

Criteria of the validation of experimental and evaluated covariance data

S. Badikov^a

Central Research Institute of Management, Economics and Information (Atominform), 127434 Moscow, Russia

Abstract. Criteria of the validation of experimental and evaluated covariance data are reviewed. In particular: a) criterion of the positive definiteness for covariance matrices, b) relationship between the “integral” experimental and estimated uncertainties, c) validity of the statistical invariants, d) restrictions imposed to correlations between experimental errors, are described. The examples of applying these criteria in practice are presented.

1 Introduction

The experimental and evaluated covariance data is of special importance for following reasons: 1) covariance matrices enter the equations of main methods of the theory of estimates (maximum likelihood method, generalized least squares method, Bayes method); 2) adequate covariance data is a key information for a construction of the distributions of the random vectors (in case of normal distribution the covariances determine the distribution completely); 3) in some studies (analysis of time serieses in economics, meteorology, etc.) the empirical covariance matrices are of direct interest since some characteristics of the matrices (for example, eigenvalues) admit physical interpretation.

This paper is devoted to a description of the criteria of the validation of experimental and evaluated covariance data which can conventionally be divided in 2 groups: 1) criteria designed for the refinement of calculated experimental and estimated covariance matrices from unphysical structures: a) checking the covariance matrices for positive definiteness, b) analysis of the correspondence between shapes of uncertainties and model function), 2) criteria for checking the results of statistical analysis: a) comparison of experimental and estimated integral uncertainties, b) verification of the statistical invariants.

2 Checking the covariance matrices for positive definiteness

The variances of physical values must be real and positive. The requirement of positive definiteness for covariance and correlation matrices is an extension of this restriction to multivariate objects – random vectors. This requirement can be justified as follows. Physically a correlation between any two random values (let's call them “secondary” ones) is induced by contribution of third random value (let's call it “primary” one) to each of the “secondary” values. The covariance matrix \mathbf{W} of the “secondary” values can be reduced to the covariance matrix \mathbf{U} of the “primary” values by an

orthogonal transformation \mathbf{F} : $\mathbf{U} = \mathbf{F}^T \mathbf{W} \mathbf{F}$. The variances of the “primary” values will be positive only in case of positive definiteness of the covariance matrix \mathbf{W} of the “secondary” values.

2.1 Transformation of random vectors and positive definiteness of the covariance matrices

An arbitrary transformation of a random vector $\vec{\theta}$ (dimension L) with a covariance matrix \mathbf{W} is designated as

$$\vec{\mu} = \vec{F}(\vec{\theta}) = (F_1(\vec{\theta}), \dots, F_M(\vec{\theta}))^T. \quad (1)$$

Using the error propagation law the covariance matrix \mathbf{U} of the vector $\vec{\mu}$ can be expressed through covariance matrix \mathbf{W}

$$\mathbf{U} = \mathbf{F}^T \mathbf{W} \mathbf{F}, \quad (2)$$

where $\mathbf{F}_{ki} = \frac{\partial F_i(\vec{\theta})}{\partial \theta_k}$ – sensitivity coefficient. The following question is of interest. Is the matrix \mathbf{U} positive definite in case of positive definiteness of the covariance matrix \mathbf{W} ? Let's consider a quadratic form

$$S = \vec{z}^T \mathbf{U} \vec{z} \quad (3)$$

of M variables z_1, \dots, z_M . In accordance with definition the matrix \mathbf{U} is a positive (semi)-definite if for any non-trivial vector \vec{z} the quadratic form S is positive (non-negative). After substitution (2) in (3) and designating $\vec{x} = \mathbf{F} \vec{z}$ the quadratic form S is transformed to

$$S = \vec{z}^T \mathbf{F}^T \mathbf{W} \mathbf{F} \vec{z} = \vec{x}^T \mathbf{W} \vec{x}. \quad (4)$$

Since \mathbf{W} is a positive definite by assumption, the quadratic form S is positive for any non-trivial vector \vec{x} . Further, we will find the vectors \vec{z} which give zero vectors \vec{x} after multiplication by \mathbf{F} . If $L < M$ rank $\mathbf{F} < M$, and a system algebraic linear equations $\mathbf{F} \vec{z} = 0$ has endless number of non-trivial solutions. Thus, positive definiteness of covariance matrix \mathbf{W} leads to positive semi-definite matrix \mathbf{U} at $L < M$. If $L \geq M$ the number of solutions of the system $\mathbf{F} \vec{z} = 0$ depends on value of rank \mathbf{F} : at rank $\mathbf{F} = M$ the system has a trivial solution $\vec{z} = 0$ while

^a e-mail: badikov@ainf.ru

at rank $\mathbf{F} < M$ it has endless non-trivial solutions. So, at $L \geq M$ the positive definiteness of covariance matrix \mathbf{W} results in a) the positive semi-definite matrix \mathbf{U} if rank $\mathbf{F} < M$, b) the positive definite matrix \mathbf{U} if rank $\mathbf{F} = M$. These conclusions can be schematically presented in a such way:

$$W > 0 \Rightarrow \begin{cases} U \geq 0 & L \geq M & (\text{rank } F < M) \text{ or } L < M \\ & \text{at} & \\ U > 0 & L \geq M & (\text{rank } F = M) \end{cases} \quad (5)$$

where $\mathbf{W} > 0$ means a positive definiteness of matrix \mathbf{W} .

2.2 Examples

Few transformations of a random vector are widely used in neutron and reactor physics. Preserving the positive definiteness of the covariance matrices in these transformations can be studied on the basis of analysis given in section 2.1. Most important examples are given below.

- a) **Estimation of unknown vector of parameters and their covariances** within the framework of generalized least squares method on the basis of correlated measurements. Main expressions of generalized least squares method in assumption of linear regression model $f(E, \vec{\theta})$ are as follows:

$$\hat{\vec{\theta}} = [\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{V}^{-1} \vec{y} \quad (6)$$

$$\mathbf{W} = [\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}]^{-1}. \quad (7)$$

In (6) and (7) \vec{y} and $\hat{\vec{\theta}}$ -vectors of N measurements and L estimated parameters ($N > L$), \mathbf{V} and \mathbf{W} – covariance matrices of measurements and estimated parameters correspondingly, \mathbf{X} – a matrix of sensitivity coefficients.

Let's find a constraint on covariance matrix \mathbf{V} of measurements which guarantees positive (semi-) definiteness of the matrix \mathbf{W} of estimated parameters. A statement **I** (its proof is straightforward) will be used in this consideration: if symmetrical matrix is positive definite then its inverse matrix is also positive definite.

If the matrix \mathbf{V} is positive definite then the matrix \mathbf{V}^{-1} is also positive definite in accordance with statement **I**. Then as follows from (5) an informational matrix $\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}$ is a positive definite if rank $\mathbf{X} = L$. Finally, \mathbf{W} is a positive definite as the matrix inverse to informational one $\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}$ according to statement **I**.

- b) Usually in reactor calculations the values of estimated function $f(E, \hat{\vec{\theta}})$ averaged over energy groups are used:

$$\bar{f}_i = \int_{E_i}^{E_{i+1}} f(E, \hat{\vec{\theta}}) dE, \quad i = 1, \dots, M, \quad (8)$$

M – number of energy groups. Covariance matrix \mathbf{U} of averaged values \bar{f}_i is presented by formula (2) where

\mathbf{W} – covariance matrix of estimated parameters $\hat{\vec{\theta}}$, $\mathbf{F}_{ki} = \int_{E_i}^{E_{i+1}} \frac{\partial f(E, \hat{\vec{\theta}})}{\partial \theta_k} dE$ – sensitivity coefficient of the estimated function in the i -th group relative to the k -th parameter.

As a rule the number of groups M is larger the number of parameters L . In this case covariance matrix \mathbf{U} is a positive semi-definite in accordance with (5).

- c) **Splitting the groups.** This procedure is widely used in reactor calculations. Let f_i and f'_i – values of function $f(E)$ within L and M ($M > L$) groups: $[E_i, E_{i+1}]$, $i = 1, \dots, L$ and $[E'_j, E'_{j+1}]$, $j = 1, \dots, M$ before and after splitting. Each initial group $[E_i, E_{i+1}]$ includes n_i new groups $[E'_j, E'_{j+1}]$. In formalized form splitting can be presented as

$$\vec{f}' = \mathbf{F}^T \vec{f} \quad (9)$$

where

$$\mathbf{F}_{ij} = \begin{cases} 1 & k_1 \leq j \leq k_2 & k_1 = \sum_{k=1}^{i-1} n_k + 1 & k_2 = \sum_{k=1}^i n_k \\ 0 & \text{other } j \end{cases} \quad (10)$$

Covariance matrix \mathbf{U} of values f'_i is given by formula (2) if \mathbf{W} – covariance matrix of values f_i . According to (5) \mathbf{U} – positive semi-definite matrix since $M < L$.

- d) **Collapsing the groups.** Let f_i and f'_i – values of function $f(E)$ within L and M ($M < L$) groups: $[E_i, E_{i+1}]$, $i = 1, \dots, L$ and $[E'_j, E'_{j+1}]$, $j = 1, \dots, M$ before and after collapsing. Each new group $[E'_j, E'_{j+1}]$ includes n_j initial groups $[E_i, E_{i+1}]$. In formalized form collapsing can be presented by formula (9) where

$$\mathbf{F}_{ij} = \begin{cases} (E_{i+1} - E_i) / \sum_{k=k_1}^{k_2} (E_{k+1} - E_k) & k_1 \leq i \leq k_2 \\ 0 & \text{other } i \end{cases}$$

$$k_1 = \sum_{k=1}^{j-1} n_k + 1 \quad k_2 = \sum_{k=1}^j n_k.$$

Covariance matrix \mathbf{U} of values f'_i is given by formula (2) if \mathbf{W} – covariance matrix of values f_i . Evidently that rank $\mathbf{F} = M$. Then the matrix \mathbf{U} is a positive definite in accordance with (5).

3 Comparison of the “integral” experimental and estimated uncertainty

Conception of “integral” uncertainty is of special importance for a validation of the results of evaluation carried out with statistical methods. The results are usually presented as a vector of estimated values and corresponding covariance matrix. The covariances of experimental and estimated errors – quadratic tables of numbers – can only be compared on the basis of integral measure of uncertainty. Exceeding estimated “integral” uncertainty over experimental one indicates error in statistical processing. In addition, covariance data from different evaluations can also be compared with using the “integral” uncertainty – of few evaluations the one with most small “integral” uncertainty (under other conditions being equal) must be recommended for use in applications. Of all the integral measures of uncertainty for a random vector only the

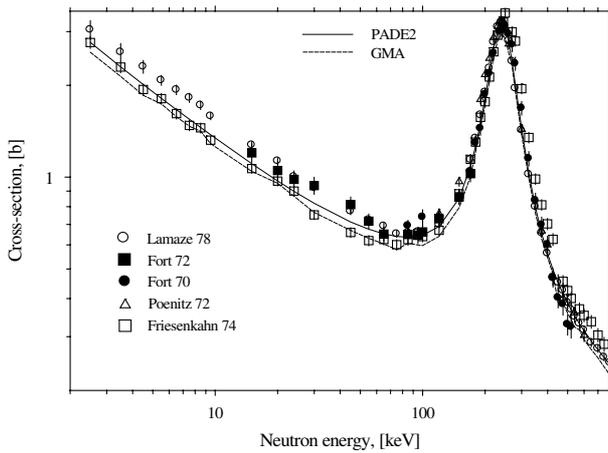


Fig. 1. Results of fitting the ${}^6\text{Li}(n,t)$ reaction cross section measurements by rational (PADE2) and stepwise (GMA) model functions.

Table 1. Limits of energy intervals used for averaging the experimental and estimated covariance information for the ${}^6\text{Li}(n,t)$ reaction cross section.

N	limits, keV	N	limits, keV	N	limits, keV
1	2.5–8.5	4	100–200	7	300–425
2	8.5–45	5	200–245	8	425–570
3	45–100	6	245–300	9	570–800

square root of determinant of a covariance matrix was applied in this study.

For illustration the results of statistical processing 5 sets of the ${}^6\text{Li}(n,t)$ reaction cross section measurements (*Lamaze 78* [1], *Fort 72* [2], *Fort 70* [3], *Poenitz 72* [4], *Friesenkahn 74* [5]) are presented (fig. 1).

Rational and stepwise model functions were used for fitting the experimental data with PADE2 and GMA code [6]. Output data for the GMA code – cross sections estimated at a set of neutron energies and corresponding covariances while for PADE2 code – estimated resonance parameters and their covariances. The experimental data and output data for both the codes were reduced to comparable values – covariances for average (over energy intervals) cross sections. Two runs of calculations (for 2 and 5 experimental data sets) were carried out. Cross sections and covariances estimated in GMA and PADE2 calculations (see fig. 1) were folded in 9 energy intervals with limits given in table 1.

The experimental information (cross sections and covariances) was also averaged. Thus, in each run 3 covariance matrices (corresponding to the experimental data, GMA and PADE2 results) were calculated. The characteristics of the matrices – square roots of the determinant ($\sqrt{\det C}$) and trace of matrix ($\sqrt{\text{tr} C}$) – are given in tables 2, 3. Average pointwise experimental and estimated cross section uncertainties (δ) are also presented in tables 2, 3.

As seen from tables 2, 3 two effects are observable. First one is expectable. Both integral and differential uncertainties for estimated cross sections are smaller the corresponding values for measured cross sections. The reduction increases with including additional experimental data in statistical processing. The second effect is non-evident. Model and

Table 2. “Integral” and average pointwise uncertainties of experimental and estimated ${}^6\text{Li}(n,t)$ reaction cross sections in case of statistical processing 2 sets of measurements [1,2].

uncertainty	experiment	GMA	PADE2	GMA/PADE2
$\sqrt{\det C}, b^9$	3.077–18	2.418–18	1.906–18	1.27
$\sqrt{\text{tr} C}, b$	8.791–02	8.465–02	3.692–02	2.29
$\delta, \%$	3.58	2.93	1.80	1.63

Table 3. “Integral” and average pointwise uncertainties of experimental and estimated ${}^6\text{Li}(n,t)$ reaction cross sections in case of statistical processing 5 sets of measurements [1–5].

uncertainty	experiment	GMA	PADE2	GMA/PADE2
$\sqrt{\det C}, b^9$	4.775–19	1.253–19	1.098–19	1.14
$\sqrt{\text{tr} C}, b$	6.542–02	5.138–02	2.305–02	2.23
$\delta, \%$	4.04	2.12	1.17	1.81

“non-model” fits result in approximately the same integral uncertainty for the vectors of average estimated cross sections while differential (pointwise) uncertainties for cross sections differ essentially. Close values for the integral uncertainties can be explained by a variation in correlation coefficients. The variation in correlations of the covariance matrix allows to preserve a value of the determinant after decreasing uncertainties. It should be noted that at smallest dimension ($n=1$) the procedure described above reduces to a comparison of the uncertainties of average (over energy range of interest) experimental and evaluated cross sections.

4 Verifying the statistical invariants

For controlling results of the large-scale statistical calculations (tens and hundreds of parameters, hundreds and thousands of measurements) preserving values – so called statistical invariants – are of importance. The invariants are preserved in statistical processing independently on 1) type or dimension of a model, 2) type of error (experimental or estimated). Let’s consider N measurements of an arbitrary model function $f(x, \vec{\theta})$, M – dimension of the vector $\vec{\theta}$ to be estimated, \mathbf{V} – covariance matrix of unbiased errors of the measurements, \mathbf{X} – matrix of sensitivity coefficients, \mathbf{W} – covariance matrix of estimated parameters:

$$\mathbf{W}_{\alpha\beta} = \left(\sum_{k,l} X_{\alpha k}^T V_{kl}^{-1} X_{l\beta} \right)^{-1}, \quad (11)$$

\mathbf{U} – covariance matrix of errors of estimated function

$$U_{ij} = \sum_{\alpha,\beta} X_{i\alpha} W_{\alpha\beta} X_{\beta j}^T. \quad (12)$$

Below few invariants are enumerated [7,8]:

$$\sum_{i,j} U_{ij} V_{ij}^{-1} = M, \quad (13)$$

$$\sum_{i,j} V_{ij} = \sum_{i,j} U_{ij}, \quad (14)$$

$$\frac{\sum_{i,k,l,j} V_{ik}^{-1} U_{kl} V_{lj}^{-1}}{\left(\sum_{ij} V_{ij}^{-1}\right)^2} = \left(\sum_{i,j} V_{ij}^{-1}\right)^{-1}. \quad (15)$$

The expression (13) is true for any model function. The sum on the left of the expression depends only on the number of parameters of the model. It should be noted that in case of non-linear model the matrices \mathbf{V} and \mathbf{U} are approximate covariance matrices. The expressions (14) and (15) are assumed to be true for any linear (relative to parameters) model function. This statement was proved for polynomials. For non-linear model functions the expressions (14) and (15) are approximate ones.

5 Checking the correspondence between shapes of uncertainties and model function

It was shown within one-parametric exactly solved models that starting from some value of the correlation a variance of the estimate falls into unphysical range; this value $\rho_0 = \sigma_i/\sigma_j$ can be defined as a *limit correlation* [9]. Physically possible values of correlation are determined by inequality:

$$\rho_{ij} < \sigma_i/\sigma_j, \quad \text{if } \sigma_i < \sigma_j, \quad (16)$$

where σ_i and σ_j – uncertainties of experimental errors ε_i and ε_j , ρ_{ij} – correlation between ε_i and ε_j . In exactly solved problem with non-constant model function the inequality (16) is transformed in following one

$$\rho_{ij} < \rho_0 = (y(x_j)/y(x_i))/(\sigma_j/\sigma_i). \quad (17)$$

The inequality (17) is transformed into restriction (16) if a model function is a constant. So, inequality (17) is a generalized one. The ratios $r = y(x_j)/y(x_i)$, σ_j/σ_i can be defined as parameters of the *shape* for model function and uncertainties respectively. As seen from (17) *limit correlation* equals unit only in case when *shape's* parameters equal each other ($y(x_j)/y(x_i))/(\sigma_j/\sigma_i) = 1$). The inequality (17) can be rewritten in following form

$$\rho < \rho_0 = (\sigma_i/y(x_i))/(\sigma_j/y(x_j)) \quad (18)$$

Thus, the inequality (18) imposes restriction on the experimental uncertainties in relative (to approximant's value) units. Consequently, even for linear model function a statistical analysis of measurements with generalized least squares method must be carried out in iterative way. The

inequality (18) must be applied for refinement of calculated covariance matrices of experimental errors from unphysical structures.

6 Summary

Following criteria of the validation of experimental and estimated covariance data are reviewed: a) criterion of the positive definiteness for covariance matrices, b) relationship between the “integral” experimental and estimated uncertainties, c) validity of the statistical invariants, d) restrictions (which are induced by differences in the shapes of uncertainties and model function) imposed to correlations between experimental errors. Applying these criteria in nuclear data evaluation was considered. In particular, a) preserving positive definiteness of covariance matrices in case of arbitrary transformation of a random vector was considered, properties of the covariance matrices in operations widely used in neutron and reactor physics (splitting and collapsing energy groups, averaging the physical values over energy groups, estimation parameters on the basis of measurements by means of generalized least squares method) were studied; b) an algorithm for comparison of experimental and estimated “integral” uncertainties was developed, square root of determinant of a covariance matrix is recommended for use in nuclear data evaluation as a measure of “integral” uncertainty for vectors of experimental and estimated values; c) a set of statistical invariants-values which are preserved in statistical processing-was presented; d) the inequality which signals on falling a correlation between experimental errors in unphysical range is given.

References

1. G.P. Lamaze et al., Nucl. Sci. Eng. **68**, 183 (1978).
2. E. Fort, J.P. Marquette, Report EANDC(E)-148“U”, 1972.
3. E. Fort, in *Proceedings of Second IAEA Conference on Nuclear Data for Reactors, Helsinki, Finland, 15–19 June 1970*, Vol. 1, (1970), p. 253.
4. W.P. Poenitz, J.W. Meadows, in *Proceedings of Conference: Panel on Neutron Standard Reference Data, Vienna, Austria, (1972)*, p. 102.
5. S.J. Friesenhahn et al., Report Intel – RT7011-001, 1974.
6. S.A. Badikov, E.V. Gai, International Atomic Energy Agency, Report INDC(NDS)-453, 2004, p. 207.
7. E.V. Gai, V.G. Pronyaev (these proceedings).
8. S.A. Badikov (private communication to V.G. Pronyaev).
9. E.V. Gai, S.A. Badikov, International Atomic Energy Agency, Report INDC(NDS)-453, 2004, p. 139.