NASA Technical Memorandum

NASA TM -75669

CONCERNING AN APPLICATION OF THE METHOD OF LEAST SQUARES WITH A VARIABLE WEIGHT MATRIX

A. A. Sukhanov

Translation of "O primenenii metoda naimen'shikh kvadratov s peremennoy vesovoy matritsyey", Academy of Sciences USSR, Institute of Space Research, Moscow, Report Pr 481, 1979, pp 1-33

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

STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-75669		
4. Title and Subtitle CONCERNING AN APPLICATION OF THE METHOD OF LEAST SQUARES WITH A VARIABLE WEIGHT MATRIX.		5. Report Date 1979
		6. Performing Organization Code
7. Author(s) A. A. Sukhanov		8. Performing Organization Report No.
		10. Work Unit No.
9. Performing Organization Name and Address SCITRAN Box 5456, Santa Barbara CA 93108		11. Contract or Grant No. NASw-2791
		13. Type of Report and Period Covered translation
12. Sponsoring Agency Name and Address National Aeronautics & Space Administration Washington, D. C. 20546		14. Sponsoring Agency Code

15. Supplementary Notes Translation of "O primenenii metoda naimen'shikh kvadratov s peremennoy vesovoy matritsyey", Academy of Sciences USSR, Institute of Space Research, Moscow, Report Pr-481, 1979, pp 1-33

16. Abstract
The problem considered is that of obtaining an estimate of a vector involving the parameters for the state of a physical system when the weight matrix in the method of least squares is a function of this vector. An iterative procedure is proposed for calculating the desired estimate. We obtain conditions for the existence and uniqueness of the limit of this procedure, and a domain is found which contains the limit estimate. We also propose a second method for calculating the desired estimate which reduces to the solution of a system of algebraic equations. We consider the question of applying Newton's method of tangents to solving the given system of equations and obtain conditions for the convergence of the modified Newton's method. Certain properties of the estimate obtained are presented together with example.

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. Price
unclassified	unclassified	29	

ANNOTATION

The problem considered is that of obtaining an estimate of a vector involving the parameters for the state of a physical system when the weight matrix in the method of least squares is a function of this vector. An iterative procedure is proposed for calculating the desired estimate. We obtain conditions for the existence and uniqueness of the limit of this procedure, and a domain is found which contains the limit estimate. We also propose a second method for calculating the desired estimate which reduces to the solution of a system of algebraic equations. We consider the question of applying Newton's method of tangents to solving the given system of equations and obtain conditions for the convergence of the modified Newton's method. Certain properties of the estimate obtained are presented together with examples.

1. Statement of the problem

We shall consider the problem of determining the state parameters of some physical system from the results of measurements (for example, the problem of determining the trajectory parameters for a space /3* vehicle from the trajectory measurements). Let q be the m-dimensional vector of the parameters for the state of the system which are to be determined, q_* its actual value, $d(q)$ the n-dimensional vector of the functions being measured, $\bar{d}$ the vector of the measurements (i.e., the measured value of the vector $d(q_*)$). We shall assume that the function $d(q)$ is given.

The given problem of determining the vector q will be solved for the case when the precision of the measurements is a function of q. Such a function can stem from the following causes.

1. The precision of the measurements can depend on the magnitude of the measured function. For example, if distance is measured in terms of the travel time of a signal, then the precision of the measurements can deteriorate with an increase in distance, due to the weakness of the reflected signal, due to an imprecise value for the propagation velocity of the signal, etc. Taking into account the function $d(q)$, the precision of the measurements in the case under consideration is a function of q.

2. The components of the vector q can also be included in the functional relationship of the accuracy of measurements to the /4 sources of errors. For example, if moments of time are measured in which some object traverses the field of view of an optical sensor, then the measurement error is proportional to the time interval during which the object crosses the entire field of view of the sensor, but this interval depends on the velocity of the object, the distance to it, etc.

3. When interfering parameters are present (i.e., parameters of the model used which are not included in the number of parameters being determined, and are assumed equal to a priori values) the

*Numbers in margin indicate foreign pagination.

measurement error vector in the linear approximation is calculated by the formula [1]

$$\xi = \xi_0 + B\alpha$$

where ξ_0 is the instrument errors of the measurement, α is the vector of the errors in the knowledge of the interfering parameters, and $B = \partial d/\partial \alpha$. In the general case, the matrix B, and, therefore, the errors in the measurement depend on q. In the problem of statistical regression for the controlled variables which contain the errors, there is an analogous dependence of the measurement errors on the vector of the parameters to be measured [2].

In the given case the covariant matrix D of the measurement errors is also a function of the vector q . Henceforth, we shall assume that the function $D(q)$ is known.

The basic method for determining the vector q is the method of least squares (MLS). According to this method, the estimate $\hat{q}$ of the vector q_u is found from the condition for the quadratic form*

$$[\tilde{d} - d(q)]^T P [\tilde{d} - d(q)] \tag{1.1}$$

where P is the weight matrix of the MLS. If the function $d(q)$ (or is linearized in some neighborhood of the vector q_u), then the weight matrix which ensures the minimum dispersion for the error in determining any parameter linear with respect to $\hat{q}$ is the matrix /5

$$P = D^{-1}(q_u) \tag{1.2}$$

according to the Gauss-Markov theorem. In the given nonlinear problem we shall assume that the weight matric (1.2) is also in some sense the best matrix. However, since q_u is unknown, the weight matrix will be computed by the formula

$$P(\tilde{q}) = D^{-1}(\tilde{q})$$

where $\tilde{q}$ is some value of the vector q. In the general case $\tilde{q}$ can be distinguished from the value of the vector q used to compute the func-

* In (1.1) and hereafter the symbol "T" denotes the transpose of a vector and a matrix.

tions being measured. Then the MLS quadratic form (1.1) and the system of normal MLS equations can be written in the form

$$\varphi(q,\tilde{q}) = [\tilde{d} - d(q)]^T P(\tilde{q})[\tilde{d} - d(q)], \tag{1.3}$$

$$\beta(q,\tilde{q}) = -\frac{1}{2}\left[\frac{\partial\varphi(q,\tilde{q})}{\partial q}\right]^T = A^T(q)P(\tilde{q})[\tilde{d} - d(q)], \tag{1.4}$$

where $A(q) = \partial d/\partial q_0$. According to the assumption that the weight matrix (1.2) is optimum, the estimate

$$\hat{q}_* = \arg\min_q \varphi(q, q_*) \tag{1.5}$$

is the best MLS estimate. Since $\hat{q}_*$ is unknown, it is necessary to find the best approximation to the estimate (1.5). As criteria for the closeness of an estimate $\hat{q}$ to $\hat{q}_*$ or q_*, we shall assume the quantities $\|\hat{q} - \hat{q}_*\|$ and $\|\hat{q} - q_*\|$. When selecting a vector norm it must be kept in mind that the components of q have different dimensions. Henceforth, we shall assume that the vector q is reduced to dimensionless form as follows (likewise for $\tilde{q}$):

$$q = Sq_p,$$

where q_p is the input vector of the parameters to be determined whose components are written in the given units of the dimensions, S is a/6 constant diagonal matrix with positive elements which can be selected quite arbitrarily. The elements of S can be, e.g., those required for precision in determining the components of the vector q_p. The matrix S must be identical for all the vectors q and $\tilde{q}$. Any of the norms used can be applied to the dimensionless vector q thus obtained.

Let q_0 be an a priori given value of the vector q which belongs to a sufficiently small neighborhood of the vector q_*. We shall assume that the value $P(q_0)$ of the weight matrix makes it possible to $\hat{q}_1 = \arg\min_q \varphi(q, q_0)$ such that $\|\hat{q}_1 - q_*\| < \|q_0 - q_*\|$. Since the vector $\hat{q}_1$ is closer to q_* than q_0, we may expect that the weight matrix $P(\hat{q}_1)$ will permit further improvement in the estimate of the vector q

Continuing to reason analogously, we arrive at the recursion procedure*

$$\hat{q}_0 = q_0, \quad \hat{q}_{\kappa+1} = \arg\min_{q} \varphi(q, \hat{q}_\kappa) \quad (\kappa = 0, 1, \ldots). \tag{1.6}$$

In Sections 2 and 3 we shall formulate the conditions under which the procedure (1.6) converges to a unique estimate, sufficiently close to the value (1.5). Note that all the quantities which enter into the conditions of the theorems to be proved below are, generally speaking, random. Therefore, the conditions of the theorems must be considered either for some concrete sample, or as realizable with a certain probability.

2. The conditions for convergence of the procedure /7

Let q_0 be a fixed vector. We shall consider two neighborhoods of the vector q_0:

$$\Omega(q_0) = \{q : \|q - q_0\| < \omega\}, \quad \Omega'(q_0) = \{q : \|q - q_0\| < \nu\omega\}, \tag{2.1}$$

where $0 < \omega < \infty$, and

$$\nu = \begin{cases} 1 & \text{when } \|q\| = \max\limits_{1 \le i \le m} |q_i| \text{ is the norm used} \\ \sqrt{2} & \text{when } |q| = \sqrt{q^T q} \text{ is the norm used} \\ 2 & \text{when } \|q\| = \sum\limits_{i=1}^{m} |q_i| \text{ is the norm used} \end{cases} \tag{2.2}$$

(here q_i is the i the component of the vector q). Let us also consider the matrices $\Theta_1 = diag(\vartheta_1, \vartheta_m)$, $\Theta_2 = diag(\vartheta_1, \vartheta_{2m})$, $0 < \vartheta_i < 1$ $(i = \overline{1, 2m})$, (2.3)
I denotes the identity matrix. The following lemma is needed to prove the theorem given below.

Lemma 1. For any $q_1, q_2 \in \Omega(q_0)$ and any matrix Θ_1, defined by (2.3), the vector

$$q = q_1 + \Theta_1 (q_2 - q_1) \tag{2.4}$$

belongs to the neighborhood $\Omega'(q_0)$.

Proof: Let us consider the norm $\|q\| = \max\limits_{1 \le i \le m} |q_i|$. Denote by q_{0i}, q_{1i}, q_{2i} the i th components of the vectors q_0, q_1, q_2. From (2.4), (2.3) for all $i = \overline{1, m}$

* An analogous procedure is considered in [2].

$$|q_i - q_{ai}| = |(1-\vartheta_i)(q_{1i}-q_{ai}) + \vartheta_i(q_{2i}-q_{ai})| \leq (1-\vartheta_i)\|q_1 - q_a\| + \vartheta_i\|q_2 - q_a\| \leq (1-\vartheta_i+\vartheta_i)\omega = \omega.$$

Hence $\|q-q_a\| = \max_{1\leq i\leq m} |q_i - q_{ai}| \leq \omega.$

/8

Next, consider the norm $\|q\| = \sqrt{q^T q}$. From (2.4)

$$\|q-q_a\| = \|q_1 + \Theta_1(q_2-q_1) - \tfrac{1}{2}(q_1+q_2) + \tfrac{1}{2}(q_1+q_2) - q_a\| \leq$$
$$\leq \|(\tfrac{1}{2}I - \Theta_1)(q_1 - q_a + q_a - q_2)\| + \tfrac{1}{2}\|q_1 - q_a + q_2 - q_a\|. \tag{2.5}$$

From the definition of the matrix Θ_1, $\|\frac{1}{2}I - \Theta_1\| < \frac{1}{2}$. Let us introduce the notation $X_1 = q_1 - q_a$, $X_2 = q_2 - q_a$, $\cos\gamma = \frac{X_1^T X_2}{\|X_1\|\|X_2\|}$. Since $q_1, q_2 \in \Omega(q_a)$, it follows that $\|X_1\| \leq \omega$ and $\|X_2\| \leq \omega$. From this and (2.5) we obtain

$$\|q-q_a\| < \tfrac{1}{2}\|X_1 - X_2\| + \tfrac{1}{2}\|X_1 + X_2\| =$$
$$= \tfrac{1}{2}\sqrt{\|X_1\|^2 + \|X_2\|^2 - 2\|X_1\|\cdot\|X_2\|\cos\gamma} + \tfrac{1}{2}\sqrt{\|X_1\|^2 + \|X_2\|^2 + 2\|X_1\|\cdot\|X_2\|\cos\gamma} \leq$$
$$\leq \tfrac{\omega}{2}\left(\sqrt{2-2\cos\gamma} + \sqrt{2+2\cos\gamma}\right) = \omega\left(\sin\tfrac{\gamma}{2} + \cos\tfrac{\gamma}{2}\right) \leq \sqrt{2}\,\omega.$$

Finally, consider the norm $\|q\| = \sum_{i=1}^{m} |q_i|$. From (2.4) and (2.3) we have:

$$\|q-q_a\| = \|q_1 + \Theta_1(q_2-q_1) - q_a\| = \|(I-\Theta_1)(q_1-q_a) + \Theta_1(q_2-q_a)\| \leq$$
$$\leq \|(I-\Theta_1)(q_1-q_a)\| + \|\Theta_1(q_2-q_a)\| < \|q_1 - q_a\| + \|q_2 - q_a\| \leq 2\omega$$

Q.E.D.

Denote by Ω, Ω' the neighborhoods (2.1) when $q_a = q_0$ or $q_a = q_*$, i.e.,

$$\Omega = \Omega(q_0),\ \Omega' = \Omega'(q_0) \text{ or } \Omega = \Omega(q_*),\ \Omega' = \Omega'(q_*). \tag{2.6}$$

Also let

$$G = G(q,\tilde{q}) = \frac{\partial f(q,\tilde{q})}{\partial q},\quad H = H(q,\tilde{q}) = \frac{\partial f(q,\tilde{q})}{\partial \tilde{q}}, \tag{2.7}$$

/9

$$\Delta = \|\hat{q}_* - q_*\|, \tag{2.8}$$

where $\hat{q}_*$ is defined by (1.5). Along with the vectors q, $\tilde{q}$ we shall also use the composite vectors

$$z = \begin{bmatrix} q \\ \tilde{q} \end{bmatrix}, \quad z_k = \begin{bmatrix} \hat{q}_{k+1} \\ \hat{q}_k \end{bmatrix} \quad (k = 0, 1, \ldots), \tag{2.9}$$

where the estimates $\hat{q}_k$, $\hat{q}_{k+1}$ are determined by (1.6). Then the functions introduced above can be written in the form $\varphi(z)$, $f(z)$, $G(z)$, and $H(z)$.

Theorem 1. Suppose that the following conditions are satisfied:

1) $q_k, \hat{q}_k, q_0 \in \Omega$;
2) all the terms of the sequence $\{\hat{q}_k\}$ constructed according to (1.6) belong to Ω.

Moreover, assume that for any $z \in \Omega' \times \Omega''$ the following conditions are satisfied:

3) the derivatives $G(z)$, and $H(z)$ exist;
4) the matrix $G^{-1}(z)$ exists;
5) $\|G^{-1}(z) H(z)\| \leq \delta_0, \quad 0 < \delta_0 < 1.$

Then the following statements are valid:

a) for any term $\hat{q}_k$ of the sequence $\{\hat{q}_k\}$, the quantity $\hat{q}_{k+1}$ is uniquely determined in Ω by (1.6) (i.e., $\min_q \varphi(q, \hat{q}_k)$ is unique in Ω); the estimate $\hat{q}_k$ is also unique in Ω;

b) the limit $q_* \in \Omega$ of the sequence $\{\hat{q}_k\}$ exists, q_* is a solution of the equation $f(q, q) = 0$, and this solution is unique in Ω;

c) the estimate q_* satisfies the inequality

$$\|q_* - \hat{q}_k\| \leq \frac{\delta_0 \Delta}{1 - \delta_0}, \quad \|q_* - q_k\| \leq \frac{\Delta}{1 - \delta_0}, \tag{2.10}$$

where Δ is defined by (2.8).

Proof: /10

a) assume that for some $\hat{q}_k \in \Omega$ there exist $\hat{q}^1_{k+1}, \hat{q}^2_{k+1} \in \Omega$. It follows from (1.4) and (1.6) that $f(\hat{q}^1_{k+1}, \hat{q}_k) - f(\hat{q}^2_{k+1}, \hat{q}_k) = 0$. There exists a vector $q = \hat{q}^1_{k+1} + \Theta_1 (\hat{q}^2_{k+1} - \hat{q}^1_{k+1})$, where the matrix Θ_1 is defined by (2.3) and $q \in \Omega'$ by Lemma 1, so that

$$f(\hat{q}^2_{\kappa+1}, \hat{q}_\kappa) - f(\hat{q}^1_{\kappa+1}, \hat{q}_\kappa) = G(q, \hat{q}_\kappa)(\hat{q}^2_{\kappa+1} - \hat{q}^1_{\kappa+1}) = 0.$$

By condition 4) of the theorem, the matrix G is nonsingular everywhere in Ω'; hence $\hat{q}^1_{\kappa+1} = \hat{q}^2_{\kappa+1}$.

The uniqueness of the estimate $\hat{q}_0$ is proved similarly.

b) by virtue of (1.4), (1.6) and taking into account the notation (2.9), $f(z_\kappa) = f(z_{\kappa-1}) = 0$. Hence it is possible to write the equality

$$f(z_\kappa) - f(z_{\kappa-1}) = \frac{\partial f(z)}{\partial z}(z_\kappa - z_{\kappa-1}) = G(z)(\hat{q}_{\kappa+1} - \hat{q}_\kappa) + H(z)(\hat{q}_\kappa - \hat{q}_{\kappa-1}) = 0, \tag{2.11}$$

where $z = z_{\kappa-1} + \theta_2(z_\kappa - z_{\kappa-1})$, the matrix θ_2 is defined by (2.3) and $z \in \Omega' \times \Omega'$ according to Lemma 1. Let us introduce the notation $\Delta_{\kappa,\ell} = \|\hat{q}_\kappa - \hat{q}_\ell\|$, where $\hat{q}_\kappa$ and $\hat{q}_\ell$ are terms of the sequence $\{\hat{q}_\kappa\}$. Then by (2.11) and conditions 4 and 5) of the theorem we obtain

$$\Delta_{\kappa+1,\kappa} = \|G^{-1}(z)H(z)(\hat{q}_\kappa - \hat{q}_{\kappa-1})\| \leq \delta_0 \Delta_{\kappa,\kappa-1} \leq \delta_0^\kappa \Delta_{1,0}.$$

From this and the triangle inequality we obtain for any K, ℓ = 0, 1, ...

$$\Delta_{\kappa+\ell,\kappa} \leq \sum_{i=\kappa}^{\kappa+\ell-1} \Delta_{i+1,i} \leq \Delta_{1,0} \sum_{i=\kappa}^{\kappa+\ell-1} \delta_0^i = \Delta_{1,0}\delta_0^\kappa \frac{1-\delta_0^\ell}{1-\delta_0} \leq \frac{\Delta_{1,0}\delta_0^\kappa}{1-\delta_0}. \tag{2.12}$$

According to condition 2) of the theorem, $\Delta_{1,0} \leq 2\omega$ and from (2.12)

$$\Delta_{\kappa+\ell,\kappa} \leq \frac{2\omega\delta_0^\kappa}{1-\delta_0}.$$

/11

Thus the conditions of the Cauchy criterion for the convergence of the sequence $\{\hat{q}_\kappa\}$, formed in accordance with (1.6), are satisfied [3]: for any $\varepsilon > 0$ there exists $K_0 = \left[\ln \frac{\varepsilon(1-\delta_0)}{2\omega} / \ln \delta_0\right]$ the brackets denote here the integral part of a number) such that for any $K > K_0$, $\ell > 0$ the inequality $\|\hat{q}_{\kappa+\ell} - \hat{q}_\kappa\| \leq \varepsilon$ is satisfied. Consequently, $\lim\limits_{\kappa\to\infty} \hat{q}_\kappa = q_*$ exists and $q_* \in \Omega$, since $\hat{q}_\kappa \in \Omega$ (K = 0,I,...). Since $f_\kappa = f(\hat{q}_{\kappa+1}, \hat{q}_\kappa) = 0$ for all K, $\lim\limits_{\kappa\to\infty} f_\kappa = f(q_*, q_*) = 0$.

Assume that there exist $q_*^1, q_*^2 \in \Omega$ such that $q_*^1 \neq q_*^2$ and $f(z_*^1) = f(z_*^2) = 0$, where $z_*^i = \begin{bmatrix} q_*^i \\ q_*^i \end{bmatrix}$ (i = 1,2). By analogy with (2.11) we write:

$$f(z_*^2) - f(z_*^1) = G(\bar{z})(q_*^2 - q_*^1) + H(\bar{z})(q_*^2 - q_*^1) = 0, \tag{2.13}$$

where $\bar{z} = z_*^1 + \theta_2(z_*^2 - z_*^1)$, and $\bar{z} \in \Omega' \times \Omega'$. From (2.13) and condition 5) of the theorem

$$\|q_*^2 - q_*^1\| = \|G^{-1}(\bar{z})H(\bar{z})(q_*^2 - q_*^1)\| \leq \delta_0 \|q_*^2 - q_*^1\| < \|q_*^2 - q_*^1\|.$$

This contradiction proves that the solution of the equation $f(q,q) = 0$ is unique in Ω.

c) in accordance with (1.4), (1.5), $f(\hat{q}_u, q_u) = 0$ and by analogy with (2.11) and (2.13) we may write

$$f(z_*) - f(z_u) = G(\bar{z})(q_* - \hat{q}_u) + H(\bar{z})(q_* - q_u) = 0, \tag{2.14}$$

where

$$z_* = \begin{bmatrix} q_* \\ q_* \end{bmatrix}, \quad z_u = \begin{bmatrix} \hat{q}_u \\ q_u \end{bmatrix}, \quad \bar{z} = z_u + \theta_2(z_* - z_u)$$

and according to Lemma 1, $\bar{z} \in \Omega' \times \Omega'$. From (2.14) and conditions 4) and 5) of the theorem

$$\|q_* - \hat{q}_u\| = \|G^{-1}(\bar{z})H(\bar{z})(q_* - \hat{q}_u + \hat{q}_u - q_u)\| \leq \delta_0(\|q_* - \hat{q}_u\| + \|\hat{q}_u - q_u\|). \tag{2.15}$$

Hence, taking into account (2.8), the first of the inequalities (2.10) follows. The second inequality follows from the first: /12

$$\|q_* - q_u\| \leq \|q_* - \hat{q}_u\| + \|\hat{q}_u - q_u\| \leq \frac{\delta_0 \Delta}{1 - \delta_0} + \Delta = \frac{\Delta}{1 - \delta_0}.$$

Q.E.D.

/13

Corollary 1. If the conditions of Theorem 1 are satisfied, then for any $q_1, q_2 \in \Omega$, the equality $f(q_2, q_1) = 0$ is a necessary and sufficient condition for the validity of the equation $\varphi(q_2, q_1) = \min_q \varphi(q, q_1)$.

Necessity follows from (1.4) and sufficiency from the uniqueness of the solution of the equation $f(q, q_1) = 0$ which is proved in a manner analogous to the proof of assertion a) of Theorem 1.

Corollary 2. The rate of convergence of the sequence $\{\hat{q}_\kappa\}$, constructed in accordance with the procedure (1.6) is estimated by the inequalities

$$\|\hat{q}_\kappa - q_*\| \le \delta_o \|\hat{q}_{\kappa-1} - q_*\| \le \delta_o^\kappa \|q_o - q_*\| , \tag{2.16}$$

$$\|\hat{q}_\kappa - q_*\| < \frac{\delta_o^\kappa}{1-\delta_o} \|\hat{q}_1 - q_o\| . \tag{2.17}$$

For the proof of (2.16) we write by analogy with (2.14)

$$G(z)(\hat{q}_\kappa - q_*) + H(z)(\hat{q}_{\kappa-1} - q_*) = 0 ,$$

where $z = z_* + \theta_1 (z_{\kappa-1} - z_*)$. Hence, and also from condition 5) of Theorem 1, the inequalities (2.16) follow. The inequality (2.17) is an immediate consequence of (2.12) when $\ell \to \infty$.

Note 1. If the derivatives (2.7) are continuous in Ω', then according to a theorem about implicit functions and condition 4) of Theorem 1, the relation $f(q, \tilde{q}) = 0$ uniquely defines for any $q, \tilde{q} \in \Omega'$ a continuously differentiable function

$$q = q(\tilde{q}) . \tag{2.18}$$

/13

It follows from condition 5) of Theorem 1 that the mapping (2.18) is compact for all $q, \tilde{q} \in \bar{\Omega}$ and Theorem 1 can be proved with the help of the fixed point principle [3].

Note 2. In the general case the limit q_* of the procedure (1.6) is not the minimum of the sequence $\{\|\hat{q}_\kappa - \hat{q}_\text{н}\|\}$. In fact, at some stage of the procedure, it may turn out that $\hat{q}_\kappa = q_\text{н}$. Then according to (1.5) and (1.6), and due to the uniqueness of the vectors $\hat{q}_{\kappa+1}$ and $\hat{q}_\text{н}$, the equalities $\hat{q}_{\kappa+1} = \hat{q}_\kappa$ and $\|\hat{q}_{\kappa+1} - \hat{q}_*\| = 0$ are valid, whereas in the general case, $\|q_* - \hat{q}_\text{н}\| = \lim_{\kappa \to \infty} \|\hat{q}_\kappa - \hat{q}_\text{н}\|$ is not zero.

Note 3. In real problems there is interest in the case $\Delta \ll \omega$. In the contrary case, any estimate $\hat{q}$ from Ω, even including $\hat{q} = q_o$ is good in the sense that $\|\hat{q} - q_\text{н}\| \sim \|\hat{q}_\kappa - q_*\|$.

Note that in real problems involving the determination of orbits, the conditions 1)-4) of Theorem 1 are usually satisfied and condition

5) is basic. Satisfiability of this condition is illustrated by an example considered in Section 7.

3. Improvement of the estimate obtained

From (1.4) and (2.7) it is easily seen that the matrix H is a linear function of the vector of the residual $\tilde{d}-d(q)$ whereas the matrix G does not depend on this vector. If the estimate (1.5) is the best MLS estimate, then the residual, computed from this estimate is minimum in the mean. Therefore, it may turn out that the quantity $\|G^{-1}H\|$, computed in a sufficiently small neighborhood of the vectors $\hat{q}_M$ and q_M will be substantially smaller than the quantity $\delta_0 \geq \max_{\Omega'}\|G^{-1}H\|$. /14 In this case a more accurate estimate of the norm $\|G^{-1}H\|$ may permit a decrease in the neighborhood of the vectors $\hat{q}_M$ and q_M which contains the limit q_* of the procedure (1.6), i.e., the right side of the inequality (2.10) is reduced. Moreover, note that δ_0 is a nondecreasing function of ω in (2.1) (in fact, when ω increases, the quantity $\max_{\Omega'}\|G^{-1}H\|$ can only increase). Since the estimates (2.10) are acceptable only when δ_0 is sufficiently small (say, not more than 1/2), then this imposes a restriction on ω. Therefore, estimates of the closeness of the quantity q_* to $\hat{q}_M$ and q_M independent of δ_0 make it possible to widen the neighborhoods Ω and Ω'. Other estimates will be obtained below.

Suppose that instead of condition 5) of Theorem 1, the following condition is satisfied: for any $q, \tilde{q} \in \Omega'$

$$\|G^{-1}(q,\tilde{q})H(q,\tilde{q})\| \leq \delta(x,y), \tag{3.1}$$

where

$$0 \leq \delta(x,y) \leq \delta_0 < 1, \tag{3.2}$$

and

$$x = \|q-q_1\|, \quad y = \|\tilde{q}-q_2\|, \tag{3.3}$$

Here q_1 and q_2 are vectors. If condition 5) of Theorem 1 is satisfied, a function $\delta(x,y)$, which satisfies (3.2), can be constructed as follows:

$$\delta(x,y) = \max_{\substack{\|q-q_1\|=x \\ \|\tilde{q}-q_2\|=y}} \|G^{-1}(q,\tilde{q})H(q,\tilde{q})\| . \tag{3.4}$$

Consider the function

$$w(u) = \max_{W(u)} \delta(x,y) , \tag{3.5}$$

where

$$W(u) = \{x,y : x \leq u+\Delta_1,\ y \leq u+\Delta+\Delta_2 ;\ q,\tilde{q} \in \Omega'\}, \tag{3.6}$$

$$\Delta_1 = \|q_1 - \hat{q}_*\| ,\quad \Delta_2 = \|q_2 - q_*\| , \tag{3.7}$$

Δ is determined by (2.8). /15

Note that when choosing the vectors q_1 and q_2 in (3.3) it is possible to make use of one of the following criteria:

—— the precision of the estimate (3.1);

—— the structural simplicity of the function $\delta(x,y)$ (for example, the simplicity with which the maximum in (3.4) is calculated);

—— the computational simplicity of the function $\omega(u)$ which will be needed subsequently (for example, if such q_1 and q_2 can be chosen that $\delta(x,y)$ is a monotonically increasing function of both variables, then according to (3.5) and (3.6), $w(u) = \delta(u+\Delta_1, u+\Delta+\Delta_2)$.

Consider the sequence $\{U_x\}$, formed in accordance with the equalities

$$u_0 = \frac{\delta_0 \Delta}{1-\delta_0} ,\quad u_{\kappa+1} = \frac{w(u_\kappa)\Delta}{1-w(u_\kappa)} \quad (\kappa = 0, 1, \ldots) \tag{3.8}$$

and the equation

$$u = \frac{w(u)\Delta}{1-w(u)} . \tag{3.9}$$

It is clear from (3.5) and (3.2) that any solution u_* of equation (3.9) satisfies the inequality

$$0 \leq u_* \leq \frac{\delta_0 \Delta}{1-\delta_0} . \tag{3.10}$$

Lemma 2. Assume that the function $\delta(x,y)$ satisfies the inequalities (3.2) for any $q, \tilde{q} \in \Omega'$. Then the following assertions are valid:

a) the sequence $\{u_\kappa\}$, formed in accordance with (3.8), has a limit u_* which is the greatest solution of equation (3.9);

b) for any u´, the inequality

$$u' \leqslant \frac{\omega(u')\Delta}{1-\omega(u')} \tag{3.11}$$ /16

is satisfied, as is the inequality $u' \leqslant u_*$.

Proof:

a) It is easy to see that ω(u) is a nondecreasing function. Indeed, for any u', u" such that u'< u", it follows from (3.6) that W(u')CW(u") and according to (3.5), ω(u')⩽ω(u").

Consider the sequence $\{U_k\}$. From (3.5) and (3.2) we find that $\omega(U_0) \leqslant \delta_0$. Hence from (3.8) $U_1 \leqslant U_0$. Therefore, since the function ω(u) is monotone, $\omega(U_1) \leqslant \omega(U_0)$, and from (3.8), $U_2 \leqslant U_1$. Reasoning analogously, we find that $U_{k+1} \leqslant U_k$ $(k = 0, 1, \ldots)$ It is clear from (3.8) (3.5) and (3.2) that every $U_k \geqslant 0$. Consequently, $\{U_k\}$ is a monotonically decreasing sequence which is bounded below, and has a limit U_* which satisfies the inequalities (3.10). Clearly, U_* is a solution of equation (3.9). Assume that there is a solution U'_* of this equation such that $U'_* > U_*$. Note that $U'_* \neq U_k$ for all U_k determined by (3.8), since, if for any K, $U'_* = U_k$, then U_k is a solution of equation (3.9) and, according to (3.8), $U'_* = U_k = U_{k+1} = \ldots = U_*$, contrary to hypothesis. On the other hand, for U'_*, as for all solutions of equation (3.9), the inequalities (3.10) hold. Consequently, by hypothesis, $U_* < U'_* < U_0$, and there exist terms U_l and U_{l+1} of the sequence $\{U_k\}$, such that $U_{l+1} < U'_* < U_l$. Hence, and also from (3.8) and the monotonicity of the function ω(u), we find that:

$$U'_* > U_{l+1} = \frac{\omega(U_l)\Delta}{1-\omega(U_l)} \geqslant \frac{\omega(U'_*)\Delta}{1-\omega(U'_*)} .$$

/17

The solution obtained contradicts the fact that U'_* is a solution of of equation (3.9). Hence U_* is the largest solution of that equation.

b) Suppose that U_* is the largest solution of equation (3.9). Assume that some $u' > U_*$ satisfies the inequality (3.11). Construct a sequence $\{U'_k\}$ in accordance with the equalities

$$u_0' = u', \quad u_{\kappa+1} = \frac{w(u_\kappa')\Delta}{1 - w(u_\kappa')} \quad (\kappa = 0, 1, \ldots). \tag{3.12}$$

Hence, in view of (3.11), it is clear that $u_1' \geq u_0'$. Consequently, $w(u_1') \geq w(u_0')$, and from (3.12), $u_2' \geq u_1'$. Continuing this line of reasoning, we obtain $u_{\kappa+1}' \geq u_\kappa'$ $(\kappa = 0, 1, \ldots)$. Here, since $w(u) \leq \delta_0$ for all u_κ', the inequality $u_\kappa' \leq \frac{\delta_0 \Delta}{1 - \delta_0}$ is satisfied. Consequently,

$\{u_\kappa'\}$ is a monotonically increasing sequence which is bounded above, and this sequence has a limit u_α^0 which is a solution of the equation (3.9), and $u_\alpha^0 \geq u' > u_\alpha$. But, according to this condition u_α is the largest solution of equation (3.9). This contradiction proves assertion b) of the lemma, and the lemma is proved.

Theorem 2. Suppose that conditions 1)-4) of Theorem 1 are satisfied, and for any $q, \tilde{q} \in \Omega'$ so are the conditions (3.1) and (3.2). Then assertions a) and b) of Theorem 1 hold and

$$\|q_* - \hat{q}_u\| \leq u_\alpha, \quad \|q_\alpha - q_u\| \leq u_\alpha + \Delta, \tag{3.13}$$

where u_* is the limit of the sequence $\{u_\kappa\}$ formed according to (3.8), and is the largest solution of equation (3.9).

Proof:

The validity of assertions a) and b) of Theorem 1 are obvious, since from (3.1) and (3.2) condition 5) of Theorem 1 follows.

Consider the equality (2.14). The vectors q and $\tilde{q}$, components of the vector Z in (2.14), satisfy the inequalities $\|q - \hat{q}_u\| \leq \|q_* - \hat{q}_u\|$, $\|\tilde{q} - q_u\| \leq \|q_* - q_u\|$. Hence, making use of the triangle inequality, taking account of the notations (3.3), (3.7), (2.8) and /18

$$u' = \|q_* - \hat{q}_u\|. \tag{3.14}$$

we obtain:

$$x \leq u' + \Delta_1, \quad y \leq u' + \Delta + \Delta_2. \tag{3.15}$$

From (3.1), (3.5), (3.6), (3.15) follows

$$\|G^{-1}(z)H(z)\| \leq \delta(x, y) \leq w(u'). \tag{3.16}$$

From (2.14) and (3.16), taking into account (2.8) and (3.14), we obtain

$$u' \leq w(u')(u' + \Delta).$$

similar to (2.15). Hence, it follows that U' satisfies (3.11). According to Lemma 2, $u' \leq u_*$, and hence, by virtue of (3.14), the inequalities (3.13) follow. Q. E. D.

The results of Theorem 2 make it possible to reduce substantially the domain in which the estimate q_* lies in case its magnitude is close to unity and $w(u_*) \ll \delta_0$. This is graphically illustrated in the example considered in Section 7.

4. The second method for obtaining an estimate /19

In computational terms the procedure (1.6) is laborious, since at each step an ordinary MLS problem must be solved (with a constant weight matrix). Therefore, we shall make use of the fact that if the conditions of Theorems 1 and 2 are satisfied, then the procedure (1.6) is equivalent to finding a solution to the equation

$$f(q, q) = 0 \tag{4.1}$$

(the equivalence follows from assertion b) of Theorem 1). From a computing standpoint, the procedure involved in solving equation (4.1) may turn out to be simpler than the procedure (1.6), since it allows us to avoid inserting iterative processes.

Let us consider solving equation (4.1) by Newton's method of tangents. According to this method, given an initial approximation $\tilde{q}_*$, a solution $\{q_k\}$ is found as the limit of the sequence formed in accordance with the recursion procedure [3-5]

$$q_{k+1} = q_k - \left[\frac{\partial f(q, q)}{\partial q}\bigg|_{q=q_k}\right]^{-1} f(q_k, q_k) \quad (k = 0, 1, \ldots). \tag{4.2}$$

From (2.7) we have $\frac{\partial f(q,q)}{\partial q} = G(q,q)+H(q,q)$

and the procedure (4.2) assumes the form

$$q_{k+1} = q_k - (G_k+H_k)^{-1} f_k \qquad (k=0,1,\ldots), \tag{4.3}$$

where

$$G_k = G(q_k, q_k),\; H_k = H(q_k, q_k),\; f_k = f(q_k, q_k).$$

We shall show that the inverse matrix in (4.3) exists.

Lemma 3. If the conditions of Theorem 1 are satisfied, then for any $Z \varepsilon \Omega' x \Omega'$, $[G(Z)+H(Z)]^{-1}$ exists.

Proof:

/20

Let us show that the matrix $G^{-1}(G+H)=I+G^{-1}H$ is nonsingular; then since the matrix G^{-1} is nonsingular, we shall have proved the nonsingularity of the matrix G+H. Suppose q is any vector. According to condition 5) of Theorem 1, we have

$$\|(I+G^{-1}H)q\| \geq \|q\| - \|G^{-1}Hq\| \geq \|q\|(1-\delta_0). \tag{4.4}$$

Since $\delta_0 < 1$, it follows from (4.4) that the equation $(I+G^{-1}H)q=0$ has only a zero solution. Consequently, the matrix $I+G^{-1}H$, and with it the matrix G+H also, is nonsingular, Q. E. D.

We shall assume that the remaining conditions for the convergence of Newton's method are satisfied*. Then the procedure (4.3) converges to the same limit q_α as the procedure (1.6).

5. Simplification of the computing procedure

The procedure (4.3), along with previous procedures, also has defects in comparison with the procedure (1.6), including the necess-

* Various conditions for the convergence of Newton's method are considered in [3-5]. For example, a sufficient condition for convergence is the existence of a bounded second derivative of the function [4].

ity of computing the matrix H at each iteration (it is easy to see that when minimizing the function $\varphi(q,\hat{q}_\alpha)$ with respect to q by Newton's method it is not necessary to compute this matrix). Therefore, let us consider the conditions under which the procedure /21

$$q_0' = q_0, \quad q_{\kappa+1}' = q_\kappa' - G_\kappa'^{-1} f_\kappa' \quad (\kappa = 0, 1, \ldots), \tag{5.1}$$

where

$$G_\kappa' = G(q_\kappa', q_\kappa'), \quad f_\kappa' = f(q_\kappa', q_\kappa'), \tag{5.2}$$

converges to the same limit q_* as the procedures (1.6) and (4.3). Note that (5.1) is a modification not only of the procedure (4.3), but also the procedure (1.6) in the case when the minima in (1.6) are found by Newton's method. Indeed, if in this case the procedure (1.6) is changed as follows: each minimum is not calculated individually after fixing the weight matrix, but at each iteration of Newton's method the value of the vector q obtained in the preceeding interation is substituted into the weight matrix, then we arrive at procedure (5.1).

In order for the procedure (5.1) to converge to q_*, it is sufficient for any terms of the sequence $\{q_\kappa'\}$ determined by (5.1) to satisfy the inequality

$$\|q_{\kappa+1}' - q_*\| \leq \beta \|q_\kappa' - q_*\|, \quad 0 < \beta < 1. \tag{5.3}$$

We shall formulate a theorem in which the convergence of the procedure (5.1) is proved under a condition of quite weak dependence of the matrix $G(q,\tilde{q})$ on q.

Theorem 3. Suppose that the conditions of Theorem 1 are satisfied together with the following conditions:

1) all terms of the sequence $\{q_\kappa'\}$ constructed according to (5.1) belong to the neighborhood Ω determined in (2.6) and (2.1);

2) for each term q_κ' of the sequence $\{q_\kappa'\}$ there exists such a vector $q_\kappa'' \in \Omega$ that $f(q_\kappa'', q_\kappa') = 0$, and for all $q = q_\kappa'' + \theta_1(q_\kappa' - q_\kappa'')$, where the matrix θ is determined by (2.5), the inequality

$$\| I - G_\alpha'^{-1} G(q, q_\alpha') \| \leq \alpha_0, \quad 0 < \alpha_0 < \frac{1-\delta_0}{1+\delta_0} , \tag{5.4}$$

is satisfied /22

where G_α' is determined by (5.2) and δ_0 is determined by condition 5) of Theorem 1.

Then the procedure (5.1) converges to the same limit q_0 as procedure (1.6).

Proof:

As follows from the conditions of the theorem, for any term of the sequence $\{q_\alpha'\}$ the quantity f_α' determined by (5.2) can be represented in the form

$$f_\alpha' = f(q_\alpha'', q_\alpha') + G \cdot (q_\alpha' - q_\alpha'') = G \cdot (q_\alpha' - q_\alpha'') , \tag{5.5}$$

where $G = G(q, q_\alpha')$, $q = q_\alpha'' + \theta_1 (q_\alpha' - q_\alpha'')$ and $q \in \Omega'$ in accordance with Lemma 1.

From (5.1) and (5.5) we have

$$q_{\alpha+1}' = q_\alpha' - G_\alpha'^{-1} G \cdot (q_\alpha' - q_\alpha'') = q_\alpha'' + (I - G_\alpha'^{-1} G)(q_\alpha' - q_\alpha''). \tag{5.6}$$

The inequality

$$\| q_\alpha'' - q_0 \| \leq \delta_0 \| q_\alpha' - q_0 \| , \tag{5.7}$$

is proved entirely in the same way as the inequality (2.16) (taking into account the fact that $f(q_\alpha'', q_\alpha') = 0$). Consequently, we obtain

$$\| q_{\alpha+1}' - q_0 \| = \| q_\alpha'' - q_0 + (I - G_\alpha'^{-1} G)(q_\alpha' - q_0 + q_0 - q_\alpha'') \| \leq$$

$$\leq \| q_\alpha'' - q_0 \| + \| I - G_\alpha'^{-1} G \| (\| q_\alpha' - q_0 \| + \| q_\alpha'' - q_0 \|) \leq [\delta_0 + \alpha_0 (1 + \delta_0)] \| q_\alpha' - q_0 \| .$$

from (5.4), (5.6) and (5.7). Let $\beta = \delta_0 + \alpha_0 (1 + \delta_0)$; then by (5.4), $\beta < 1$. Hence, for any terms of the sequence $\{q_\alpha'\}$, the inequality (5.3) is valid, Q. E. D.

/23

It is easy to see that in linear problems (i. e., if the matrix A in (1.4), and , therefore, according to (2.7), the matrix G also, does not depend on q, the procedures (1.6) and (5.1) coincide.

We shall prove one more theorem about the convergence of the procedure (5.1).

Theorem 4. Suppose that the conditions of Theorem 2 and the following conditions also are satisfied:

1) all the terms of the sequence $\{q'_K\}$ constructed according to (5.1) belong to the neighborhood Ω' defined in (2.6) and (2.1);

2) for each term q'_K of the sequence $\{q'_K\}$ and for the vector

$$q''_K = q'_K - (G'_K + H'_K)^{-1} f'_K , \tag{5.8}$$

where G'_K and f'_K are determined by (5.2) and $H'_K = H(q'_K, q'_K)$, the inequality

$$\|q''_K - q_*\| \leq \alpha_K \|q'_K - q_*\| , \quad 0 < \alpha_K \leq \frac{1-\delta_K}{1+\delta_K} - \varepsilon , \tag{5.9}$$

is satisfied, where q_* is the limit of the procedure (1.6) or (4.3), $\delta_K = \delta(\|q'_K - q_*\|, \|q'_K - q_*\|)$ (the function $\delta(x,y)$ is defined in Section 3), $\varepsilon > 0$ is an arbitrarily small quantity, the same for all K = 0, I.

Then the procedure (5.1) converges to q_*.

Proof:

From (5.8)

$$f'_K = (G'_K + H'_K) \cdot (q'_K - q''_K). \tag{5.10}$$

From (5.1) together with (5.10) we obtain

$$q'_{K+1} = q'_K - G'^{-1}_K (G'_K + H'_K)(q'_K - q''_K) = q''_K - G'^{-1}_K H'_K (q'_K - q''_K). \tag{5.11}$$

Hence from (3.1) and (5.9)

$$\begin{aligned} \|q'_{K+1} - q_*\| &= \|q''_K - q_* - G'^{-1}_K H'_K (q'_K - q_* + q_* - q''_K)\| \leq \\ &\leq \|q''_K - q_*\| + \|G'^{-1}_K H'_K\| \cdot (\|q'_K - q_*\| + \|q''_K - q_*\|) \leq [\alpha_K + \delta_K (1+\alpha_K)] \|q'_K - q_*\|. \end{aligned} \tag{5.12}$$

Let $\beta = \sup_{\alpha}[\alpha_\alpha + \delta_\alpha(1+\alpha_\alpha)]$; by (5.9), $\beta \leq 1-\varepsilon < 1$. Hence, for any terms of the sequence $\{q'_k\}$ the inequality (5.3) holds, Q. E. D.

Note 1. If the center of the neighborhoods Ω and Ω' is the vector q_0 (i.e., in (2.1), $q_* = q_0$), then to satisfy the condition 1) of Theorem 4, it is sufficient to satisfy the inequality

$$\|q_0 - q_*\| \leq \frac{\vartheta\omega}{2}. \tag{5.13}$$

where ϑ and ω are determined by (2.1) and (2.2). In particular, when ϑ = 2 condition 1) is always satisfied, since according to Theorem, $q_* \in \Omega$.

To prove this assertion, set K = 0 in (5.11). Since $q'_0 = q_0 \in \Omega'$, then for G'_0 and H'_0 conditions 3) and 4) of Theorem 1 and (3.1) and (3.2), and for q''_0 (determined by (5.8)), condition 2) of Theorem 4 are satisfied. Hence, for q'_1 the inequality (5.12) is satisfied. Taking into consideration the inequality $\alpha_k + \delta_k(1+\alpha_k) < 1$ $(K = 0, 1, \ldots)$ we obtain

$$\|q'_1 - q_0\| \leq \|q'_1 - q_*\| + \|q_0 - q_*\| < 2\|q_0 - q_*\| \leq \vartheta\omega .$$

from (5.12) and (5.13). Consequently, $q'_1 \in \Omega'$. By analogous reasoning for K = 1,2....it follows that all $q'_k \in \Omega'$, Q. E. D.

Note 2. Theorem 4 is also valid of instead of conditions 1) and 2) of Theorem 1, only the existence of a solution $q_* \in \Omega$ (and even $q_* \in \Omega'$) of equation (4.1) is required. However, it must be kept in mind that, in this case, the inequalities (3.13) cannot be satisfied and nothing can be said about the closeness of the solution q_* and $\hat{q}_H$ and q_H.

Note that the results of Theorems 3 and 4 are applicable to the general case of the solution for a system of equations represented in the form (4.1).

6. The properties of the estimate obtained.

Let us consider the properties of the estimate q_* obtained as the limit of the procedure (1.6) or as the solution of equation (4.1). Here we shall assume that in all the cases considered below the conditions of Theorems 1 and 2 are satisfied.

1. The algorithm for estimation with whose help the estimate q_* is computed, is single-valued and unbiased (i.e., the estimate q_* is unique and in the absence of measurement errors it coincides with q_κ [1]).

The singlevaluedness of the algorithm was proved in Theorem I, and the absence of bias follows from the lack of bias of the estimate $\hat{q}_\kappa$ [1]: In the absence of measurement errors in (2.10), Δ = 0 and consequently, $q_* = \hat{q}_\kappa = q_\kappa$.

2. If the functions being measured are linearly dependent on q, and the measurement errors are not biased, then the estimate q_* is unbiased when the weight matrix is fixed (i.e., if the matrix $P_* = P(q_*)$ is fixed, then $M(q_*) = q_\kappa$).

This property follows from the absence of bias in the MLS estimate with a constant weight matrix in linear problems [1]. Consequently, any estimate $\hat{q}_\kappa$ determined by (1.6) is unbiased and a limit estimate is also not biased.

3. If the estimate $\hat{q}_M$ is consistent and effective [1], then the estimate q_* is also consistent and converges with respect to probability to the effective estimate.

This property follows from (2.10), since if the estimate $\hat{q}_M$ is consistent, then the quantity Δ in (2.10) converges with respect to probability to zero.

It is easy to see that all the properties considered are also valid for the estimates $\hat{q}_\kappa$ determined by (1.6).

The limit properties of the estimate given below are proved in [2] for the case of a diagonal weight matrix and do not require that condition 5) of Theorem 1 be satisfied (however, the existence of the derivatives with respect to q of the measured functions up to and including the third order and of the weight matrix elements is required):

1) the procedure (1.6) converges with respect to probability;

2) the estimate q_* obtained as the limit of the procedure (1.6) has an asymptoticly normal distribution;

3) the estimate q_* is consistent in the sense of convergence almost certainly.

Moreover, in [2] it is shown that the estimate obtained by minimization with respect to q of the function $\Psi(q,q)$ defined by (1.3) is not consistent even in the case of a diagonal weight matrix. In the first approximation this estimate coincides with the estimate obtained by the method of maximum probability [1] with normally distributed measurement errors. In fact, in view of (1.3) and the relation $P(q)=D^{-1}(q)$, this method reduces to finding the maximum of the function

$$L(q)=\det[P(q)]\exp[-\varphi(q,q)]. \tag{6.1}$$

If we neglect the fact that the determinant of the weight matrix depends on q, then maximization in (6.1) reduces to minimizing the function $\Psi(q,q)$.

7. An illustrative example

Suppose that q is a scalar (i.e., M = 1) and the functions to be measured are linear with respect to q. In this case we may assume that the quantity q itself is measured, i.e., the functions to be measured are reduced to the form

$$d_i = q \quad (i=\overline{1,n}). \tag{7.1}$$

by elementary transformations. It will be assumed that the measurement errors are not biased and not correlated among themselves and their mean square deviations (m.s.d.) are linearly dependent on $\mathring{q}$

(cf. Section 1):

$$\sigma_i = \sigma_{oi} + \sigma_{1i}(\tilde{q} - q_u), \quad |\sigma_{1i}| \le \varepsilon \ll 1 \quad (i = \overline{1,n}). \tag{7.2}$$

The smallness of the factor σ_{1i} in (7.2) is illustrated by the following example: if distance is being measured, then as distance increases, let us say, per 1 km the error in the measurements increases substantially less than by 1 km. It is to be expected that in this example ε does not exceed a magnitude on the order of between 10^{-4} and 10^{-6}.

From (1.4) and (7.1) we have:

$$f(q,\tilde{q}) = \sum \frac{\xi_i - (q - q_u)}{\sigma_i^2}, \tag{7.3}$$

where $\xi_i = \tilde{q}_i - q_u$ are the measurement errors. From (2.7), (7.3) and (7.2) we obtain

$$G(\tilde{q}) = -\sum_{i=1}^{n} \frac{1}{\sigma_i^2}, \quad H(q,\tilde{q}) = -2\sum_{i=1}^{n}[\xi_i - (q - q_u)]\frac{\sigma_{1i}}{\sigma_i^3}. \tag{7.4}$$

Obviously, if for all $\tilde{q} \in \Omega'$ $\sigma_i \ne 0$ $(i = \overline{1,n})$, then conditions 3) and 4) of Theorem 1 are satisfied. The msd of the errors ξ_i equal σ_{oi}; we shall assume that

$$|\xi_i| \le 3\sigma_{oi} \quad (i = \overline{1,n}). \tag{7.5}$$

Then follows from (7.4), (7.2), (7.5) /28

$$|G^{-1}H| \le 2\sum_{i=1}^{n} \frac{|\sigma_{1i}|}{\sigma_i^3}(3\sigma_{oi} + |q - q_u|) \Big/ \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \le$$
$$\le 2\left(3\varepsilon \max_{1 \le i \le n} \frac{\sigma_{oi}}{\sigma_i} + |q - q_u| \max \frac{|\sigma_{1i}|}{\sigma_i}\right) \tag{7.6}$$

Assume Ω and Ω' as neighborhoods of the vector q_u (i.e., in (2.1), $q_o = q_u$). Note that in (2.1), $\nu = 1$ (i.e., $\Omega = \Omega'$), since for M = 1 all three norms in (2.2) coincide. It is easy to see than when $|q - q_u| = |\tilde{q} - q_u| = \omega$, the right side of the inequality (7.6) assumes the greatest value in Ω. Set the quantity δ_0 equal to this maximum; we obtain from (7.6) and (7.2)

$$|G^{-1}H| \leq 2\left(3\varepsilon \max_{1\leq i\leq n} \frac{\sigma_{0i}}{\sigma_{0i} - |\sigma_{1i}|\omega} + \max_{1\leq i\leq n} \frac{|\sigma_{1i}|\omega}{\sigma_{0i} - |\sigma_{1i}|\omega}\right) = 2\,\frac{3\varepsilon s + \omega}{s - \omega} = \delta_0\,, \tag{7.7}$$

where

$$s = \min_{1\leq i\leq n} \frac{\sigma_{0i}}{|\sigma_{1i}|} \tag{7.8}$$

From (7.7)

$$\omega = \frac{\delta_0 - 6\varepsilon}{2 + \delta_0}\, s\ . \tag{7.9}$$

It is evident from (7.7) that if $\omega < \frac{s}{3}$ and $\varepsilon \ll 1$, then condition 5) of Theorem 1 is satisfied (i.e., $\delta_0 < 1$). Here, according to (7.2) and (7.8)

$$\max_{q\in\Omega} \sigma_i \Big/ \min_{q\in\Omega} \sigma_i \leq \frac{s+\omega}{s-\omega} < 2\,,$$

i.e., the msd of the errors in the measurements can vary less than by a factor of 2.

Let $q_0 \in \Omega$ and $\varepsilon \ll \Omega$ (the latter condition is satisfied for sufficiently large δ_0 according to (7.2)). Let us show that in this case conditions 1) and 2) of Theorem 1 are satisfied. /29

We shall assume that

$$\Delta = |\hat{q}_{\nu} - q_{\nu}| \leq 3\sqrt{D(q_{\nu})}\ . \tag{7.10}$$

The deviation occurring in (7.10) with respect to the error in determining the parameter $\hat{q}_{\nu}$ by MLS is calculated by means of the formula [1]

$$D(\hat{q}_{\nu}) = |G^{-1}(q_{\nu})|\ . \tag{7.11}$$

From (7.4), (7.2) and (7.8) we obtain

$$|G(q_{*})| = \sum_{i=1}^{n} \frac{1}{\sigma_{0i}^{2}} \geq \frac{1}{\varepsilon^{2}} \sum_{i=1}^{n} \left(\frac{\sigma_{1i}}{\sigma_{0i}}\right)^{2} \geq \frac{1}{\varepsilon^{2}} \max_{1 \leq i \leq n} \left(\frac{\sigma_{1i}}{\sigma_{0i}}\right)^{2} = \frac{1}{(\varepsilon S)^{2}} \tag{7.12}$$

and from (7.10)-(7.12)

$$\Delta \leq 3\varepsilon S \ . \tag{7.13}$$

Therefore, when $\varepsilon \ll \delta_0$, we have $\Delta \ll \omega$ and $\hat{q}_{*} \in \Omega$, according to (7.9) and (7.13). Since $q_{*}, q_{0} \in \Omega$, condition 1) of Theorem 1 is satisfied.

We shall show by induction with respect to K that condition 2) of Theorem 1 is satisfied when $q_0 \in \Omega$ and $\varepsilon \ll \delta_0$. Since $\hat{q}_0 = q_0$, it follows that $\hat{q}_0 \in \Omega$. Assume that for some K = 0,1,..., $\hat{q}_K \in \Omega$. Since the problem under consideration is linear, then, as remarked in Section 5, the procedures (1.6) and (5.1) coincide and, according to (5.1), (7.3) and (7.4)

$$\hat{q}_{K+1} - q_{*} = \hat{q}_{K} - q_{*} - G^{-1}(\hat{q}_K)\, f(\hat{q}_K, \hat{q}_K) = \sum_{i=1}^{n} \frac{\xi_i}{\sigma_i^{(K)2}} \Big/ G(\hat{q}_K), \tag{7.14}$$

where $\sigma_i^{(K)} = \sigma_{0i} + \sigma_{1i}(\hat{q}_K - q_{*})$. We shall assume that

$$|\hat{q}_{K+1} - q_{*}| \leq 3\sqrt{D(\hat{q}_{K+1})} \ ; \tag{7.15}$$

From (7.14) follows:

$$D(\hat{q}_{K+1}) = M\left[(\hat{q}_{K+1} - q)^2\right] = \sum_{i=1}^{n} \frac{1}{\sigma_i^{(K)2}} \left(\frac{\sigma_{0i}}{\sigma_i^{(K)}}\right)^{2} \Big/ G^{2}(\hat{q}_K). \tag{7.16}$$

/ 30

We obtain from (7.2), (7.8), (7.9) and (7.12)

$$\begin{aligned} &\sum_{i=1}^{n} \frac{1}{\sigma_i^{(K)2}} \left(\frac{\sigma_{0i}}{\sigma_i^{(K)}}\right)^{2} \leq \frac{|G(\hat{q}_K)|}{\left(1 - \frac{\omega}{S}\right)^{2}} = \left(\frac{2+\delta_0}{2+6\varepsilon}\right)^{2} |G(\hat{q}_K)| < \left(1 + \frac{\delta_0}{2}\right)^{2} |G(\hat{q}_K)|, \\ &|G(\hat{q}_K)| = \sum_{i=1}^{n} \frac{1}{\sigma_i^{(K)2}} = \sum_{i=1}^{n} \frac{1}{\sigma_{0i}^{2}} \left(\frac{\sigma_{0i}}{\sigma_i^{(K)}}\right)^{2} \geq \frac{|G(\ \ )|}{\left(1 + \frac{\omega}{S}\right)^{2}} = \\ &= \left(\frac{2+\delta_0}{2+2\delta_0 - 6\varepsilon}\right)^{2} |G(q_{*})| > \left(\frac{1 + \frac{\delta_0}{2}}{1+\delta_0} \cdot \frac{1}{\varepsilon S}\right)^{2} . \end{aligned} \tag{7.17}$$

Then from (7.15)-(7.17) we find that

$$|\hat{q}_{K+1} - q_{*}| < 3(1+\delta_o)\varepsilon s \, ,$$

(7.18)

and when $\varepsilon \leq \frac{\delta_o}{3(4+3\delta_o+\delta_o^2)}$ so that according to (7.9), $\hat{q}_{K+1} \in \Omega$ Q.E.D.

Let us now find the quantity u_* which occurs in the inequality (3.13). Set

$$\delta(x,y) = 2 \cdot \frac{3\varepsilon s + x}{s - y} \quad , \tag{7.19}$$

where X and Y are determined by (3.3) when $q_1 = q_2 = q_M$. According to (7.6), (7.2) and (7.8), $|G^{-1}H| \leq \delta(x,y)$. Since the function (7.19) is monotonically increasing, it follows from (3.5) and (3.6) that $\omega(u) = \delta(u + \Delta, u + \Delta)$, and the equation (3.9) assumes the form

$$u = 2\Delta \frac{3\varepsilon s + \Delta + u}{s - 6\varepsilon s - 3u - 3\Delta}$$

and u_* is the greatest solution of this equation when $\omega(u) < \delta_o$. Taking (7.2) into account we obtain

$$u_* \approx \frac{2\Delta}{s}(3\varepsilon s + \Delta) \leq 12\varepsilon\Delta \leq 36\varepsilon^2 s \, .$$

(7.20)

with accuracy with respect to small ε up to the second order. It is clear from (3.13) and (7.20) that when ε is sufficiently small, the quantity q_* practically coincides with $\hat{q}_M$, and the estimate of /31 closeness q_* and q_M for the quantity δ_o close to unity is better by almost a factor of two than the analogous estimate for K = o,1,...) according to (3.13) and (7.13).

In conclusion, the author expresses appreciation to G. A. Mersov for valuable comments and advice regarding the present paper.

REFERENCES

1. ---Yacberg, P. Ye., Opredeleniye dvizheniya po rezul'tatam izmeriniy (Determination of Motion from the Results of Measurements). Moscow, "Nauka", 1976.

2.

3. Kantorovich, L. V., Akilov, G. P., Funktsional'nyy analiz (Functional Analysis). Moscow, "Nauka", 1977.

4. Bakhvalov, N. S., Chislennyye metody (Numerical Methods), v. 1, Moscow, "Nauka", 1973.

5. Berezin, I. S., Didkov, N. P., Metody vychisleniy (Computational Methods), v. 1. Moscow, "Nauka", 1966.