EPJ Web of Conferences **106**, 07003 (2016)
DOI: 10.1051/epjconf/201610607003
© Owned by the authors, published by EDP Sciences, 2016

Some Insight into the Generalized Linear Least Squares Parameter Adjustment Methodology

J.J. Wagschal

Racah Institute of Physics, Hebrew University of Jerusalem, Edmond J. Safra Campus, 91904 Jerusalem, Israel

Abstract. Some features of the generalized linear least squares parameter adjustment procedure have been discussed and proved. In particular: the equivalence of the adjusted measured response values and their recalculated values with the adjusted parameters, the proper way to iterate when the responses are not a linear function of the parameters and the equivalence of the simultaneous adjustment of the parameters with two not correlated measured responses and the consecutive adjustment first with one response and then with the second response.

1. Introduction

The generalized linear least squares adjustment methodology has been used for almost fifty years in nuclear data evaluation and testing, in reactor dosimetry as well as in criticality safety analysis [1–4]. Many novel users of the methodology either lack the knowledge of some of the properties of the methodology or try to reinvent some of its properties. In this paper we are going to elaborate on and document some of the less known properties of the methodology.

2. The Adjustment Formulas

The outcome of a generalized linear least squares adjustment can be presented by the following equations:

$$p' = p + \left(C_{pr} - C_p S^{\dagger}\right) C_d^{-1} d \tag{1}$$

$$r' = r + \left(C_r - C_{rp} S^{\dagger}\right) C_d^{-1} d \tag{2}$$

where p denotes a column vector of N parameters to be adjusted by a series of I measured responses presented by the column vector r. Any quantity with a prime symbol indicates the respective adjusted, i.e. posterior, quantity. The corresponding calculated response values using the parameters are given by the column vector $\bar{r}(p)$. Since usually the calculated response values differ from their respective measured values, we define their difference vector by $d = \bar{r}(p) - r$. The partial derivatives of the calculated responses with respect to the various parameters, the sensitivities, are presented by the I*N

This is an Open Access article distributed under the terms of the Creative Commons Attribution License 2.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article available at http://www.epj-conferences.org or http://dx.doi.org/10.1051/epjconf/201610607003

dimensional matrix S, and its transpose, N*I sensitivity matrix, is denoted by a dagger. The linearity is expressed by $\bar{r}\left(p'\right)-\bar{r}\left(p\right)=S\times\left(p'-p\right)$, which is equivalent to

$$\bar{r}\left(p'\right)-r=S\times\left(p'-p\right)+d\,. \tag{3}$$

The respective uncertainties are presented by the square variance-covariance matrices C_p, C_r and by the I*N response-parameter covariance matrix C_{rp}. The uncertainty in the difference vector, stemming from the respective uncertainties in the parameters and in the measured response, is given by

$$C_d=\left\langle\delta\left(\bar{r}-r\right)\delta\left(\bar{r}-r\right)^{\dagger}\right\rangle=SC_pS^{\dagger}-SC_{pr}-C_{rp}S^{\dagger}+C_r. \tag{4}$$

Usually the measured responses are not correlated with the parameters, however even if there is no prior response-parameter correlation the application of the generalized linear least squares adjustment results in such correlations. The posterior uncertainties in the responses, in the parameters and the resulting response-parameter correlations are given by:

$$C_{r'}=C_r-\left(C_r-C_{rp}S^{\dagger}\right)C_d^{-1}\left(C_r-SC_{pr}\right) \tag{5}$$

$$C_{p'}=C_p-\left(C_{pr}-C_pS^{\dagger}\right)C_d^{-1}\left(C_{rp}-SC_p\right) \tag{6}$$

$$C_{p'r'}=C_{pr}-\left(C_{pr}-C_pS^{\dagger}\right)C_d^{-1}\left(C_r-SC_{pr}\right). \tag{7}$$

Equations (1-7) were derived by minimizing the quadratic loss function $\left(r'-r,p'-p\right)^{\dagger}$ $C^{-1}\left(r'-r,p'-p\right)$, where the covariance matrix C is $\begin{pmatrix} C_r & C_{rp} \\ C_{pr} & C_p \end{pmatrix}$, subject to the linearity constraint, Eq. (3), using Lagrange multipliers, in [5, 6], and in a more elegant way in [7]. This method has important advantages over the alternative method of using the linearity, Eq. (3), explicitly in the loss function and minimizing it in a straightforward way. The two methods give identical results.

For the sake of simplicity, let us rewrite Eq. (1) for the case of vanishing response-parameter correlations. The parameters adjustment is $p'-p=-C_pS^{\dagger}C_d^{-1}d$ and the alternative derivation [5], without using Lagrange multipliers, results in $p'-p=-\left(C_p^{-1}+S^{\dagger}C_r^{-1}S\right)^{-1}S^{\dagger}C_r^{-1}d$. One can easily see that $C_pS^{\dagger}C_d^{-1}=\left(C_p^{-1}+S^{\dagger}C_r^{-1}S\right)^{-1}S^{\dagger}C_r^{-1}$, by multiplying both sides of the equation on the left by $\left(C_p^{-1}+S^{\dagger}C_r^{-1}S\right)$ and on the right by C_d which is $\left(C_r+S^{\dagger}C_pS\right)$. The straightforward minimization of the loss function necessitates the inversion of two large matrices and Eqs. ((1–2), (5–7)) require only the inversion of C_d, which has the dimension of the number of responses that is usually much smaller than the number of parameters.

2.1 Adjusted Responses and Re-calculated Parameters and Responses

The calculated response vector using the adjusted parameters and the linearity property is:

$$\bar{r}\left(p'\right)=\bar{r}\left(p\right)+S\left(p'-p\right)=\bar{r}\left(p\right)+S\left(C_{pr}-C_pS^{\dagger}\right)C_d^{-1}d \tag{8}$$

however, since $S\left(C_{pr}-C_pS^{\dagger}\right)=-C_d-C_{rp}S^{\dagger}+C_r$,

$$\begin{aligned}\bar{r}\left(p'\right)&=\bar{r}\left(p\right)+S\left(p'-p\right)=\bar{r}\left(p\right)+\left(-C_d-C_{rp}S^{\dagger}+C_r\right)C_d^{-1}d\\&=\bar{r}\left(p\right)-d+\left(-C_{rp}S^{\dagger}+C_r\right)C_d^{-1}d=r+\left(C_r-C_{rp}S^{\dagger}\right)C_d^{-1}d=r'.\end{aligned} \tag{9}$$

This means that, if the linearity assumption is correct, the adjusted response vector and the corresponding response vector recalculated with the adjusted parameters are just the same, thus, $\bar{r}\left(p'\right)=r'$. Although there is an uncertainty in $\bar{r}\left(p'\right)$ due to the propagated uncertainties in the adjusted parameters p' and there is an uncertainty in the adjusted response r', there is no uncertainty in their difference which is always equal to zero. However, if the calculated response is not a linear function of the parameters then the real recalculated value, using the adjusted parameters, is not equal to the adjusted response r'.

2.2 Adjustment Iterations

In practice, since the linearity is only an approximation, the explicitly recalculated response using the adjusted parameters may not coincide with the adjusted response value and thus d' does not vanish. It is tempting to try to re-adjust (adjust again) the new parameters p' and the new responses r' using their new respective uncertainties $C_{p'}$ and $C_{r'}$. According to the adjustment formulas the re-adjusted response vector r'' is: $r'' = r' + \left(C_{r'} - C_{r'p'}S^{\dagger}\right) C_{d'}^{-1} d'$. Inserting the adjusted uncertainty matrices into the coefficient of $C_{d'}^{-1} d'$ we get

$$\begin{aligned}
&C_{r'} - C_{r'p'}S^{\dagger}\\
&= \left[C_r - \left(C_r - C_{rp}S^{\dagger}\right) C_d^{-1} \left(C_r - SC_{pr}\right)\right] - \left[C_{rp} - \left(C_r - C_{rp}S^{\dagger}\right) C_d^{-1} \left(C_{rp} - SC_p\right)\right] S^{\dagger}\\
&= C_r - C_{rp}S^{\dagger} - \left(C_r - C_{rp}S^{\dagger}\right) C_d^{-1} \left(C_r - SC_{pr} - C_{rp}S^{\dagger} + SC_pS^{\dagger}\right)\\
&= C_r - C_{rp}S^{\dagger} - \left(C_r - C_{rp}S^{\dagger}\right) C_d^{-1} C_d = 0
\end{aligned} \tag{10}$$

which means that if we do not recalculate the sensitivities at p', which is time consuming, and use in the re-adjustment the original sensitivities, that have been calculated at p, the re-adjusted response values just do not change although d'does not vanish. Similarly, since $C_{p'r'} - C_{p'}S^{\dagger} = 0$ the re-adjusted parameters also do not change, and since these two coefficients are equal to zero the adjusted covariance matrices also do not change. The sensitivities have to be recalculated after each iteration.

2.3 Reduced Uncertainties

The common notion is that the need to adjust parameters, e.g. neutron cross sections, arises when there is a discrepancy between the calculated response values and their corresponding measured values, e.g. dosimeter reaction rates. However, it should be remembered that even when there is full agreement between calculated and measured response values, i.e. $d = 0$, and the parameters and response values do not change, the prior uncertainties in the parameters as well as in the measured response values are reduced. This reduction reflects the fact that there is no prior discrepancy.

Even if there is a prior discrepancy, one important result of an adjustment campaign is that the posterior response values and the posterior parameters have a finite reduced uncertainty. The posterior response uncertainty given by the covariance matrix, $C_{r'} = C_r - \left(C_r - C_{rp}S^{\dagger}\right) C_d^{-1} \left(C_r - SC_{pr}\right)$, is the difference between two covariance matrices. The first matrix is the prior covariance matrix of the measured responses and the second matrix is the covariance matrix of the response adjustment, $r' - r$. Since all covariance matrices are positive definite, the uncertainty is reduced. Similarly, the posterior uncertainty in the parameters, $C_{p'} = C_p - \left(C_{pr} - C_pS^{\dagger}\right) C_d^{-1} \left(C_{rp} - SC_p\right)$is also the difference of two covariance matrices, the second being the covariance matrix of the parameters adjustment, $p' - p$.

3. Consecutive Adjustments

A frequently asked question, particularly in the reactor physics community and in the cross section evaluators community, is whether an adjusted parameter library can be further modified by an adjustment with a new measured response or whether one has to go back to the original parameter library and perform a new adjustment campaign with all responses. In this section we will demonstrate this question by the simultaneous adjustment with two non correlated measured responses and then by the consecutive adjustment of the parameter library by the two responses.

The 2*N sensitivity matrix, S, has two rows. The first row, S_1, the sensitivity of the first response to all the parameters and the second row, S_2, the sensitivity of the second response to all the parameters. All notations will have an index denoting their relationship to response number one or response number two.

3.1 Simultaneous Adjustment

The joint uncertainty matrix of the deviations of the two calculated response values from their corresponding uncorrelated measured response values is given by

$$C_d = \begin{pmatrix} C_{d_1} & S_1 C_p S_2^{\dagger} \\ S_2 C_p S_1^{\dagger} & C_{d_2} \end{pmatrix} \tag{11}$$

and its inverse is given by

$$\begin{aligned} C_d^{-1} &= \frac{1}{C_{d_1} C_{d_2} - \left(S_1 C_p S_2^{\dagger}\right)\left(S_2 C_p S_1^{\dagger}\right)} \begin{pmatrix} C_{d_2} & -S_1 C_p S_2^{\dagger} \\ -S_2 C_p S_1^{\dagger} & C_{d_1} \end{pmatrix} \\ &= \frac{1}{\det C_d} \begin{pmatrix} C_{d_2} & -S_1 C_p S_2^{\dagger} \\ -S_2 C_p S_1^{\dagger} & C_{d_1} \end{pmatrix}. \end{aligned} \tag{12}$$

The simultaneous adjustment in the response values is given by $r' - r = C_r C_d^{-1} d$.

In particular, the adjustment of the first response is given by

$$r_1' - r_1 = \frac{C_{r_1}}{\det C_d} \left[C_{d_2} d_1 - \left(S_1 C_p S_2^{\dagger}\right) d_2 \right] \tag{13}$$

and the adjustment of the second response is

$$r_2' - r_2 = \frac{C_{r_2}}{\det C_d} \left[-\left(S_2 C_p S_1^{\dagger}\right) d_1 + C_{d_1} d_2 \right]. \tag{14}$$

The simultaneous adjustment in the parameter values is given by

$$\begin{aligned} p' - p &= -C_p S^{\dagger} C_d^{-1} d \\ &= \frac{-C_p S^{\dagger}}{\det C_d} \begin{pmatrix} C_{d_2} & -S_1 C_p S_2^{\dagger} \\ -S_2 C_p S_1^{\dagger} & C_{d_1} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \frac{-C_p S^{\dagger}}{\det C_d} \begin{pmatrix} C_{d_2} d_1 - \left(S_1 C_p S_2^{\dagger}\right) d_2 \\ -\left(S_2 C_p S_1^{\dagger}\right) d_1 + C_{d_1} d_2 \end{pmatrix}. \end{aligned} \tag{15}$$

All adjusted quantities, the parameters as well as the two previously non correlated measured responses carry now information from each of the adjusted responses.

3.2 Adjustment by the First Response

The adjustment of the parameters by the first response results in the adjustment of the first response by $r_1' - r_1 = C_{r_1} C_{d_1}^{-1} d_1$ and of the parameters by $p_1' - p = -C_p S_1^{\dagger} C_{d_1}^{-1} d_1$. The symbol p_1' indicates that these are adjusted parameters (the prime indicating adjustment) and that the adjustment was performed with the first response (sub 1). The uncertainty in these parameters which have to be used in the further adjustment by response number two is $C_{p_1'} = C_p - C_p S_1^{\dagger} C_{d_1}^{-1} S_1 C_p$.

All adjusted items are the result of the applications of the adjustment formulas given in Sect. 2.

3.3 Consecutive Adjustment by the Second Response

We proceed with the adjustment of the parameters with response number two. Response number two is now calculated not with the original parameters p but rather with the adjusted parameters that have already been adjusted by response number one. The resulting deviation of its calculated value from its measured value is $\bar{r}_2\left(p_1'\right) - r_2$ and the inverse of the uncertainty of this deviation, needed for the next adjustment is

$$C_{d_2}^{-1}\left(p_1'\right) = \left[C_{r_2} + S_2 C_{p_1'} S_2^{\dagger}\right]^{-1}. \tag{16}$$

The adjustment in response two by response two, calculated with parameters that have been previously adjusted, is

$$r_2'\left(p_1'\right) - r_2 = C_{r_2} C_{d_2}^{-1}\left(p_1'\right)\left[\bar{r}_2\left(p_1'\right) - r_2\right]. \tag{17}$$

Note that the deviation of the calculated response from its measured value, used in the second adjustment is not d_2, since the response was calculated with the adjusted parameters p_1' and not with the original parameters p. We use the linearity in calculating the shift of the response value resulting from the shift of the parameters due to the first adjustment. This deviation, can be reformatted as follows

$$\begin{aligned} \bar{r}_2\left(p_1'\right) - r_2 &= \left[\bar{r}_2\left(p_1'\right) - \bar{r}_2(p)\right] + \left[\bar{r}_2(p) - r_2\right] \\ &= S_2\left(p_1' - p\right) + d_2 = S_2\left(-C_p S_1^{\dagger} C_{d_1}^{-1} d_1\right) + d_2 \\ &= -S_2 C_p S_1^{\dagger} C_{d_1}^{-1} d_1 + d_2. \end{aligned} \tag{18}$$

The adjustment in response two is

$$\begin{aligned} r_2'\left(p_1'\right) - r_2 &= C_{r_2} C_{d_2}^{-1}\left(p_1'\right)\left[\bar{r}_2\left(p_1'\right) - r_2\right] \\ &= C_{r_2} C_{d_2}^{-1}\left(p_1'\right)\left(-S_2 C_p S_1^{\dagger} C_{d_1}^{-1} d_1 + d_2\right) \\ &= C_{r_2}\left[C_{r_2} + S_2\left(C_p - C_p S_1^{\dagger} C_{d_1}^{-1} S_1 C_p\right) S_2^{\dagger}\right]^{-1}\left(-S_2 C_p S_1^{\dagger} C_{d_1}^{-1} d_1 + d_2\right) \\ &= \frac{C_{r_2}\left(-S_2 C_p S_1^{\dagger} C_{d_1}^{-1} d_1 + d_2\right)}{C_{r_2} + S_2\left(C_p - C_p S_1^{\dagger} C_{d_1}^{-1} S_1 C_p\right) S_2^{\dagger}} = \frac{-C_{r_2}\left[\frac{\left(S_2 C_p S_1^{\dagger}\right)}{C_{d_1}} d_1 - d_2\right]}{\frac{C_{d_2} C_{d_1} - S_2 C_p S_1^{\dagger} S_1 C_p S_2^{\dagger}}{C_{d_1}}} \\ &= C_{r_2}\frac{-\left(S_2 C_p S_1^{\dagger}\right) d_1 + C_{d_1} d_2}{C_{d_2} C_{d_1} - S_2 C_p S_1^{\dagger} S_1 C_p S_2^{\dagger}}. \end{aligned} \tag{19}$$

The denominator can be recognized as $\det C_d$. Thus, the total adjustment of response two is the same as its simultaneous adjustment with response one.

3.4 The Adjusted Parameters

The original parameters have been adjusted twice, first by response one and then the adjusted parameters have been adjusted once more by response two. Let us denote these parameters as $p_{1,2}''$, the double prime indicating two adjustments and 1,2 indicating that the adjustments have been first by response one and then by response two.

$$p''_{1,2} - p = p''_{1,2} - p'_1 + \left(p'_1 - p\right) = \left(p''_{1,2} - p'{}_1\right) - C_p S_1^\dagger C_{d_1}^{-1} d_1. \tag{20}$$

The total parameter adjustment is the sum of the first adjustment by response one, starting with the original parameters p, and presented in 3.2, followed by the second adjustment by response two starting with p'_1, derived in Sect. 3.3. The second adjustment is:

$$p''_{1,2} - p'_1 = -C_{p'_1} S_2^\dagger \left\{C_{d_2}^{-1}\left(p'_1\right)\left[\bar{r}_2\left(p'_1\right) - r_2\right]\right\}. \tag{21}$$

Plugging in the various items that have been derived in earlier sections and manipulating the expression in a similar way we end up with a compact expression for $p''_{1,2} - p$

$$\begin{aligned} p''_{1,2} - p &= \frac{C_p S_2^\dagger\left(S_2 C_p S_1^\dagger\right) - C_p S_1^\dagger C_{d_2} C_{d_1} C_{d_1}^{-1}}{\det C} d_1 + \frac{C_p S_1^\dagger\left(S_1 C_p S_2^\dagger\right) - C_p S_2^\dagger C_{d_1}}{\det C} d_2 \\ &= \frac{-C_p S^\dagger}{\det C}\begin{pmatrix} C_{d_2} d_1 - \left(S_1 C_p S_2^\dagger\right) d_2 \\ -\left(S_2 C_p S_1^\dagger\right) d_1 + C_{d_1} d_2 \end{pmatrix}. \end{aligned} \tag{22}$$

The consecutively adjusted parameters are equal to the adjusted parameters resulting from a simultaneous adjustment of the parameters and the two measured responses, as derived in 3.1.

3.5 The Measured Response one After the Second Adjustment

Response one has been adjusted by itself and is equal to its calculated value with the corresponding adjusted parameters. The following adjustment of these parameters, by the second response, necessitates the calculation of the value of response one once again, this time by the twice adjusted parameters. The final calculated value of response one is the sum of its former adjusted value by itself, and using the linearity, the shift of the calculated response due to the shift of the parameters from p'_1 to $p''_{1,2}$.

$$\begin{aligned} \bar{r}_1\left(p''_{1,2}\right) - r_1 &= \left[\bar{r}\left(p''_{1,2}\right) - \bar{r}_1\left(p'_1\right)\right] + \left[\bar{r}_1\left(p'_1\right) - r_1\right] \\ S_1 &= \left(p''_{1,2} - p'_1\right) + C_{r_1} C_{d_1}^{-1} d_1. \end{aligned} \tag{23}$$

After inserting the parameter shift into the last equation and manipulating the terms once again we get

$$r'_1 - r_1 = \frac{C_{r_1}}{\det C_d}\left[C_{d_2} d_1 - \left(S_1 C_p S_2^\dagger\right) d_2\right] \tag{24}$$

which is equal to the result of the simultaneous adjustment.

We have shown that the simultaneous adjustment of a set of parameters by two uncorrelated responses is equivalent to adjusting the parameters by one of the responses and modifying these adjusted parameters by a second adjustment by the second measured response.

4. Summary

After briefly introducing the motivation for writing this paper and the introduction of the generalized linear least squares adjustment symbols and equations we pointed out a few less known and less documented properties of the methodology. We showed that the equality of adjusted responses and their corresponding calculated values with the adjusted parameters is a direct consequence of the linearity assumption in the linear least squares adjustment methodology. We have shown that unless one recalculates the sensitivities of the responses to the parameters for the adjusted parameters, adjustment iterations are superfluous. We pointed out why adjustment reduces the uncertainties of both responses and parameters and finally compared the adjustment results of an adjustment utilizing simultaneously two uncorrelated measured responses to the sequential adjustment of the parameters and the responses showing that they are identical.

References

[1] G. Cecchini, U. Farinelli, A. Gandini and M. Salvatores, "Analysis of integral data for few-group parameter evaluation of fast reactors", *Proc. 3rd Intern. Conf. Peaceful Uses Atomic Energy* **2**, 388–397 (1964)

[2] M. Humi, J.J. Wagschal and Y. Yeivin, "Multi-Group Constants From Integral Data," *Proc. 3rd Intern. Conf. Peaceful Uses Atomic Energy* **2**, 398–402 (1964)

[3] J.J. Wagschal, R.E. Maerker and Y. Yeivin, "Extrapolation of Surveillance Dosimetry Information to Predict Pressure Vessel Fluences," Trans. Am. Nucl. Soc. **34**, 631–632 (1980)

[4] B. L. Broadhead, B. T. Rearden, C. M. Hopper, J. J. Wagschal and C. V. Parks, "Sensitivity – and Uncertainty – Based Criticality Safety Validation Techniques," Nucl. Sci. Eng. **146**, 340–366 (2004)

[5] J.J. Wagschal and Y. Yeivin, "The Significance of Lagrange Multipliers in Cross-Section Adjustment," Trans. Am. Nucl. Soc. **34**, 776–777 (1980), also ESIS Newsletter **33**, 3 (1980)

[6] R.E. Maerker, B.L. Broadhead and J.J. Wagschal, "Theory of a New Unfolding Procedure in Pressurized Water Reactor Pressure Vessel Dosimetry and Development of an Associated Benchmark Data Base," Nucl. Sci. Eng. **91**, 369–392 (1985)

[7] D. Shalitin, J.J. Wagschal and Y. Yeivin, "A Contribution to the Theory of Reactor Pressure Vessel Dosimetry," *Reactor Dosimetry: Methods, Applications, and Standardization*," ASTM STP-**1001**, Harry Farrar, IV and E.P. Lippincott eds., American Society for Testing and Materials, Philadelphia, 399–404 (1989)