Statistical Invariants in the Least Squares Method

S.A.Badikov

PI “Atomstandard”, Rosatom, Moscow

Technical Meeting on Neutron Data Standards,
January 26-30, 2026 IAEA, Vienna

Description of the estimation problem

It's necessary to estimate the unknown vector of parameters $\vec{\beta}$ on the basis of results of n measurements y_i of the model function $f(x,\vec{\beta})$

$$y_i = f(x_i,\vec{\beta}) + \varepsilon_i , \qquad i = 1,\dots,n .$$

The results of the measurements are distorted by the experimental errors ε_i . The variations of the experimental errors are described by the covariances V_{ij}

$$V_{ij} = cov(\varepsilon_i, \varepsilon_j)$$

The model function $f(x,\vec{\beta})$ is an element of a vector space of dimension *L*

.

Interpretation of the estimation process

- The set of the experimental data y_i, $i = 1,\dots,n$ with covariances V_{ij} can be interpreted as a system of *n* particles with coordinates y_i ; the interaction between particles is described by the values V_{ij}
- in turn, statistical processing (application of LSM) can be interpreted as a transition F of the *n* – particle system from one state (y_i , V_{ij}) to another one ($\hat{y}_i$, R_{ij}) :

$$F : \quad (y_i , V_{ij}) \quad \Rightarrow \quad (\hat{y}_i , R_{ij})$$

- we are looking for quantities which are stay unchanged at transition

Definition of a scalar product.

A scalar product in the normalized vector space $\boldsymbol{\Omega}$ can be defined as follows

$$\langle f_k(\vec{x}) \bullet f_l(\vec{x}) \rangle = \sum_i \sum_j f_k(x_i)(V^{-1})_{ij} f_l(x_j)$$

Such the definition meets all the requirements for the scalar product

- commutativity
- distributivity
- uniformity
- positive definiteness (if V - positive definite matrix)

Representation of the model function through the basis functions

If a set of functions $\varphi_0(x),\dots,\varphi_{L-1}(x)$ form a basis in the space Ω then the function $f(x,\vec{\beta})$ can be represented as a linear combination of these functions

$$f(x,\vec{\beta}) \equiv \sum_{m=0}^{L-1} \gamma_m \varphi_m(x)$$

Using the standard procedure of orthogonalization it's possible to transform the initial basis $\varphi_0(x),\dots,\varphi_{L-1}(x)$ into the orthogonal one $\psi_0(x),\dots,\psi_{L-1}(x)$

$$\langle \psi_k(x) \cdot \psi_l(x) \rangle = \delta_{kl}$$

Correspondingly, the model function takes the form

$$f(x,\vec{\beta}) \equiv g(x,\vec{\theta}) = \sum_{m=1}^{L} \theta_m \psi_m(x)$$

Representation of the model function through the basis functions

If a set of functions $\varphi_0(x),...,\varphi_{L-1}(x)$ form a basis in the space Ω then the function $f(x,\vec{\beta})$ can be represented as a linear combination of these functions

$$f(x,\vec{\beta}) \equiv \sum_{m=0}^{L-1} \gamma_m \varphi_m(x)$$

Using the standard procedure of orthogonalization it's possible to transform the initial basis $\varphi_0(x),...,\varphi_{L-1}(x)$ into the orthogonal one $\psi_0(x),...,\psi_{L-1}(x)$

$$\langle \psi_k(x) \cdot \psi_l(x) \rangle = \delta_{kl}$$

Correspondingly, the model function takes the form

$$f(x,\vec{\beta}) \equiv g(x,\vec{\theta}) = \sum_{m=1}^{L} \theta_m \psi_m(x)$$

Estimation problem after transformation of the basis

After orthogonalization of the basis in the space Ω the regression equation can be written as follows

$$y_i = \sum_{m=0}^{L-1} \theta_m \psi_m(x) + \varepsilon_i \; , \qquad\qquad i = 1,\dots,n \, .$$

$$V_{ij} = cov(\varepsilon_i, \varepsilon_j)$$

where the basis functions $\psi_0(x),\dots,\psi_{L-1}(x)$ are orthogonal ones, $\vec{\theta}$ - vector to be estimated.

Thus, the initial estimation problem with an arbitrary model function from the space Ω was reduced to the problem with a linear model function.

The LSM estimate

The LSM estimate $\hat{\vec{\theta}}$ with the covariance matrix W for the linear model function is well known

$$\hat{\vec{\theta}} = (X^T V^{-1} X)^{-1} X^T V^{-1} \vec{y}$$

$$W = (X^T V^{-1} X)^{-1}$$

where $X_{ij} = \left. \frac{\partial f(x, \vec{\beta}(\vec{\theta}))}{\partial \theta_i} \right|_{x=x_j} = \left. \frac{\partial \left\{ \sum_{m=0}^{L-1} \theta_m \psi_m(x) \right\}}{\partial \theta_i} \right|_{x=x_j}$ is the matrix of the sensitivity coefficients.

The covariance matrix of estimated values $\hat{y}_i = f(x_i, \hat{\vec{\beta}}(\hat{\vec{\theta}})) \equiv \sum_{m=0}^{L-1} \hat{\theta}_m \psi_m(x)$ of the model function is given by the following expression

$$R_{ij} = cov(\hat{y}_i, \hat{y}_j) = XWX^T$$

Representation of the LSM estimate in the orthonormal basis

$$\hat{\vec{\theta}} = X^T V^{-1} \vec{y}$$

$$W = E \qquad \left(W_{kl} = \sum_{i=1}^{n} \sum_{j=1}^{n} \psi_k(x_i)(V^{-1})_{ij} \psi_l(x_j) = \delta_{kl} \right)$$

$$R = XX^T$$

where

$$X_{ij} = \left.\frac{\partial f(x, \vec{\beta}(\vec{\theta}))}{\partial \theta_i}\right|_{x=x_j} = \left.\frac{\partial \left\{ \sum_{m=0}^{L-1} \theta_m \psi_m(x) \right\}}{\partial \theta_i}\right|_{x=x_j} = \psi_i(x_j)$$

Statistical invariants

There are **strict** relationships between the characteristics of the system in original and final states (for **nonlinear** model function the relationships are **approximate**)

$$\sum_i c_i \hat{y}_i = \sum_i c_i y_i \qquad \sum_i \sum_j c_i R_{ij} c_j = \sum_i c_i V_{ij} c_j$$

where weights c_i are determined as follows

$$c_i = \frac{\sum_j (V^{-1})_{ji}}{\sum_k \sum_j (V^{-1})_{jk}}$$

Thus, the evaluated values $\hat{y}_i$ and their covariances R_{ij} are result of a **redistribution** of the experimental values y_i and their covariances V_{ij} .

The redistribution is managed by the weights c_i .

Interpretation of weights

$$c_i = \frac{\sum_j (V^{-1})_{ji}}{\sum_k \sum_j (V^{-1})_{jk}}$$

a share of overall information on the uncertainty of the multipoint system related to the point *i*

Interpretation of the statistical invariants

The invariants have a clear statistical interpretation

$$\sum_i c_i \hat{y}_i = \sum_i c_i y_i$$

Average (weighted in special way) value of the model function in the range under consideration

$$\sum_i \sum_j c_i R_{ij} c_j = \sum_i c_i V_{ij} c_j$$

Variance of the average (weighted in special way) value of the model function in the range under consideration

Side results during the derivation process

Trace of the matrix RV^{-1} is equal to the dimension L of the basis (= dimension of the vector space ***Ω***) : $Tr(RV^{-1}) = L$

$$\sum_{i=1}^{n}(RV^{-1})_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n}R_{ij}(V^{-1})_{ji} =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\left\{\sum_{k=0}^{L-1}\psi_k(x_i)\psi_k(x_j)\right\}(V^{-1})_{ji} =$$

$$= \sum_{k=0}^{L-1}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\psi_k(x_i)(V^{-1})_{ji}\psi_k(x_j)\right\} = \sum_{k=0}^{L-1}\delta_{kk} = L$$

Useful enequalities for the experimental covariances

$$\frac{\sum_j (V^{-1})_{ji}}{\sum_k \sum_j (V^{-1})_{jk}} \geq 0$$

from positivity of weights

$$\sum_k \sum_j (V^{-1})_{jk} > 0$$

from positivity of variance

of the physical quantity

Example 1

Ω – the space of polynomials of degree lower than *L*

1. Basis $\varphi_0(x),\ldots,\varphi_{L-1}(x)$:

$$1, \quad x, \ldots, \quad x^{L-1}$$

2. Orthonormal basis $\psi_0(x),\psi_1(x),\ldots,\psi_{L-1}(x)$ is constructed as follows

$$\lambda\psi_j(x) = (x-\alpha_{j-1})\psi_{j-1}(x) - \alpha_{j-2}\psi_{j-2}(x) - \cdots - \alpha_0\psi_0(x)$$

where coefficients $\alpha_0,\alpha_1\ldots,\alpha_{j-1}$ are calculated from requirements of orthogonality

$$\langle \psi_j(x)\cdot\psi_k(x)\rangle = 0, \qquad k = 1,\ldots, j-1$$

or

$$\alpha_k = \frac{\langle x\psi_{j-1}(x)\cdot\psi_k(x)\rangle}{\langle \psi_k(x)\cdot\psi_k(x)\rangle}, \quad k = 1,\ldots, j-1$$

Example 2

Ω – the space of piece-wise constant functions

$$f(x,\vec{\beta}) = \beta_k \quad \text{if} \quad x \in [x_{i_k}, x_{i_{k+1}}), \qquad k = 1,\dots,L$$

1. Basis $\varphi_0(x),\dots,\varphi_{L-1}(x)$:

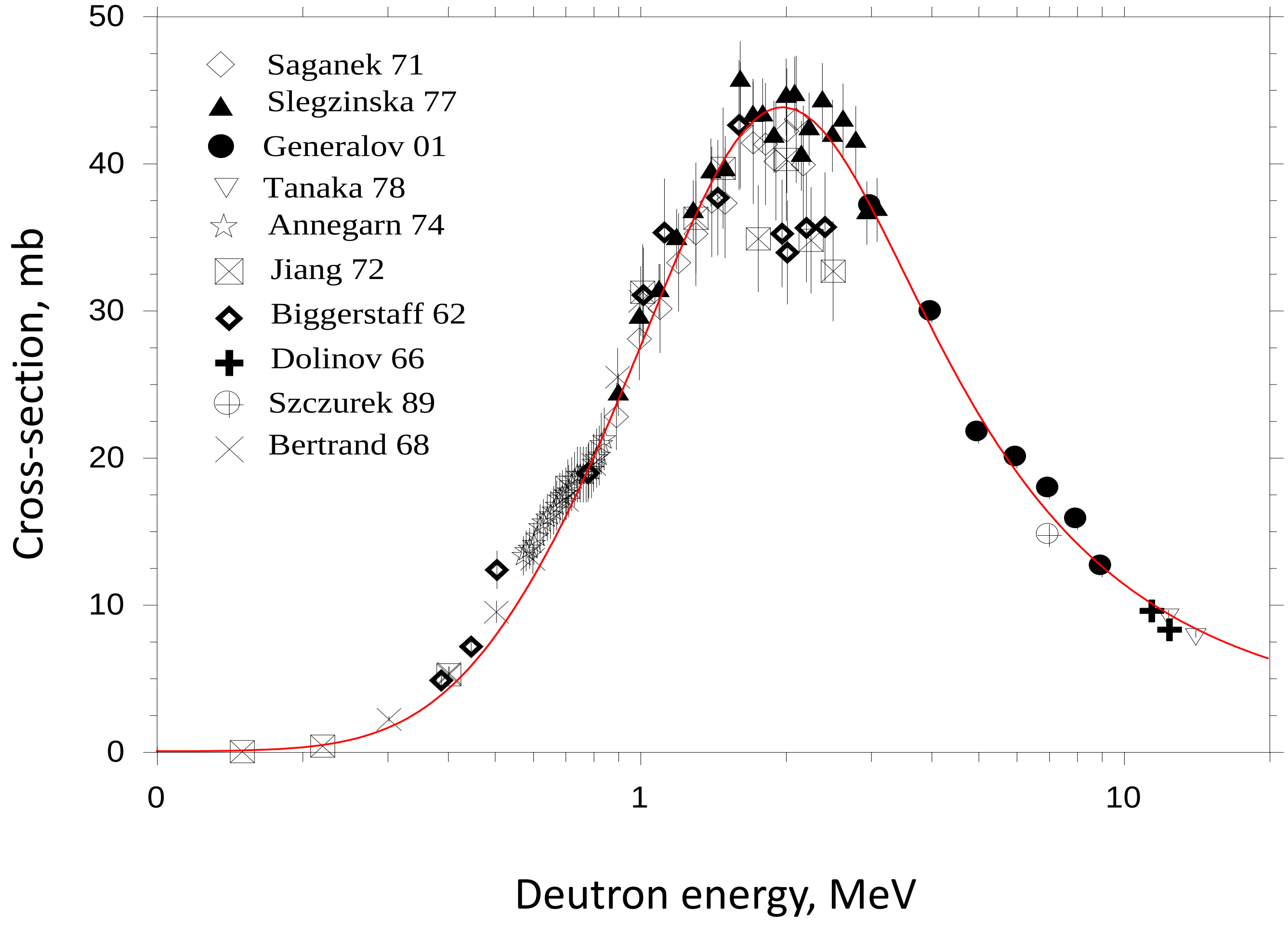
$$\varphi_k(x) = \begin{cases} 1 & if \quad x \in [x_{i_k}, x_{i_{k+1}}) \\ 0 & if \quad x \notin [x_{i_k}, x_{i_{k+1}}) \end{cases}$$

Example 3

Ω – the space of continuous functions in the range under consideration

- As known any continuous function can be uniformly approximated by a sequence of polynomials
- As already shown for any function from the finite-size polynomial space the statistic invariants are true
- Consequently, it should be expected that for continuous functions the statistic invariants will be approximately true.

Checking the statistical invariants.
Evaluation of the Be9(d,α0) reaction cross-section



Evaluation of the 9Be(d,α0) reaction cross-section.

Checking the statistical invariants

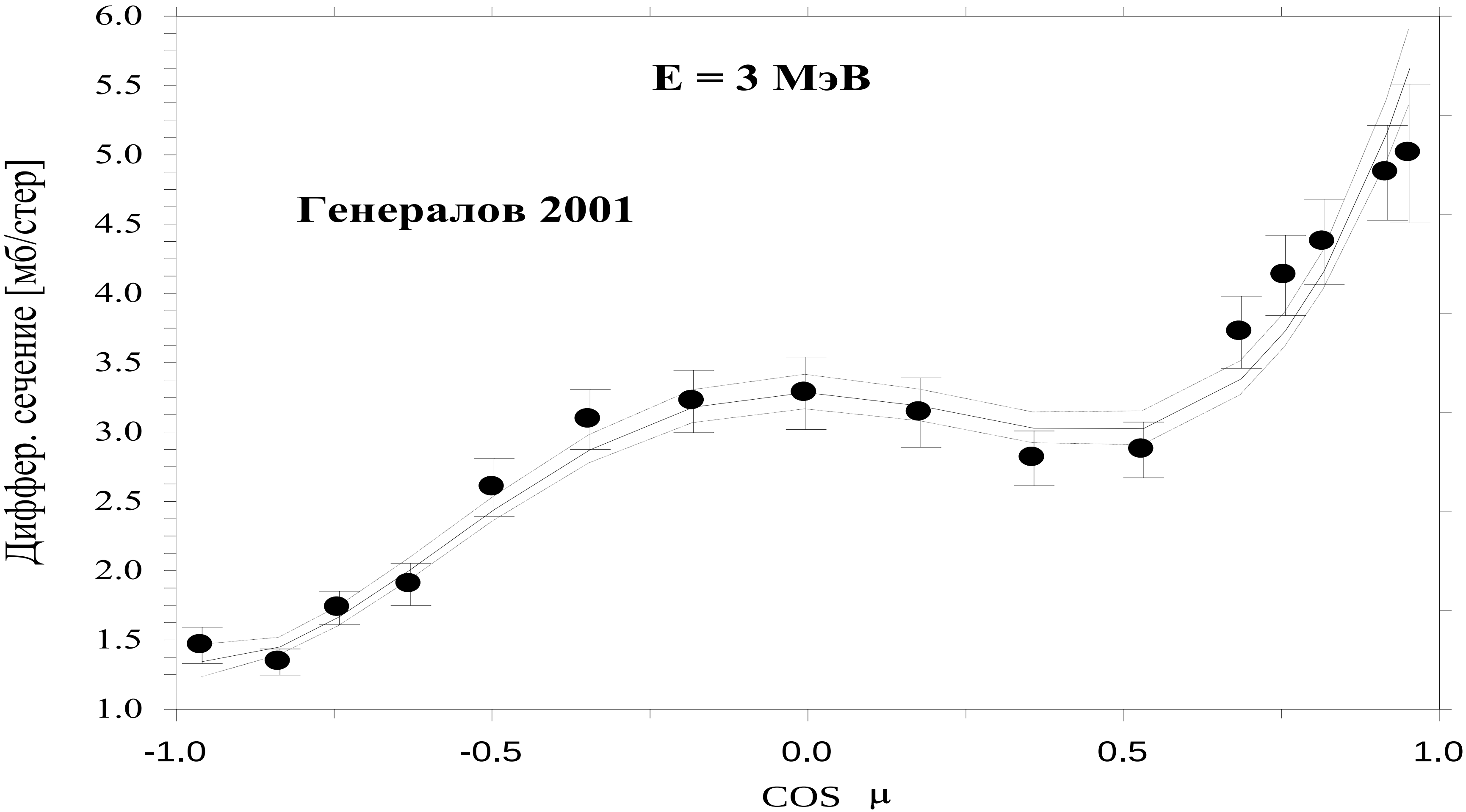
$\sum_i c_i y_i ,mb$	$\sum_i c_i \hat{y}_i ,mb$	$\sum_i \sum_j c_i V_{ij} c_j ,mb^2$	$\sum_i \sum_j c_i R_{ij} c_j ,mb^2$
$0.4548 \cdot 10^{-3}$	$0.4548 \cdot 10^{-3}$	$0.8793 \cdot 10^{-9}$	$0.8793 \cdot 10^{-9}$

Statistical invariants. Example.
Evaluation of the 9Be(d,α0) differential reaction cross-section at deutron energy 3 MeV. Results of measurements [2]

angle, grad	c-section, mb/ster	Uncertainty,%	Angle, grad	c-section, mb/ster	Uncertainty,%
17.7	5.01	10	90.2	3.28	8
23.5	4.87	7	100.4	3.22	7
35.2	4.37	7	110.2	3.09	7
40.9	4.13	7	119.8	2.60	8
46.7	3.72	7	129.0	1.90	8
57.9	2.87	7	137.9	1.73	7
69.0	2.81	7	146.7	1.34	7
79.7	3.14	8	163.5	1.46	9

[2] Generalov L.N. et al., "LI Meeting on Nuclear Spectroscopy and Nuclear Structure", P. 187. Sarov, RFNC-VNIIEF, 2001 [in Russian], EXFOR F0530

Statistical invariants. Example.
Evaluation of the 9Be(d,α0) differential reaction cross-section at deutron energy 3 MeV. Plot of the experimental data



Statistical invariants. Example.

Evaluation of the 9Be(d,α0) differential reaction cross-section.

Evaluated coefficients of Legendre polynomial

$$\sigma(\mu, E) = \sum_{l=0}^{N} \theta_N^l P_l(\mu)$$

•

$\boldsymbol{\theta_4^0}$	$\boldsymbol{\theta_4^1}$	$\boldsymbol{\theta_4^2}$	$\boldsymbol{\theta_4^3}$	$\boldsymbol{\theta_4^4}$
2.968	**1.464**	**0.01839**	**1.020**	**0.8686**

Statistical invariants. Example.

Evaluation of the 9Be(d,α0) differential reaction cross-section.

Covariances (x1000) of evaluated coefficients of Legendre polynomial

$$\sigma(\mu,E)=\sum_{l=0}^{N}\theta_N^l P_l(\mu)$$

	Number	0	1	2	3	4
θ_4^0	0	3.517				
θ_4^1	1	1.770	7.577			
θ_4^2	2	-1.874	4.912	17.55		
θ_4^3	3	1.238	0.3945	6.929	23.78	
θ_4^4	4	2.701	0.1747	-0.7937	10.96	24.58

Evaluation of the 9Be(d,α0) differential reaction cross-section.

Checking the statistical invariants

$\sum_i c_i y_i , \frac{mb}{ster}$	$\sum_i c_i \hat{y}_i , \frac{mb}{ster}$	$\sum_i \sum_j c_i V_{ij} c_j , \frac{mb^2}{ster^2}$	$\sum_i \sum_j c_i R_{ij} c_j , \frac{mb^2}{ster^2}$
2.260	2.260	2.091-3	2.091-3

Summary

- **Input experimental data (results of measurements and their covariances) predetermine the evaluated data and their covariances calculated by the LSM for the model function;**

- **Weighted sum of elements of the covariance matrix is a natural measure of the integral uncertainty for the random vector.**

- **As follows from the conservation laws relative decreasing (increasing) uncertainties of the evaluated data leads to pumping uncertainty information into the off-diagonal covariances**

- **strict relationships between input experimental data and output evaluated data, restrictions imposed to the covariances of the experimental errors provide verification both the final and intermediate results of calculations**

Some statements of the DDEP methodology “to correct” results of the evaluation

- Uncertainty of the recommended value can not be lower the most accurate uncertainty of the experimental data

- If contribution of some measurement into statistical sum is larger than 50% than uncertainty of this result of the measurement is expended to get the 50% contribution.