NASA TECHNICAL MEMORANDUM NASA TM-75922

COMPARATIVE ANALYSIS OF VARIOUS APPROACHES TO EVALUATING THE PARAMETERS OF MULTIDIMENSIONAL MODELS

L. S. Guran

Translation of "Sravnitel'nyy analiz razlichnykh podkhodov k otsenke parametrov mnogomernykh modeley," Academy of Sciences USSR, Institute of Space Research, Moscow, Report Pr-350, 1977, pp. 1-14.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546 OCTOBER 1979

STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-75922		

4. Title and Subtitle	5. Report Date
COMPARATIVE ANALYSIS OF VARIOUS APPROACHES TO EVALUATING THE PARAMETERS OF MULTIDIMENSIONAL MODELS	October 1979
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
L. S. Guran	
	10. Work Unit No.
9. Performing Organization Name and Address	11. Contract or Grant No.
SCITRAN Box 5456 Santa Barbara, CA 93108	NASW-3198
	13. Type of Report and Period Covered
	Translation
12. Sponsoring Agency Name and Address	14. Sponsoring Agency Code
National Aeronautics and Space Administration Washington, D.C. 20546	

15. Supplementary Notes

Translation of "Sravnitel'nyy analiz razlichnykh podkhodov k otsenke parametrov mnogomernykh modeley," Academy of Sciences USSR, Institute of Space Research, Moscow, Report Pr-350, 1977, pp. 1-14.

16. Abstract

The problem of evaluating parameters of a multidimensional model according to measurement results is examined. The conditions on which the problem solution depends are itemized.

17. Key Words (Selected by Author(s))	18. Distribution Statement
	Unclassified - Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22.
Unclassified	Unclassified	15	

COMPARATIVE ANALYSIS OF VARIOUS APPROACHES FOR EVALUATION OF PARAMETERS OF MULTIDIMENSIONAL MODELS

L. S. Gurin

I. Formulation of the Problem

/3*

Let us examine the problem of evaluating parameters of a multidimensional model according to measurement results: Let $X^T=(x_1,\ldots,x_m)$ -- vector of independent variables, and $\theta^T=(\theta_1,\ldots,\theta_k)$ -- vector of evaluated parameters. Further, the measured value 3 may be presented in the form

$$y(x)=\eta(x,\theta)+\xi\,, \tag{1}$$

where the form of function $\eta(x,\theta)$, which represents the model studied, may be either given or selected on the basis of measurement results, and certain assumptions may be made about random error ξ. We will suppose that measurements are conducted at points x_i, $i=\overline{1,N}$, determined by the plan of the experiment. It is necessary to find the best evaluation of the unknown vector of parameters $\hat{\theta}$. The possible approaches to solution of this problem depend on the following conditions:

1. Is the model given or is it selected?
2. If the model is given, is the experiment plan also given? If the plan is not given, it becomes necessary to select the best plan (the problem of planning the experiment).
3. Is the law of error distribution ξ given?

---

*Numbers in margin indicate pagination of original foreign text.

4. What are the criteria for selecting the best evaluation?
5. What is the purpose of conducting the research?

The last point is the most important, since it determines the dependence between the remaining points. Therefore, we shall formulate the goal of the investigation conducted in this work.

Due to the complexity of multidimensional dependences, /4
many researchers consider it more expedient to produce decomposition of a model, ie., to study the dependence of the final value on each of the unknown variables x_i individually, fixing whenever possible the values of the remaining independent variables. If such decomposition is performed before the beginning of the experiment, it is expressed in the appropriate selection of experiment plan. This approach is severely criticized in [1] from the standpoint of planning the experiment. If the experiment has already been conducted and its plan did not correspond to the idea of decomposition, it is possible to perform approximate decomposition during processing of the results of the experiment (so-called plotting segments, see [2]).

On the other hand, when the experiment has been planned according to the scheme of decomposition, during processing we may perform reverse composition, returning to the model (1) in its overall multidimensional form.

Our goal is to clarify the dependence of the feasibility of decomposition on the circumstances formulated above.

2. Some General Considerations

Returning to [1], let us stress that here the authors examine only the case where the model is given. If the model is to be selected, then decomposition is necessary at the planning stage of the experiment or at least at the processing stage in order to define the problem and to at least outline a certain number of competitive models for subsequent study.

However, let us examine more closely the case where the experiment has already been conducted according to the method of decomposition and we are speaking only of selecting the method of processing. /5

We will begin with the question of criteria (question 4).

We will utilize the concept of conditionally effective evaluation introduced in [3], i.e., evaluation which has better precision characteristics with consideration of limitations according to other criteria (labor intensiveness of the algorithm, stability in relation to change in the law of error distribution). In each individual situation (except the most simple), the solution may be obtained only with application of the statistical modeling method on an electronic computer with consideration for machine time expenditure. The methodology of such an investigation in our case is analogous to that used in [3]. However, the compared evaluation algorithms will also be distinguished by whether or not composition was employed in them. It is clear that the result of such an investigation in its overall form cannot be predicted. In the simplest case, conclusions may be obtained analytically and speak in favor of composition.

3. Investigation of the Simplest Case

As the simplest case, let us examine a linear model with normal law of error distribution. As is known [4], evaluations of the method of least squares are the best in this case by their precision characteristics. Due to the fact that there are simple analytical expressions for these evaluations, the criterion of algorithm labor intensiveness may be considered insignificant. Thus, we have only to compare algorithms using and not using decomposition by their precision characteristics. For greater clarity (although this is not necessary), we will limit ourselves to the case m = 2, k = 2, i.e., we will examine model /6

$$y = \theta_1 x_1 + \theta_2 x_2 + \xi, \tag{2}$$

where error ξ -- independent normally distributed random values with zero mathematical expectation and dispersion σ^2. We will take the overall number of measurements equal to $N = n^2$ and we assume that plan matrix X has the form

$$X = \begin{Vmatrix} 0 & 0 \\ \frac{1}{n} & 0 \\ \frac{2}{n} & 0 \\ \vdots & \vdots \\ \frac{n-1}{n} & 0 \\ 0 & \frac{1}{n} \\ 0 & \frac{2}{n} \\ \vdots & \vdots \\ 0 & \frac{n-1}{n} \end{Vmatrix}. \tag{3}$$

Here informational matrix X^TX is equal to

$$X^TX = \begin{Vmatrix} \frac{(n-1)(2n-1)}{6} & \frac{(n-1)^2}{4} \\ \frac{(n-1)^2}{4} & \frac{(n-1)(2n-1)}{6} \end{Vmatrix}, \tag{4}$$

and error matrix

$$(X^TX)^{-1} = \left\| \begin{array}{cc} \frac{24(2n-1)}{(n-1)(7n^2+2n-5)} & -\frac{36}{7n^2+2n-5} \\ -\frac{36}{7n^2+2n-5} & \frac{24(2n-1)}{(n-1)(7n^2+2n-5)} \end{array} \right\| . \quad (5)$$

Thus, in processing without decomposition, the covariational evaluation matrix $(\hat{\theta}_1, \hat{\theta}_2)$ is equal to $\sigma^2(X^TX)^{-1}$, i.e. $\hat{\theta}_1$ and $\hat{\theta}_2$ separately are determined with dispersion

$$\sigma^2_{\hat{\theta}_1} = \sigma^2_{\hat{\theta}_2} = \frac{24\sigma^2(2n-1)}{(n-1)(7n^2+2n-5)} . \quad (6)$$

In the case of processing utilizing decomposition instead of model (2), we first obtain n of individual models /7

$$y_j = \theta_1 x_1 + C_j + \xi \; ; \; C_j = \theta_2 y_j \; ; \; j = \overline{1,n} \quad (7)$$

with the same plan matrix

$$X_j = \left\| \begin{array}{cc} 0 & 1 \\ \frac{1}{n} & 1 \\ \dots & \dots \\ \frac{n-1}{n} & 1 \end{array} \right\| . \quad (8)$$

In this case the error matrix appears as

$$(X_j^TX_j)^{-1} = \left\| \begin{array}{cc} \frac{12n}{n^2-1} & -\frac{6}{n+1} \\ -\frac{6}{n+1} & \frac{2(2n-1)}{n(n+1)} \end{array} \right\| . \quad (9)$$

By n we have obtained evaluation $\hat{\theta}_{1j}$ and $\hat{C}_j$ with dispersions determined by σ^2 and (9). From here for resulting evaluation $\hat{\theta}_{1d}$ (d -- index marking the use of decomposition), we will obtain dispersion

$$\sigma^2_{\hat{\theta}_{1d}} = \frac{12\sigma^2}{n^2-1} . \quad (10)$$

For evaluation $\hat{\theta}_{2d}$, we have problem

$$\hat{C}_j = \frac{j-1}{n}\theta_2 + \xi_c \qquad (11)$$

with plan matrix

$$X_2 = \left\| \begin{matrix} 0 \\ \frac{1}{n} \\ \cdots \\ \frac{n-1}{n} \end{matrix} \right\| ; \quad (X_2^T X_2)^{-1} = \frac{6n}{(n-1)(2n-1)} \qquad (12)$$

and dispersion ξ_c, equal to $\frac{2\sigma^2(2n-1)}{n(n+1)}$ (see (9)).

From here we quickly obtain /8

$$\sigma^2_{\hat{\theta}_{2d}} = \frac{2\sigma^2(2n-1)}{n(n+1)} \cdot \frac{6n}{(n-1)(2n-1)} = \frac{12\sigma^2}{n^2-1} . \qquad (13)$$

Comparing formulas (10) and (13) with formula (6), we obtain increase in dispersion due to decomposition (we omit indexes 1, 2 and $\hat{\theta}$):

$$\lambda_n = \frac{\sigma^2_{\hat{\theta}_d}}{\sigma^2_{\hat{\theta}}} = \frac{7n^2+2n-5}{2(n+1)(2n-1)} ; \qquad (14)$$

$$\lim_{n\to\infty} \lambda_n = \frac{7}{4} . \qquad (15)$$

To explain reasons for reduced precision with decomposition, we will note that in both cases the evaluations are linear functions of measurements y, but with decomposition the coefficients during individual measurements are not optimal, since they have the form $a_i b_j (i=\overline{1,n}; j=\overline{1,n}$ correspond to values $x_1 = \frac{i-1}{n}$, $x_2 = \frac{j-1}{n}$) instead of c_{ij}. Concrete expressions for

a_i, b_j and c_{ij} (which for brevity we do not write out) show that $c_{ij} \neq a_i b_j$.

We will note in conclusion that an analogous calculation may be performed for more complex linear models as well as for nonuniform distances between measurement points according to individual variables.

4. Example of a More Complex Problem

We will limit ourselves once again to the case of two independent variables, which may be more conveniently expressed as x and y. Moreover, $x_i = \frac{i-1}{n}$, $y_j = \frac{j-1}{n}$ $(i,j = \overline{1,n})$ the dependent variable will be expressed as Z. Further, let /9

$$z_{ij} = f(x_i, y_j, \theta) + \xi_{ij}, \tag{16}$$

where ξ_{ij} are again independent normally distributed random values with zero mathematical expectation and dispersion σ^2, and $f(x,y,\theta)$ has the form

$$f(x,y,\theta) = \frac{\alpha}{2\pi\sigma_{x_1}\sigma_{y_1}\sqrt{1-\rho_1^2}} \exp\left\{-\frac{1}{2(1-\rho_1^2)}\left[\frac{(x-a_1)^2}{\sigma_{x_1}^2} - \frac{2\rho_1(x-a_1)(y-b_1)}{\sigma_{x_1}\cdot\sigma_{y_1}} + \frac{(y-b_1)^2}{\sigma_{y_1}^2}\right]\right\} + \frac{1-\alpha}{2\pi\sigma_{x_2}\sigma_{y_2}\sqrt{1-\rho_2^2}} \exp\left\{-\frac{1}{2(1-\rho_2^2)}\left[\frac{(x-a_2)^2}{\sigma_{x_2}^2} - \frac{2\rho_2(x-a_2)(y-b_2)}{\sigma_{x_2}\sigma_{y_2}} + \frac{(y-b_2)^2}{\sigma_{y_2}^2}\right]\right\}. \tag{17}$$

Thus, in our example

$$\theta = (\alpha, a_1, b_1, \sigma_{x_1}, \sigma_{y_1}, \rho_1, a_2, b_2, \sigma_{x_2}, \sigma_{y_2}, \rho_2), \tag{18}$$

that is, $K = 11$.

Solution of the problem without decomposition using the method of least squares is reduced to minimalization of the expression

$$S = \sum_{i=1}^{n} \sum_{j=1}^{n} [z_{ij} - f(x_i, y_j, \theta)]^2 \qquad (19)$$

according to θ, with which we obtain evaluation $\hat{\theta}$. We will note that the problem of minimalization of S may turn out to be multiextremal. Let us examine a series of algorithms A_μ of minimalization. (Also included in the algorithm is the given rule of the end of computation, for example -- number of iterations.) Corresponding to each algorithm will be a certain computation time t_μ and a set of precision evaluation characteristics $\hat{\theta}_\mu$, for example, displacement vector $A\hat{\theta}_\mu$ and covariational matrix $B\hat{\theta}_\mu$. The problem of selecting μ becomes the problem of vector optimalization. Putting aside the question of selecting μ for the time being, let us examine the variant of applying decomposition.

With fixed $y=y_j$, we obtain, instead of (17), an individual model (we will note that decomposition is not simple by parameters): /10

$$f_1(x, \theta_1) = a \exp\{-b(x-a_1)^2 + c(x-a_1) + d\} + f \exp\{-l(x-a_2)^2 + p(x-a_2) - q\}, \qquad (20)$$

in which there are ten parameters:

$$\theta_1 = (a_1, a_2, a, b, c, d, f, l, p, q),$$

and

$$\left.\begin{aligned} a &= \frac{\alpha}{2\pi\sigma_{x_1}\sigma_{y_1}\sqrt{1-\rho_1^2}}; & f &= \frac{1-\alpha}{2\pi\sigma_{x_2}\sigma_{y_2}\sqrt{1-\rho_2^2}}; \\ b &= \frac{1}{2(1-\rho_1^2)\sigma_{x_1}^2}; & l &= \frac{1}{2(1-\rho_2^2)\sigma_{x_2}^2}; \\ c &= \frac{\rho_1(y-b_1)}{(1-\rho_1^2)\sigma_{x_1}\sigma_{y_1}}; & p &= \frac{\rho_2(y-b_2)}{(1-\rho_2^2)\sigma_{x_2}\sigma_{y_2}}; \\ d &= \frac{(y-b_1)^2}{2(1-\rho_1^2)\sigma_{y_1}^2}; & q &= \frac{(y-b_2)^2}{2(1-\rho_2^2)\sigma_{y_2}^2}. \end{aligned}\right\} \qquad (21)$$

To evaluate θ_1, it is necessary to minimize expression (n times with respect to j)

$$S_{1j} = \sum_{i=1}^{n} [z_{ij} - f_1(x_i, \theta_1)]^2. \qquad (22)$$

We are examining a series of algorithms A_γ. At the next stage, considering each value of $\hat{\theta}_{1j}$ as a measurement of initial vector of parameters θ, we obtain a system which looks like

$$\left.\begin{aligned} \hat{a}_{1j} &= a_1 + \xi_{1j}; \\ \hat{a}_{2j} &= a_2 + \xi_{2j}; \qquad j = \overline{1,n}; \\ \hat{a}_j &= \frac{\alpha}{2\pi\sigma_{x_1}\sigma_{y_1}\sqrt{1-\rho_1^2}} + \xi_{3j}; \\ &\dots \end{aligned}\right\} \qquad (23)$$

From this system, 10_n measurements are required to evaluate θ. The difficulty here lies in the fact that errors $\xi_{\chi j}$ $(\chi = \overline{1,10})$ have already been correlated and, generally speaking, do not have zero mathematical expectation.

In studying the problem by the method of statistical examination, we know the characteristics of errors and may continue /11

evaluation of θ by minimalization of the corresponding overall square form. However, in solving a real problem we have no such opportunity. Therefore, for comparative evaluation of the variant of decomposition, we will use ordinary sums of squares (here it becomes possible to achieve partial separation of parameters).

$$S_{d1} = \sum_{j=1}^{n} (\hat{a}_{1j} - a_1)^2 ; \quad S_{d2} = \sum_{j=1}^{n} (\hat{a}_{2j} - a_2)^2 ;$$

$$S_{d3} = \sum_{j=1}^{n} \left[\left(\hat{a}_j - \frac{\alpha}{2\pi\sigma_{x_1}\sigma_{y_1}\sqrt{1-\rho_1^2}} \right)^2 + \left(\hat{b}_j - \frac{1}{2(1-\rho_1^2)\sigma_{x_1}^2} \right)^2 + \right.$$

$$\left. + \left(\hat{c}_j - \frac{\rho_1 (y_j - b_1)}{\sigma_{x_1}\sigma_{y_1}(1-\rho_1^2)} \right)^2 + \left(\hat{d}_j - \frac{(y_j - b_1)^2}{2(1-\rho_1^2)\sigma_{y_1}^2} \right)^2 \right] ; \qquad (24)$$

$$S_{d4} = \sum_{j=1}^{n} \left[\left(\hat{f}_j - \frac{1-\alpha}{2\pi\sigma_{x_2}\sigma_{y_2}\sqrt{1-\rho_2^2}} \right)^2 + \cdots \right] .$$

Expressions of items not written out for S_{d4} are evident from (21) by analogy with S_{d3} .

From expressions for S_{d1} and S_{d2} , we promptly obtain

$$\hat{a}_1 = \frac{1}{n} \sum_{j=1}^{n} \hat{a}_{1j} \; ; \quad \hat{a}_2 = \frac{1}{n} \sum_{j=1}^{n} \hat{a}_{2j} \; . \qquad (25)$$

For minimalization of S_{d3} and S_{d4} , we use a set of algorithms $A_{\nu''}$, $A_{\nu'''}$. We will designate as A_ν the overall algorithm for solving the problem in case of decomposition, i.e., $\nu = (\nu', \nu'', \nu''')$. For each A_ν we will again obtain the overall time for solving the problem t_ν and a set of precision characteristics $A_{\hat{\theta}_\nu}$, $B_{\hat{\theta}_\nu}$.

Now the question of feasibility of decomposition may be answered.

First, it is necessary to select an approach to the problem of vector optimalization. For example, let the computation time /12 be given by limitation

$$t(\theta) \leq t_c , \tag{26}$$

and the vector of precision characteristics convoluted in scalar criteria -- average quadratic deviation of a certain linear form from components $\hat{\theta}$, i.e.,

$$\varepsilon(\theta) = M[\beta(\theta - \hat{\theta})]^2 . \tag{27}$$

There remains the difficulty that both t and ε depend on the value of the evaluated parameter θ. Let θ change in a certain domain Ω, Then the minimaxal approach may be used and (26) and (27) be replaced by

$$t = \sup_{\theta \in \Omega} t(\theta) \leq t_0 ; \tag{28}$$

$$\varepsilon = \sup_{\theta \in \Omega} \varepsilon(\theta) . \tag{29}$$

Now t and ε depend only on selected algorithms A_μ and A_ν. Leaving only those A_μ and A_ν for which (28) is being fulfilled and selecting one of the remaining algorithms from the condition of minimum ε, we solve the problem to its end. Let A_o be the algorithm selected. If $A_o = A_{\mu o}$, i.e., one of the algorithms is without decomposition, not only is the nonfeasibility of decomposition demonstrated, but the best algorithm for solving the problem is found. If $A_o = A_{\nu o}$, i.e., one of the algorithms has decomposition, then decomposition is feasible and again the best algorithm is found. We will note that we were able to include in this comparison also various realizations of decomposition itself (i.e., not only of type (20), but also others).

Realization of the described procedure with the aid of statistical modeling on an electronic computer is possible with a small number of values of μ and ν. A number of views on such a realization, particularly on multiextremal and vector optimalization, is given in [5]. /13

5. Conclusions

Thus, in evaluating parameters of multidimensional models, three cases are possible.

1. If the problem consists of selecting the best model, then it is expedient to perform decomposition either at the stage of planning the experiment or, if this is impossible, at the processing stage.

The latter may be done by constructing segments [2].

2. For linear models under conditions of utilizing the method of least squares (for example, with normal law of error distribution), decomposition is not feasible. If the experiment has already been conducted according to a scheme of decomposition, processing of results should be done according to the general model.

3. In all intermediate cases, the solution to the question of feasibility of decomposition depends on the concrete problem and may be obtained on the bases of a special investigation utilizing the method of statistical modeling, with consideration of all significant criteria, including labor intensiveness of the evaluation algorithm. The methodology of such an investigation is given above (section 4).

REFERENCES

/14

1. V. V. Nalimov, T. I. Golikova. Logicheskiye osnovaniya planirovaniya eksperimentov ("Logical Bases of Planning Experiments"), Moscow, "Metallurgiya," 1976, 128 p.

2. L. S. Gurin, N. F. Smirnova. Postroyeniye razrezov mnogomernykh mnogo-obraziy, poluchayemykh pri issledovanii zavisimostey mezhdu fizicheskimi velichinami ("Constructing Segments of Multidimensional Multivariants Obtained During Investigation of Dependences Between Physical Values"), Preprint IKI AN SSSR, Pr-253, Moscow, 1975, 22 p.

3. L. S. Gurin, K. A. Tsoy. Application of Statistical Modeling in Development and Comparative Analysis of Methods of Parameter Evaluation. In "Obraborka kosmicheskoy informatsii" ("Processing Cosmic Information"), Moscow, Nauka Press, 1976, 36-37.

4. S. R. Rao. Lineynyye statisticheskiye methody i ikh primeneniya ("Linear Statistical Methods and Their Applications"), Moscow, Nauka Press, 1968, 548 p.

5. L. S. Gurin, K. A. Tsoy. Uslovno-effektivnyye otsenki parametrov nelineynykh modeley ("Conditionally Effective Evaluations of Parameters of Nonlinear Models"), Preprint No. 257, IKI AN SSSR, Moscow, 1975, 34 p.