объединенный
институт
ядерных
исследований
дубна

E5-85-50

K.Mühle, Š.Šujan

ON INVERSE PROBLEMS OF EXPERIMENTAL PHYSICS: FINITE-DIMENSIONAL LINEAR THEORY

Submitted to "Inverse Problems"

1985

INTRODUCTION

In this paper we consider inverse problems in a very wide sense as problems of reconstructing the inner state of a physical object based on its outer (observable) properties. Problems of both deterministic and stochastic nature fall within that setup, and we shall focus ourselves upon the statistical ones.

At the heart of our exposition there is the notion of a (parametric) regression experiment. Thus, the relations between the inner state and the observable properties are supposed to be of known analytic form, but depending on an unknown parameter vector (which is supposed on its own to characterize the state of the object considered). In this way, the inverse problem may be reformulated either as estimation of the unknown "true" parameter vector (estimation aposteriori) or, as optimum experimental design (estimation apriori).

Since inverse problems in experimental physics are quite frequently ill-posed [1], one cannot expect reasonable results without introducing somehow prior information. Prior information can concern the model of experiment itself (e.g., introduction of some prior probability distribution on the set of estimated parameters) as well as the estimates themselves (e.g., assignation of weights to the components of the parameter vector, employment of loss functions, etc.).

Choosing the basic model, introducing the prior information, determining the criteria of "quality" of solutions we can reduce the inverse problem to an associated estimation problem. In this paper we attempt to give a brief account of finite-dimensional linear theory [2,3]. We note that finite-dimensional models are the most common statistical models, but they can arise also when aiming at calculable results within general infinite-dimensional theory (e.g., linear estimation and regularization problems in Banach spaces [4]).

Now, the basic problem can be described as that of solving a linear system $A\vec{x} = \vec{y}$, where $\vec{y}$ is the vector of observed values (= experimental data). The following features are met typically in realistic problems:

(a) the dimensions of $\vec{x}$ and $\vec{y}$ are different, and A does not have full rank (= the problem is ill-posed),

(b) measurement errors when observing different components of $\vec{y}$ are statistically dependent, and

(c) the joint distribution of the error vector is unknown. While (c) and partly (b) exclude methods like maximum likelihood, (a) makes the problem unsolvable in a unique manner. Fortunately, we can achieve uniqueness provided some rather natural constraints will be imposed.

1. BASIC FEATURES OF LINEAR SYSTEMS

In all considerations, if $\vec{x} \in R^k$ is a row vector ($\vec{x} = \vec{x}_{1,k}$, say), then $\vec{x}' \in R^k$ is the corresponding column vector ($\vec{x}' = \vec{x}_{k,1}$). If $A_{m,n}$ is a matrix, then $A'_{n,m}$ is its transpose. Dimensions of vectors and matrices will be written occasionally to make the formulae more transparent. Let

$$A = A_{M,N} \ (N \leq M), \quad R(A) = r \leq N \tag{1.1}$$

(here $R(A)$ stands for the rank of A). Consider the linear system

$$A_{M,N}\vec{x}'_{N,1} = \vec{y}'_{M,1}. \tag{1.2}$$

The system (1.2) is called consistent if there exists at least one its solution $\vec{x}$. Necessary and sufficient conditions for consistency can be expressed in terms of g-inverses to A. The class $\mathcal{A}^-$ of all g-inverses to A ($\mathcal{A}^- \neq \emptyset$ for any matrix A) consists of all matrices $G = G_{N,M}$ such that

$$AGA = A. \tag{1.3}$$

Then (1.2) is consistent if and only if there exists a matrix $A^- \in \mathcal{A}^-$ such that

$$AA^-\vec{y} = \vec{y} \tag{1.4}$$

(and in this case (1.4) is valid for all $A^- \in \mathcal{A}^-$). Note that when $R(A) = N$ (full rank), it is easy to find an element $A^- \in \mathcal{A}^-$. Namely, take

$$A^- = (A'A)^{-1}A'. \tag{1.5}$$

Then $A^-A = I_{N,N}$, where I is the diagonal unit matrix. In the general case, the solutions to (1.2) form a parametric family ($A^- \in \mathcal{A}^-$ a fixed element):

$$\{A^-\vec{y} + (I - A^-A)\vec{z} : \vec{z} \in R^N\}. \tag{1.6}$$

If $R(A) = N$ then $I - A^-A = \vec{0}$. Consequently, up to the choice of $A^- \in \mathcal{A}^-$ there is a unique solution $A^-\vec{y}$ of (1.2). Thus, $AA^-\vec{y} = A\vec{x} = \vec{y}$, i.e., any system of full rank is consistent.

The basic role will be played by the linear space $\mathfrak{M}(C)$ spanned by the columns of C. If $C = A'$ then the following obvious criterion takes place:

$$\vec{p} \in \mathfrak{M}(A') \iff \exists\, L : A'L = \vec{p}. \tag{1.7}$$

2. UNIVERSAL LINEAR MODEL

Let $\vec{n}_{1,M}$ be a random perturbation of the right-hand side of (1.2):

$$A_{M,N}\vec{x}'_{N,1} = \vec{y}'_{M,1} + \vec{n}'_{M,1}. \tag{2.1}$$

Let $\tilde{\vec{y}} = \vec{y} + \vec{n}$. If $E\vec{n} = \vec{0}$ (E = expectation), then $E\tilde{\vec{y}}' = A\vec{x}'$. If $E\vec{n} \neq \vec{0}$ then $\vec{n} = E\vec{n} + \tilde{\vec{n}}$, where $E\tilde{\vec{n}} = \vec{0}$. Then $\tilde{Y} = A\vec{x} + \tilde{\vec{n}}$. Consequently, we may suppose that the perturbed system assumes on the form

$$\left.\begin{aligned} E\tilde{\vec{y}}'_{M,1} &= A_{M,N}\vec{x}'_{N,1}, \\ E\vec{n}'_{M,1} &= \vec{0}_{M,1}. \end{aligned}\right\} \tag{2.2}$$

In order the problem be fully determined it is necessary to define the correlation structure of the perturbation. We shall consider our problem under the following assumptions (we suppose throughout that the correlation matrix $R^{\tilde{y}} = R^n$ does not depend on $\vec{x}$):

$$\text{(I)} \quad R^{\tilde{y}}_{M,M} = \sigma^2 V_{M,M}, \tag{2.3}$$

where σ^2 and V are known,

$$\text{(II)} \quad R^{\tilde{y}}_{M,M} = \Sigma_{i=1}^{s} \sigma_i^2 V^i_{M,M}, \tag{2.4}$$

where σ_i^2, V^i are known,
(III) the case (I) with σ^2 unknown,
(IV) the case (III) with σ_i^2 unknown, $1 \leq i \leq s$.
The models (I) and (III) will be denoted as

$$[\tilde{\vec{y}}, A\vec{x}, \sigma^2 V], \tag{2.5}$$

and the models (II) and (IV) as

$$[\tilde{\vec{y}}, A\vec{x}, \Sigma_{i=1}^{s} \sigma_i^2 V^i], \tag{2.6}$$

and called a universal linear model [3] (sometimes it is called a generalized Gauss-Markov model). Note that the existence of V^{-1} (of $(V^{i})^{-1}$) is not required, the model allows for correlated and arbitrarily distributed measurement errors, and furthermore, the model includes also linear constraints on $\vec{x}$. To see this, suppose the constraints are given in the form $R\vec{x}=\vec{c}$. Let us put $\vec{y}^*=(\vec{y}, \vec{c})'$, $A^*=(A,R)'$, In case (2.3) we get

$$\sigma^2 V^* = \mathfrak{D}^2(\vec{y}^*) = \begin{pmatrix} V & \vec{0} \\ \vec{0} & \vec{0} \end{pmatrix} \sigma^2,$$

and thus the extended model reads $[\vec{y}^*, A^*\vec{x}, \sigma^2 V^*]$(cf.(2.5)).

Essentially, we shall seek for efficient estimates $\hat{\vec{x}}$. This leads naturally to the requirement of unbiased estimates, for which

$$E\,\hat{\vec{x}} = \vec{x}. \qquad (2.7)$$

In conclusion of this section observe the following. If V is a singular matrix, then we arrive at natural constraints upon the vector $\vec{y}$ these constraints are necessary to verify in order to exclude obvious errors in model formulation [3]. In more detail, let us consider the model (2.5). Let α be a vector such that $\alpha' A=\vec{0}$ and $\alpha' V=\vec{0}$. Then $E(\alpha'\vec{y}) = \alpha' A\vec{x} = \vec{0}$ and $E(\alpha'\vec{y})^2 = \vec{0}$ (since this latter expectation involves $\alpha' V = \vec{0}$). Hence $\mathrm{Prob}[\alpha'\vec{y} = \vec{0}] = 1$, i.e., $\alpha \perp \vec{y}$ with probability one. But since $\alpha \perp A$ and $\alpha \perp V$, we see that $\vec{y} \in \mathfrak{M}(A \,\vdots\, V)$, i.e., our system is consistent with probability one.

3. SOLUTION IN CASE (I)

First of all we describe general results for the case (I). In the subsequent section we shall present several methods of regularization of the estimation problem with A being of non-full rank.

Recall [2,3] that within the universal model a functional $f(\vec{x}) = \vec{p}'\vec{x} : R^n \to R^1$ ($\vec{p} \in R^n$) is unbiasedly estimable if and only if $\vec{p} \in \mathfrak{M}(A')$ (cf. (1.7)). If $\vec{p} \in \mathfrak{M}(A')$ then the (linear) unbiased estimate of $\vec{p}'\vec{x}$ is of the form

$$L'\vec{y} = [(A')^- \vec{p}]'\vec{y}, \qquad (3.1)$$

and its generalized variance is determined by

$$\mathfrak{D}^2(L'\vec{y}) = \sigma^2 L' V L. \qquad (3.2)$$

Let N be a positive semidefinite matrix. The N-seminorm of a vector $\vec{z}$ is defined by

$$\|\vec{z}\|_N = (\vec{z}' N \vec{z})^{1/2}. \qquad (3.3)$$

We let $\mathcal{A}^-_{m(N)}$ denote the class of all g-inverses to A having minimal N-seminorm. By definition, $G_{N,M} \in \mathcal{A}^-_{m(N)}$ if $A = AGA$ (i.e., $G \in \mathcal{A}^-$) and

$$(\forall \vec{y} \in \mathfrak{M}(A))\ (\forall \vec{x} : A\vec{x} = \vec{y})\ (G\vec{y})' N G\vec{y} < \vec{x}' N \vec{x}. \qquad (3.4)$$

Then $\mathfrak{D}^2(L'\vec{y})$ will be minimal if we take as $(A')^-$ in (3.1) the matrix

$$(A')^- = (A')^-_{m(V)}\ . \qquad (3.5)$$

A transparent geometrical explanation of the choice (3.5) can be given in case when V is a positive definite matrix. We introduce another class, $\mathcal{A}^-_{\ell(M)}$, of g-inverses as follows: $G \in \mathcal{A}^-_{\ell(M)}$ if for all $\vec{y} \in R^N$ and all $\vec{x} \in R^N$

$$\|AGA\vec{y} - \vec{y}\|_M < \|A\vec{x} - \vec{y}\|_M\ . \qquad (3.6)$$

For a positive definite matrix V we have [2,3]

$$[(\mathcal{A}')^-_{m(V)}]' = \mathcal{A}^-_{\ell(V^{-1})}\ . \qquad (3.7)$$

Consequently,

$$\mathfrak{D}^2(L'\vec{y}) = \sigma^2 \|L\|^2_V \quad \text{(cf. (3.3))}. \qquad (3.8)$$

Formula (3.8) is in a complete agreement with the formula for the volume of minimal ellipsoid of concentration for efficient estimates in regular models [2,3]. Consequently, the estimates of the type ((3.1), (3.5)) will be called efficient, too. Thus, in case (I) an efficient linear estimate of $f(\vec{x}) = \vec{p}'\vec{x}$ ($\vec{p} \in \mathfrak{M}(A')$) assumes on the form

$$\widehat{f(\vec{x})} = \vec{p}'[(A')^-_{m(V)}]'\vec{y}. \qquad (3.9)$$

As the matrix in (3.9) we may take, e.g., $(V+AA')^- A[A'(V+AA')^- A]^-$, where C^- designates an arbitrary element of the class $\mathcal{C}^-$.

Preceding considerations allow for a straightforward generalization. Let $\vec{p}_i \in \mathfrak{M}(A')$, $1 \le i \le r$, and let $P = (\vec{p}_1, \ldots, \vec{p}_r)'$. Then an efficient estimate of the vector $P\vec{x}$ is of the form

$$\widehat{P\vec{x}} = P[(A')^-_{m(V)}]'\vec{y}. \qquad (3.10)$$

At the first glance it may seem there is a large ambiguity in the choice of an efficient estimate. However, it is possible to prove that an efficient estimate is unique with probability one. This important result implies, in particular, that the estimates (3.9) and (3.10) do not depend on the particular choice of the matrix $(A')^-_{m(V)} \in (\mathcal{A}')^-_{m(V)}$. This crucial fact was the fundament of the theory of parallel computations in linear statistical theory [3].

Of course, having obtained the estimates (3.9) and (3.10) we are not yet done. In order one can assign to these estimates statistical reliability it turns out necessary to determine their statistical properties. Following the general philosophy of linear statistical methods we restrict ourselves to properties of the first and second orders. In case of estimate (3.10) we have

$$R^{P\hat{\vec{x}}} = \sigma^2 [P[(A')^-_{m(V)}]' V (A')^-_{m(V)}] P' = \sigma^2 P\{[A'(V + AA')^- A]^- - I\} P'. \quad (3.11)$$

This formula can be employed for two purposes. First of all, it makes possible to determine confidence domains for the estimate (3.10). At the same time, it is possible to use it for testing the correctness of numerical computations. Indeed, we may take some matrix $(A')^-_{m(V)}$ and determine $\sigma^2[(A')^-_{m(V)}]' V (A')^-_{m(V)}$. Then we may compare the result with $\sigma^2\{[A'(V + AA')^- A]^- - I\}$.

4. THE PROBLEM OF ILL-POSEDNESS

Suppose for a moment that A is of full rank: $R(A) = N$. Then the unit vectors $\vec{p}_i = (0,\dots,0,1,0,\dots,0) \in \mathfrak{M}(A')$ and, consequently, the matrix $P = I$ satisfies the conditions under which we were able to get the estimate (3.10). Since $P\vec{x} = \vec{x}$, we get an efficient estimate of the vector $\vec{x}$ itself:

$$\hat{\vec{x}} = [(A')^-_{m(V)}]' \vec{y}. \quad (4.1)$$

Nevertheless, even when V is invertible, we can meet serious difficulties when calculating the estimate (4.1). For the sake of simplicity let $V = I$ (noncorrelated errors). Then the estimation problem reduces to that one of solving the system of normal equations:

$$A'A\hat{\vec{x}} = A'\vec{y}. \quad (4.2)$$

As is well-known, under our hypotheses the solution of (4.2), $\hat{\vec{x}} = (A'A)^{-1}A'\vec{y}$, minimizes the form

$$\Phi(\vec{u}) = (\vec{y} - A\vec{u})'(\vec{y} - A\vec{u}). \quad (4.3)$$

However, if the matrix $A'A$ is not well-conditioned then the expectation $E\|\hat{\vec{x}} - \vec{x}\|^2 = \sigma^2 \sum_{i=1}^{k} \lambda_i^{-1}$ ($\{\lambda_1, \dots, \lambda_k\}$ is the spectrum of $A'A$) can assume on very large values. Consequently, even the efficient estimate $\hat{\vec{x}}$ will be the best in the sense of minimizing the quadratic form (4.3) within the class of all linear unbiased estimates, it will result in a bad estimate of $\vec{x}$ (thus, it will be unacceptable also from the statistical point of view, in addition to known numerical problems with calculating inverses of such bad behaved matrices).

That is why it appears reasonable not to dwell on unbiased estimates and employ an estimate, for which

$$\left.\begin{aligned} &\Phi(\vec{t}) = \Phi_{MIN} + c\ (c > 0), \\ &E\|\vec{t}(\vec{y}) - \vec{x}\|^2 < E\|\hat{\vec{x}} - \vec{x}\|^2. \end{aligned}\right\} \quad (4.4)$$

Following [5] we get a one-parameter family of ridge estimates

$$\hat{\vec{x}}(k) = (A'A + kI)^{-1} A'\vec{y}, \quad k > 0. \quad (4.5)$$

Analogous considerations yield the class of contracting estimates

$$\hat{\vec{x}}(k) = (1 + k)^{-1} (A'A)^{-1} A'\vec{y}, \quad k > 0. \quad (4.6)$$

Modifying Tikhonov's regularization method [1] to finite-dimensional spaces we can define the regularized solution to the problem of determining $\vec{x}$ as a vector $\hat{\vec{x}}(k)$ that minimizes the functional

$$\Phi_k(\vec{x}) = \|A\vec{x} - \vec{y}\|^2 + k\|\vec{x} - \vec{x}^*\|^2, \quad (4.7)$$

where $\vec{x}^*$ is an arbitrary "centering" vector (expressing the prior information about $\vec{x}$), and k is the regularization parameter [1,7]. If the norms in (4.7) are the usual Euclidean ones, we get

$$\hat{\vec{x}}^{(I)}(k) = (A'A + kI)^{-1}(A'\vec{y} + k\vec{x}^*),$$

and the choice $\vec{x}^* = \vec{0}$ leads to the class (4.5). If $\|\vec{x}\|^2 = \vec{x}'A'A\vec{x}$ then we get the class

$$\hat{\vec{x}}^{(II)}(k) = (k + 1)^{-1} (A'A)^{-1} A'\vec{y} + k(k + 1)^{-1}\vec{x}^*.$$

Again, when $\vec{x}^* = \vec{0}$, we get (4.6). However, an application of regularization methods in more general situations (such as within the universal linear model) is by far not so simple and transparent [6,7]. Consequently, it appears more advantageous to employ the above developed algebraic approach. To this end, consider the model $[\vec{y}, A\vec{x}, \sigma^2 V]$ with $\vec{p} \notin \mathcal{M}(A')$. The bias of a linear estimate $L'\vec{y}$ for $f(\vec{x}) = \vec{p}'\vec{x}$ is defined to be the function

$$b_{\vec{p}}(\vec{x}) := |E(L'\vec{y}) - \vec{p}'\vec{x}| = |(A'L - \vec{p})'\vec{x}|. \qquad (4.8)$$

Since the vector $\vec{x}$ is unknown, the concrete value $b_{\vec{p}}(\vec{x})$ for a given estimate $L'\vec{y}$ hardly can say something about its quality. If it is possible to introduce prior information with the aid of some metric (i.e., of some positive definite matrix U), then it will be reasonable to choose $L'\vec{y}$ so that the vector $A'L - \vec{p}$ will have minimal U-seminorm (cf. (3.3)). Such an estimate will be denoted by $L_0'\vec{y}$ and called U-minimally biased. Hence

$$L_0'\vec{y} = \vec{p}'[(A')_{\ell(U)}]'\vec{y} \qquad (4.9)$$

and

$$\mathfrak{D}^2(L_0'\vec{y}) = \sigma^2 \|L_0\|_V^2 . \qquad (4.10)$$

Since

$$\sigma^2 \|L_0\|_V^2 = \sigma^2 \vec{p}'[(A')_{\ell(U)}]'V(A')_{\ell(U)}\vec{p},$$

the expression (4.10) depends upon the choice of the matrix $(A')_{\ell(U)}$. The most reasonable choice is that of the best estimate within the class (4.9), i.e., the choose $L_0'\vec{y}$, where $L_0 = (A')^-_{\ell(U)}\vec{p}$ and $(\forall\, G \in (\mathfrak{A}')_{\ell(U)})\ \|G\vec{p}\|_V \ge \|L_0\vec{p}\|_V$.

We refer the reader to [2,3] for further results in this direction and restrict ourselves to the remark that one of these best estimates is of the form $\vec{p}'\{(V+AUA')^- AU[AUA'(V+AUA')^- AU]'\vec{y}$.

5. EQUIVALENCE THEORY AND CASE (III)

This section is devoted to the following three problems:

- unbiased estimation of the unknown parameter σ^2,
- the possibility of simpler calculations of estimates (3.9) and (3.10), and
- the possibility of simultaneous efficient estimation of a functional of the parameter vector $\vec{x}$ and unbiased estimation of the unit variance σ^2.

These seemingly rather different questions can be answered in a unified manner within the frame of equivalence theory, the fundaments of which were created already by Gauss [8]. First recall [3] the explicit expression for an unbiased estimate $\hat{\sigma}^2$ in case (III):

$$\hat{\sigma}^2 = (\vec{y} - A[(A')^-_{m(V)}]'\vec{y})'V^-(\vec{y} - A[(A')_{m(V)}]'\vec{y}) \times [R(V,A) - R(A)]^{-1} = s^{-1}\vec{v}'V^-\vec{v}, \qquad (5.1)$$

where $s = R(V,A) - R(A)$,

$$\vec{v} = A\hat{\vec{x}} - \vec{y} \qquad (5.2)$$

is the vector of corrections, and $\hat{\vec{x}}$ is an estimate of the following form:

$$A\hat{\vec{x}} = A[(A')^-_{m(V)}]'\vec{y}. \qquad (5.3)$$

The meaning of the estimate $\hat{\vec{x}}$ can be explained with the help of a generalization of least squares method. An M-LSE of $\vec{x}$ is, by definition, any solution $\hat{\vec{x}}$ of generalized normal equations (cf. (4.2))

$$A'MA\hat{\vec{x}} = A'M\vec{y}. \qquad (5.4)$$

If V^{-1} exists and if $\hat{\vec{x}}$ is the usual least squares estimate (LSE) of $\vec{x}$, then $P\hat{\vec{x}}$ is an efficient estimate of the vector $P\vec{x}$ for any matrix P with $\mathcal{M}(P') \subset \mathcal{M}(A')$. If V^{-1} does not exist, then the latter assertion will remain true provided we take $\hat{\vec{x}}$ to be a $(V+A'A)^{-1}$-LSE of $\vec{x}$. From the uniqueness theorem for efficient estimates and from (3.10) it follows that

$$P\hat{\vec{x}} = P[(A')^-_{m(V)}]'\vec{y} \quad \text{a.e.}, \qquad (5.5)$$

and this justifies the formula (5.3). The practical significance of (5.5) comes from the fact that LSE's can be usually calculated much more easily (e.g., V^{-1}-LSE is but the least squares method in its classical form). However, until now we do not know anything about a suitable choice of the matrix M in (5.4). Such an information gives us the following assertion on equivalence of (A) and (B), where

(A) the system (5.4) has a solution and for any $\vec{p} \in \mathcal{M}(A')$ the vector $\vec{p}'\hat{\vec{x}}$ is an efficient estimate of $f(\vec{x}) = \vec{p}'\vec{x}$, and

(B) $$M = (V + AUA')^- + K, \qquad (5.6)$$

$$R(A'MA) = R(A'), \qquad (5.7)$$

where U and K are arbitrary matrices such that

$$\mathcal{M}(V,A) = \mathcal{M}(V + AU'A') = \mathcal{M}(V + AUA'), \qquad (5.8)$$

$$VK'A = \vec{0},\ A'KA = \vec{0}\ [9]. \qquad (5.9)$$

Different choices of K and U yield different choices of M, however, the obtained estimates coincide with probability one. This is again important from the point of view of numerical calculations.

Since in case (III) σ^2 is unknown, there appears the problem of jointly estimating $\vec{p}'\vec{x}$ and σ^2. In other words, we ask under which conditions upon the matrix M (cf. (5.6)) the expression

$$\hat{\sigma}^2 = (\vec{y} - A\hat{\vec{x}})'M(\vec{y} - A\hat{\vec{x}})/[R(V,A) - R(A)] \quad (5.10)$$

will be an unbiased estimate of σ^2? In order to answer this one can again formulate the corresponding variant of equivalence theorem [9]. Nevertheless, the concrete choice of M is still open. Let us briefly describe one unexpected method for calculation of joint estimates - the Pandora box method [2,3]. The Pandora box method is defined by

$$\begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix} = \begin{pmatrix} V & A \\ A' & \vec{0} \end{pmatrix}^{-} . \quad (5.11)$$

Surprisingly enough, the matrices C_i contain all information needed for solution of our problem. In fact, for any functional $f(\vec{x}) = \vec{p}'\vec{x}$, $(\vec{p} \in \mathfrak{M}(A'))$ the estimates

$$\vec{p}'C_2'\vec{y}, \quad \vec{p}'C_3\vec{y} \quad (5.12)$$

are efficient,

$$\mathfrak{D}^2(\vec{p}'C_2'\vec{y}) = \mathfrak{D}^2(\vec{p}'C_3\vec{y}) = \sigma^2\vec{p}'C_4\vec{p}, \quad (5.13)$$

and

$$\hat{\sigma}^2 = \vec{y}'C_1\vec{y}/\mathrm{Tr}(VC_1) \quad (5.14)$$

is an unbiased estimate of the unit variance σ^2.

6. SOLUTION IN CASES (II) AND (IV)

In order to get estimates of functionals $\vec{p}'\vec{x}$ and vectors $P\vec{x}$ in cases (II) and (IV) one can employ the same methods as in the preceding sections. The estimation problem for the vector $(\sigma_1^2, \ldots, \sigma_s^2)$ (cf. (2.4)) is not completely solved.

A quadratic form $\vec{y}'M\vec{y}$ is said to be an invariant unbiased estimate with minimal norm (following Rao, such estimates got the acronym MINQUE) of a linear functional $f(\sigma_1^2, \ldots, \sigma_s^2) = \sum_{i=1}^{s} p_i\sigma_i^2$, if M is a symmetric matrix and the following conditions are satisfied:

- invariance, i.e.,

$$(\forall \vec{a} \in R^N)\ \vec{y}'M\vec{y} = (\vec{y} + A\vec{a})'M(\vec{y} + A\vec{a}), \quad (6.1)$$

- unbiasedness, i.e.,

$$(\forall \sigma_i^2 \in (0,\infty),\ 1 \le i \le s)\ E(\vec{y}'M\vec{y}) = \sum_{i=1}^{s} p_i\sigma_i^2, \quad (6.2)$$

- minimum of the norm, i.e.,

$$\|U'MU - \Delta\|^2 = \mathrm{Tr}\{(U'MU - \Delta)(U'MU - \Delta)'\} = \mathrm{MIN}, \quad (6.3)$$

where $U = (U_1, \ldots, U_s)$, $R(U_i) = R(V^i) = n_i$, $V^i = U_iU_i'$, $i \in \{1, \ldots, s\}$, and Δ is a block-diagonal matrix with diagonal blocks $(p_i/n_i)\,I_{n_i n_i}$, $1 \le i \le s$.

If the matrix $V = V^1 + \ldots + V^s$ is regular then a matrix M satisfying all requirements formulated above can be determined from the relations

$$\left.\begin{aligned} M = \sum_{i=1}^{s} [I - V^{-1}A(A'V^{-1}A)^{-}A']V^{-1}V^iV^{-1} \times \\ \times [I - A(A'V^{-1}A)^{-}A'V^{-1}]\beta_i\ , \end{aligned}\right\} \quad (6.4)$$

where the numbers β_i are determined from

$$\mathrm{Tr}(MV^i) = p_i,\ 1 \le i \le s. \quad (6.5)$$

Unfortunately, equations (6.5) do not admit, in general, a solution. In such a case it is necessary to employ the structure of concrete estimation problem, and attempt to get satisfactory results using appropriate combinations of estimable functionals (see [3] for more details and examples).

7. CONCLUSION

We did not consider the questions related to various ways of calculation of g-inverses because (a) the appropriate way depends upon concrete structure of the matrix and (b) there is an extensive literature devoted to that topic [10]. Of course, if errors are normally distributed, we can get much stronger results to the effect that we can find exact distributions for estimates presented above [2,3], these distributions being always independent of the particular choice of a g-inverse.

In conclusion, we have seen there is a well-developed theory which deals in a unified manner with finite-dimensional linear theory of inverse problems. It was our intention to call attantion of people dealing with inverse problems in experimental physics to this fact. Of course, the linear theory of experimental design can be developed much the same way as estimation theory, again without any assumptions concerning regularity of the regression models. However, this exceeds the frame of the present paper.

REFERENCES

1. Tikhonov A.N., Arsenin V.Ya. Methods of Solution of Ill-Posed Problems (in Russian). "Nauka", Moscow, 1979.
2. Rao C.R. Linear Statistical Inference and Its Applications. J.Wiley, New York, 1973.
3. Kubaček L. Fundaments of Estimation Theory (in Slovak, English version in preparation). Veda, Bratislava, 1983.
4. Tanana V.P. Doklady Akad.Nauk SSSR, 1984, 277, pp.557-559.
5. Hoerl A.E., Kennard R.W. Technometrics, 1970, 12, pp.56-68, 69-82.
6. Mathematical Theory of Experimental Design (in Russian). "Nauka", Moscow, 1983.
7. Turchin V.F., Kozlov V.P., Malkevich M.S. Usp.Fiz.Nauk, 1970, 102, pp.345-386.
8. Gauss K.-F. Werke 4-1, Göttingen, 1809.
9. Wimmer G. Aplikace mat., 1980, 25, pp.417-422, 423-431.
10. Ben-Israel A., Greville T.N.E. Generalized Inverses: Theory and Applications. J.Wiley, New York, 1974; Pringle R.M., Rayner A.A. Generalized Inverses of Matrices and Applications to Statistics. Griffin, London, 1971.

Received by Publishing Department
on January 24, 1985.

Мюле К., Шуян Ш. E5-85-50

Об обратных задачах экспериментальной физики: конечномерная линейная теория

Описаны основы статистической теории конечномерных линейных обратных задач, причем особое внимание уделяется единому подходу к исследованию как регулярных, так и некорректных задач.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна 1985

Mühle K., Šujan Š. E5-85-50

On Inverse Problems of Experimental Physics: Finite-Dimensional Linear Theory

Statistical theory of finite-dimensional linear inverse problems is outlined, with particular emphasis on the unified approach of dealing with regular and ill-posed problems.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna 1985