2 = \begin{pmatrix} 1.023319 \\ 0.031119 \end{pmatrix} \ . \qquad (21)$$

Thus $\bar{r}(p_2) = 0.991965$, which almost reproduces r = 1. Further, it turns out that the angle between s_2 and s' is just 0.66°.

And, to be more pedantic, we also note that $p_3{}^{\dagger}$ = (0.994056, –0.005577), which is close enough to the ideal (1,0), and that $\bar{r}(p_3) = 0.999615$, which indeed is practically r = 1, which means that we have reached our goal, and our proposed iteration scheme does work.

V. Prescriptions for the General Case

The extension of Eq. (8) to the case of *several*, say *I*, measured responses to be reproduced by an adjusted parameter library was shown in Ref. 1 to be

$$x = -S^{\dagger}\left(SS^{\dagger}\right)^{-1} d \ , \qquad (22)$$

where S now is the sensitivity *matrix* of which the elements are $\partial \bar{r}_i / \partial p_n$, and d the deviations *vector*: $d_i = \bar{r}_i(p) - r_i$. The expansion of Eq. (14), obviously, is also valid in our present problem, and then, by complete

analogy we arrive at the extension of Eq. (20). In short, the proposed iteration scheme progresses to convergence via the successive iterations given by

$$p_{i+1} = p + S_i^{\dagger}\left(S_i S_i^{\dagger}\right)^{-1}\left[r - \bar{r}_i + S_i\left(p_i - p\right)\right] \quad . \qquad (23)$$

This indicates how to go about extending the proposed iteration scheme when dealing with real-life adjustment problems. By this we mean considering parameter and response uncertainties, and even their possible cross covariance matrices, and aiming at both the adjusted parameters *and* the adjusted responses.

For the sake of clarity and completeness, we introduce a few definitions before spelling out the explicit general linearized least-squares prescriptions. Consider the following partitioned generalized vectors and matrices, namely the adjustment vector z, sensitivity matrix S_z, and the "grand" uncertainty matrix C:

$$z \equiv \begin{pmatrix} p' - p \\ r' - r \end{pmatrix}, \quad S_z \equiv \left(S \ -1\right), \quad C \equiv \begin{pmatrix} C_p & C_{pr} \\ C_{rp} & C_r \end{pmatrix} \quad , \qquad (24)$$

where S is the conventional response-sensitivity matrix, and 1 the I×I unit matrix. Then the (apparently) "most" general (conventional) adjustment problem is to minimize $z^{\dagger}C^{-1}z$ subject to $S_z z + d = 0$. The solution of this conditional-minimum problem, also derived in Ref. 1, is

$$z = -CS_z^{\dagger}\left(S_z C S_z^{\dagger}\right)^{-1} d \quad . \qquad (25)$$

This may be somewhat simplified, owing to the (trivial!) observation $S_z C S_z^{\dagger} = C_d$. And further, block multiplying the first two (partitioned) matrices on the right-hand side of Eq. (25), we may separately express the parameter and response adjustments:

$$p' = p + \left(C_{pr} - C_p S^{\dagger}\right) C_d^{-1} d, \quad r' = r + \left(C_r - C_{rp} S^{\dagger}\right) C_d^{-1} d \quad (26)$$

Incidentally, the only difference between these prescriptions and the ones given in Eq. (1) is that here we also consider possible non-vanishing response-parameter cross covariances.

The called for extension of Eq. (23) is rather obvious. We propose that if the conventional adjustment, for whatever reason, is deemed unsatisfactory, then adjustment should be iterated by the recursive prescription

$$p_{i+1} = p + \left(C_{pr} - C_p S_i^{\dagger}\right) C_{di}^{-1}\left[\bar{r}_i + S_i\left(p - p_i\right) - r\right] \quad , \qquad (27)$$

where by C_{di} we mean that the sensitivity matrix which appears in the explicit expression of C_d is S_i, and where, as before, p_0 is the straightforward solution of conventional adjustment, i.e. p' of Eq. (26).

VI. Recapitulation and Discussion

We have demonstrated that the proposed adjustment iteration scheme, at least in the case of the elementary exercise described in sections III and IV, indeed converges to the pair of parameters that truly minimize the quadratic form it is supposed to. On the other hand, the scheme that we (and others) had considered to be the naturally called for scheme, turned out to converge to a "library" that, though reproducing the measured response, *does not* minimize the required form.

We should now discuss the iteration scheme suggested by FG Perey[2]. He states that "solving a non-linear least-squares problem we must always linearize the model ... [and this] involves an expansion, and the best expansion to make is about the solution." His idea might have been similar to ours. In any event, in terms of our elementary problem it was easy to explain why this expansion point indeed was the best choice. However, in modifying Eq. (8) Perey's statement, taken literally, seems to imply just replacing the sensitivity profile at the given p by the profile evaluated at the current library point, so that his successive-iteration scheme would seem to be

$$p_{i+1} = p + \frac{r - \bar{r}(p)}{s_i s_i^{\dagger}} s_i^{\dagger} \quad . \qquad (28)$$

Now, suppose that the iterated sensitivity profile converges to the ideal $s' = (1,-1)$, which means that the given p moves in the right direction toward the target p'. Even then, it should be obvious that, given the non-linearity of the response, a *linear* extrapolation cannot exactly move p to p'. Indeed, introducing s' into Eq. (28), and using the other numerical values that we have evaluated for calculating the result of Eq. (11), we find $p_{\infty}^{\dagger} = (0.848368, 0.151632)$, which obviously lies on the line $b = 1 - a$ connecting p and p', but just about a third of the way from p to p'.

Such a scheme, of course, is all wrong. But, by some stunning intuition, Perey also expanded $\bar{r}(p)$, in the numerator of Eq. (28), about the current set of parameters into $\bar{r}(p_i) + s_i(p - p_i)$. And thus at one stroke Eq. (28) reduces to Eq. (20). In other words, our iteration scheme is identical to that of Perey. And we might add that our analysis provides insight and justification for Perey's iteration scheme, and demonstrates why it is just right.

Acknowledgement

We are much obliged to John G Williams whose comments during the discussion following the oral presentation of Ref. 1, and further private exchanges, have in fact triggered the investigation reported here.

References

1) RL Perel, JJ Wagschal and Y Yeivin, "PV surveillance dosimetry and adjustment: review of several 'oral laws'," in *Reactor Dosimetry: Radiation Metrology and Assessment, ASTM STP 1398*, JG Williams *et al.*, Eds., American Society for Testing and Materials, West Conshohocken. PA. (2001). p. 486.

2) FG Perey, "Spectrum unfolding by the least-squares method," prepared for the IAEA Technical Committee Meeting on Current Status of Neutron Spectrum Unfolding, Oak-Ridge, October 1977. In *Contributions to few-channel spectrum unfolding*, FG Perey. ORNL/TM–6267 (1978). p. 21.