ON THE ACTUAL DIMENSIONS OF MEASUREMENTS PERFORMED WITH A LINEAR INSTRUMENT

Yu. P. Pyt'ev, S. P. Bondarenko, and D. V. Polyakov

The concept of effective rank of the linear model of the experiment is introduced, which determines the actual dimensions of both the results of measurements and the totality of parameters to be determined for the object under study. The effective rank enables one to characterize the quality of the solution of a measurement interpretation problem: the limiting resolving power, information content of measurements, and other important characteristics of the experiment.

Basic notation
\( \mathcal{R}_m \) is the Euclidean space;
\( \dim \mathcal{R}_s = s \) is the dimension of \( \mathcal{R}_s \);
\( \mathcal{R}(A) \) is the space of values of the operator \( A \);
\( \mathcal{N}(A) \) is the null-space (kernel) of the operator \( A \);
\( \mathcal{N}^\perp(A) \) is the orthogonal complement of \( \mathcal{N}(A) \);
\( \operatorname{rank} A \) is the rank of the operator \( A \);
\( A^* \) is the operator conjugate to \( A \);
\( A^- \) is the operator pseudoinverse to \( A \);
\( P, P \) are the orthogonal projectors;
\( ||A||_2 \) is the Hilbert–Schmidt norm of the operator \( A \).

INTRODUCTION

Recently, in performing various physical experiments, considerable interest has been repeatedly shown in the question of what information on a measured quantity can be obtained with a guaranteed accuracy. Evidently, in each particular experiment the answer to this question will be determined by the technological characteristics of the instruments used, by the accuracy of the experimental model and by the noise arising in the process of measurements. It would be desirable to construct a quantitative measure that would characterize in this respect the quality of a solution to the interpretation problem. Then, using the mathematical model of the instrument and knowing the type of noise in the measurement process and the accuracy desired, the researcher would know the limiting capabilities of the instrument before starting the experiment.

Consider a mathematical model of a typical linear scheme of measurements in physical studies, which is defined by the equality
\[
\xi = Af + \nu, \tag{1}
\]
where the measured signal \( f \) is \textit{a priori} assumed to be an arbitrary vector \( \mathcal{R}_m \), \( A \in \mathcal{R}_m \rightarrow \mathcal{R}_n \) is a linear operator which acts from the Euclidean space of signals \( \mathcal{R}_m \) onto the Euclidean space of measurements \( \mathcal{R}_n \) and models the measuring instrument, \( \nu \) is a random vector \( \mathcal{R}_n \) with the preset mathematical expectation \( E\nu = 0 \) and with the covariance operator \( \Sigma \) that models the measurement error. In short, it is assumed in what follows that a model \([A, \Sigma]\) of measurement scheme (1) is specified [1]. At the beginning of the
interpretation of measurement (1), a linear operator $U \in \mathcal{R}_m \to \mathcal{R}_k$ is given which determines the parameters of interest $Uf$ of the signal $f$, and one has to define the linear operator $R \in \mathcal{R}_m \to \mathcal{R}_k$ in such a way that

$$h(R, U) = \sup_{f \in \mathcal{R}_m} E \| R\xi - Uf \|^2 \sim \min_R.$$  \hfill (2)

Problem (2) is solvable if and only if $U(I - A^{-1}A) = 0$. For a nondegenerate operator $\Sigma$ the solution of the problem is unique and is given by the equality

$$R(U) = U(A^*\Sigma^{-1}A)^{-1}A^*\Sigma^{-1}. \hfill (3)$$

The operator $R = R(U)$ defined in this a way will give a maximum exact mean square (m.s.) linear estimate $R(U)\xi$ of the vector $Uf$. The mean square error of the interpretation is then

$$h(R(U), U) = h_* (U) = \sup_{f \in \mathcal{R}_m} E \| R(U)\xi - Uf \|^2 = \text{tr} U(A^*\Sigma^{-1}A)^{-1}U^*.$$ \hfill (4)

Suppose that problem (2) is solvable for $U = I$ and it is required that the m.s. error of estimation of $f$ should not exceed $\delta$. If $h(R(I), I) > \delta$, then $f$ cannot be estimated with the desired accuracy, and a question arises of the “part” of $f$ that has a maximum dimension, which can be estimated with the desired accuracy on the basis of measurement (1). It is reasonable to call the “maximum dimension” the effective rank, which will depend on the desired accuracy and on the model of the measurement scheme (1), i. e., on the pair of operators $A$, $\Sigma$ and $\delta$. The present paper is concerned with study and applications of the effective rank of the model in the interpretation of experiments.

**THE PROBLEM OF INTERPRETATION OF THE LINEAR MEASUREMENT FUNCTION**

In some cases we deal with a problem in which it is necessary to choose a component of $f$ whose dimension is no less than the given $k \leq m$, and to find no more than $q$ linear combinations of measurements which provide a maximum accuracy of measurement of this component of $f$.

Let $S \in \mathcal{R}_n \to \mathcal{R}_q$ be a linear operator, with $q \leq n$, the rank $S = q$ and the model $[A, \Sigma]$ of scheme of measurements (1) is given. In the problem of interpretation of the linear function $S\xi$ of measurement, the linear operator $U \in \mathcal{R}_m \to \mathcal{R}_k$ is given, and the linear operator $R$ has to be determined from the condition

$$\sup_{f \in \mathcal{R}_m} E \| RS\xi - Uf \|^2 \sim \min_R.$$ \hfill (5)

The solution to problem (5) is given by the following lemma.

**Lemma 1.** Let $S \in \mathcal{R}_n \to \mathcal{R}_q$, the rank $S = q$, the operator $\Sigma$ is nondegenerate and the condition

$$U(I - (SA)^{-1}SA) = 0 \hfill (6)$$

is satisfied. Then

$$\inf_{R'} \sup_{f \in \mathcal{R}_m} E \| R'\xi - Uf \|^2 = \text{tr} U(B^*PB)^{-1}U^*,$$

is reached at $R = R(U) = U(B^*PB)^{-1}B'(SS^1/2)^{-1}$, where $P = (SS^1/2) - SS^1/2$ is the orthogonal projector on $N^1(S^1/2)$, $B = \Sigma^{-1}A$. If condition (6) is not satisfied, problem (5) is unsolvable [2].

**Proof.** By the theorem hypothesis the operator $\Sigma$ and consequently $SS^*$ are nondegenerate. Then

$$\inf_{R'} \sup_{f \in \mathcal{R}_m} E \| R'\xi - Uf \|^2 = \text{tr} U(A^*S^*(SS^*)^{-1}SA)^{-1}U^* = \text{tr} U(B^*PB)^{-1}U^*,$$

and the operator $R = R(U)$ at which the lower bound is reached has the form

$$R(U) = U(A^*S^*(SS^*)^{-1}SA)^{-1}A^*S^*(SS^*)^{-1} = U(B^*PB)^{-1}B'(SS^1/2)^{-1}.$$
If condition (6) is not satisfied, the left-hand side of (5) is equal to infinity.

Let \( U = \Pi_s \), where \( \Pi_s \) is an orthogonal projector in \( \mathcal{R}_m \) which determines the orthogonal projection \( \Pi_s f \) of the vector \( f \) on the \( s \)-dimensional subspace \( \mathcal{R}_s \subset \mathcal{R}_m \) \((\text{rank } \Pi_s = s)\). According to Eq. (6), the orthogonal projection \( \Pi_s f \) of the vector \( f \) can be estimated only provided that

\[
\Pi_s \leq (SA)^{-1} SA.
\]

Consider the following variation problem:

\[
\inf_{\Pi_s} \sup_{\mathcal{R}_m} E \| R'S\xi - \Pi_s f \|^2,
\]

where \( P \) is an orthogonal projector defined in Lemma 1.

In this problem one has to determine the orthogonal projector \( \Pi_s, k \leq s \leq m, \) on the subspace \( \mathcal{R}_s \subset \mathcal{R}_m \), which is least struck by the noise among all the subspaces \( \mathcal{R}_m \) of the dimension \( s \), as well as the orthogonal projector \( P = P_t, t \leq n, t = \text{rank } P_t \), at which the minimum is reached.

The following results are directly related to the problem in question.

**Lemma 2.** Let \( \Pi \) be an orthogonal projector on \( L(e_1, e_2, e_k) \), where \( e_i, i = 1, \ldots, k \leq m \) are the orthonormalized eigenvectors of the eigenvalue problem

\[
B^* Be_i = \beta^2_i e_i, \quad i = 1, \ldots, m, \quad \beta^2_1 \geq \beta^2_2 \geq \ldots \geq \beta^2_m.
\]

Then

\[
\text{tr } \Pi (B^* B)^{-1} = \text{tr } (\Pi B^* B \Pi)^{-1} = \text{tr } (\Pi B^* B \Pi)^{-1}.
\]

**Proof.** By the hypothesis

\[
\text{tr } \Pi (B^* B)^{-1} = \sum_{i=1}^{\text{rank}(B^* B)} \beta^2_i, \text{ where } r = \text{rank } (B^* B)
\]

\[
\Pi B^* B \Pi e_i = \begin{cases} \beta^2_i e_i & i = 1, \ldots, k, \\ 0 & i = k + 1, \ldots, m. \end{cases}
\]

Hence,

\[
\text{tr } (\Pi B^* B \Pi)^{-1} = \sum_{i=1}^{\text{rank}(B^* B)} \beta^{-2}_i.
\]

**Lemma 3.** Let \( B \in \mathcal{R}_m \rightarrow \mathcal{R}_n, \Pi \) and \( P \) be orthogonal projectors on \( \mathcal{R}_m \) and \( \mathcal{R}_n \) respectively, and also \( \Pi \leq (B^* PB)^{-1} B^* PB \).

Then \( \Pi (B^* PB)^{-1} \geq \Pi (B^* B)^{-1} \).

**Proof.** Since for any orthogonal projector \( P \) the inequality \( B^* PB \leq B^* B \) is fulfilled, then according to [2] we have

\[
(B^* PB)^{-1} \geq (B^* PB)^{-1}(B^* PB)(B^* B)^{-1}(B^* PB)^{-1}(B^* B)^{-1}(B^* PB)^{-1},
\]

hence,

\[
\Pi (B^* PB)^{-1} \geq \Pi (B^* PB)^{-1}(B^* PB)(B^* B)^{-1}(B^* PB)^{-1}(B^* PB)^{-1} = \Pi (B^* B)^{-1},
\]

as the orthogonal projector \( \Pi \) is part of the orthogonal projector \( (B^* PB)^{-1}(B^* PB)^{-1} \).

Note that condition (7) is equivalent to the condition \( \Pi_i \leq (PB)^{-1} PB \), where \( P = (S\Sigma^{1/2} S)^{-1} S\Sigma^{1/2} \), since \( N(SA) = N(S\Sigma^{1/2} S) \subset N((S\Sigma^{1/2} S)^{-1} S\Sigma^{1/2} B) \subset N(S\Sigma^{1/2} B) = N(S\Sigma^{1/2} B) = N(SA) \).

**Theorem.** Let \( \Pi \) and \( P \) be orthogonal projectors, \( k \leq m, t \leq n, k \leq t \). Then

\[
\min_{\Pi: \text{rank } \Pi \geq k, \Pi \leq (PB)^{-1} PB} \text{tr } (B^* PB)^{-1} \Pi = \sum_{i=1}^{k} \beta^{-2}_i.
\]
is achieved on the orthogonal projector \( \Pi_k \), (rank \( \Pi_k = k \)), which projects on \( \mathcal{L}(e_1, e_2, \ldots, e_k) \), where \( e_i \), \( i = 1, \ldots, k \), \( k \leq r = \text{rank}(B^*B) \) are the orthonormalized eigenvectors of problem (8), and on any orthogonal projector \( P_q \), (rank \( P_q = q \)), projecting on \( \mathcal{L}(s_1, s_2, \ldots, s_q) \), \( k \leq q \leq t \), where \( s_i = \beta_i^{-1}B e_i \), \( i = 1, \ldots, r \), are the orthonormalized system in \( \mathcal{R}_n \), \( s_{r+1}, \ldots, s_n \) are arbitrary linearly independent vectors from the orthogonal complement of \( \mathcal{L}(s_1, s_2, \ldots, s_r) \) in \( \mathcal{R}_n \).

**Proof.** According to [2],

\[
\min_{\Pi: \text{rank } \Pi \leq k, \Pi \in \mathcal{B}^-B} \text{tr } \Pi (B^*B)^{-1} \Pi = \text{tr } \Pi_k (B^*B)^{-1} \Pi_k = \sum_{i=1}^{k} \beta_i^{-2},
\]

\( \Pi_k \) is an orthogonal projector (rank \( \Pi_k = k \)) on \( \mathcal{L}(e_1, e_2, \ldots, e_k) \). It is well known (see, e.g., [2]) that the orthonormalized basis \( \{ e_i \} \) of (8) creates in \( \mathcal{R}_n \) an orthonormalized system \( s_i = \beta_i^{-1}B e_i \), \( i = 1, \ldots, r \). Let \( s_{r+1}, \ldots, s_n \) be any linearly independent vectors from the orthogonal complement of \( \mathcal{L}(s_1, s_2, \ldots, s_r) \) in \( \mathcal{R}_n \). Then, choosing \( P = P_q \), where \( P_q \) is an orthogonal projector on \( \mathcal{L}(s_1, s_2, \ldots, s_q) \), \( k \leq q \leq t \), according to Lemma 3 we obtain

\[
\Pi_k (B^*B)^{-1} \Pi_k \leq \Pi_k (B^*PB)^{-1} \Pi_k = \Pi_k (B^*B)^{-1} \Pi_k.
\]

For any orthogonal projector \( P \)

\[
\min_{\Pi: \text{rank } \Pi \leq k, \Pi \in \mathcal{B}^-B} \text{tr } \Pi (B^*PB)^{-1} \Pi \geq \min_{\Pi: \text{rank } \Pi \leq k, \Pi \in \mathcal{B}^-B} \text{tr } \Pi (B^*B)^{-1} \Pi
\]

\[
\geq \min_{\Pi: \text{rank } \Pi \leq k, \Pi \in \mathcal{B}^-B} \text{tr } \Pi (B^*B)^{-1} \Pi = \text{tr } \Pi_k (B^*B)^{-1} \Pi_k = \sum_{i=1}^{k} \beta_i^{-2},
\]

and for \( P = P_q \) the equalities in (10) are fulfilled.

**Corollary.** The minimum projector \( P \) at which minimum (10) is reached, has the rank \( k \) and is given by the equality \( P = B \Pi_k (B \Pi_k)^{-1} \).

**EFFECTIVE RANK OF THE MEASUREMENT MODEL.**

Based on the results presented in the preceding sections, we introduce the notion of the effective rank of the model \( [A, \Sigma] \) of measurement scheme (1). It is well known that for each model \( [A, \Sigma] \) there exists a corresponding specific expanding sequence of linear subspaces of the space \( \mathcal{R}_m \) of input signals. Each of these is characterized by a minimal (among all the linear subspaces of the same dimension) value of the m.s. error of estimation of signals existing in it. Each of these "extremal" subspaces \( \mathcal{L}_s \) is a linear span of the first \( s \) orthonormalized eigenvectors of problem (8). If \( \Pi_s \) is an orthogonal projector on \( \mathcal{L}_s \), then, according to expression (4) and equalities (8) (recall that \( B = \Sigma^{-1/2}A \)), the m.s. error of estimation \( R(\Pi_s) \) of the signal \( \Sigma \) satisfies the inequality

\[
h_\star(\Pi_s) = \sum_{j=1}^{s} \beta_j^{-2} \leq h_\star(U_s),
\]

where \( U_s \) is an orthogonal projector on any linear subspace of input signals with the dimension \( s \), \( s = 1, \ldots, m \).

The orthonormalized basis \( \{ e_j \} \) (8) of the Euclidean space \( \mathcal{R}_m \) is called the proper basis of the model \( [A, \Sigma] \). This basis produces in \( \mathcal{R}_m \) an orthonormalized system

\[
s_j = \beta_j^{-1} \Sigma^{-1/2} A e_j, \quad j = 1, \ldots, r \quad \beta_1^2 \geq \beta_2^2 \geq \ldots \geq \beta_r^2 \geq \beta_{r+1}^2 = \ldots = \beta_m^2 = 0
\]

(\( r = \text{rank} A \)), and in this case it is sufficient to know \( \tilde{\xi}_j = (s_j \Sigma^{-1/2} \xi) \), \( j = 1, \ldots, k \), to estimate the orthogonal projection \( \Pi_k f \), \( k \leq r \), since according to (3) and (8)

\[
\Pi_k (A^* \Sigma^{-1} A)^{-1} A^* \Sigma^{-1} \xi = \sum_{j=1}^{k} \beta_j^{-1} (s_j, \Sigma^{-1/2} \xi) e_j, \quad k = 1, \ldots, r.
\]
Hence, the minimum m.s. error of estimating the $k$-dimensional orthogonal component of $f$ is given by the formula $h = \sum_{i=1}^{k} \beta_i^{-2}$, where $\beta_i^2$, $i = 1, \ldots, r = \text{rank} \, A$ are eigenvalues of problem (8).

**Definition.** We will call the function which is defined on the half-line $\mathbb{R}_+ = [0, +\infty)$ and assumes the values 0, 1, 2, \ldots, $r = \text{rank} \, A$, the effective rank of the model $[A, \Sigma]$.

The function $\rho[A, \Sigma](h)$ is the maximum dimension of the orthogonal component of $f \in \mathbb{R}_+$ which can be estimated with a m.s. error not exceeding $h \in \mathbb{R}_+$. If the accuracy of estimation, as determined by the m.s. error $h$, is considered acceptable, then any linear combination of the first $\rho[A, \Sigma](h)$ of the eigenvectors of problem (8) can be estimated with this accuracy. Consequently, the first $\rho[A, \Sigma](h)$ of the eigenvectors of problem (8) show what details of the signal $f \in \mathbb{R}_m$ admit estimating with an acceptable accuracy.

To estimate the orthogonal component of $f$ with the dimension $\rho[A, \Sigma](h)$ with a m.s. error not exceeding $h \in \mathbb{R}_+$, one needs $\rho[A, \Sigma](h)$ linear combinations of measurements $\xi_j = (s_j, \Sigma^{-1/2} \xi)$.

Let us formulate the basic properties of the function $\rho[A, \Sigma](h)$.

**Lemma 4.**
1. The effective rank $\rho[A, \Sigma](h)$ is a nondecreasing function $h \in \mathbb{R}_+$, and $\lim_{h \to \infty} \rho[A, \Sigma](h) = r = \text{rank} \, A$.
2. If the model $[A, \Sigma]$ is uniformly no worse than the model $[\tilde{A}, \tilde{\Sigma}]$, $([A, \Sigma] \prec [\tilde{A}, \tilde{\Sigma}]$, see [1]), then $\rho[A, \Sigma](h) \geq \rho[\tilde{A}, \tilde{\Sigma}](h)$.
3. If $\Sigma > 0$ and $||\Sigma||_2 \to 0$, then $\rho[A, \Sigma](h) \to r = \text{rank} \, A$, i.e., $\lim_{||\Sigma||_2 \to 0} \rho[A, \Sigma](h) = r$.

**Proof.**
1. It immediately follows from the definition.
2. Let $\Pi_k$ and $\tilde{\Pi}_k$ be orthogonal projectors on the first $k$ vectors of the basis of the models $[A, \Sigma]$ and $[\tilde{A}, \tilde{\Sigma}]$, respectively, where $k \leq \text{min}(r, \bar{r})$. According to [1], the model $[A, \Sigma]$ is uniformly no worse than the model $[\tilde{A}, \tilde{\Sigma}]$, if $A - A \geq \tilde{A} - \tilde{A}$ and $\tilde{A} - \tilde{A}((A^* \Sigma^{-1} A)^{-1} - (\tilde{A}^* \tilde{\Sigma}^{-1} \tilde{A})^{-1}) \tilde{A} - \tilde{A} \leq 0$. Bearing in mind that the model $[A, \Sigma]$ is uniformly no worse than the model $[\tilde{A}, \tilde{\Sigma}]$, and taking into account the extremal properties of the basis of the model, we have

$$\text{tr} \, \Pi_k (A^* \Sigma^{-1} A) - \Pi_k \leq \text{tr} \, \tilde{\Pi}_k (A^* \Sigma^{-1} A) - \tilde{\Pi}_k \leq \text{tr} \, \tilde{\Pi}_k (\tilde{A}^* \tilde{\Sigma}^{-1} \tilde{A}) - \tilde{\Pi}_k \leq \text{tr} \, U_k (\tilde{A}^* \tilde{\Sigma}^{-1} \tilde{A}) - U_k,$$

where $U_k$ is the orthogonal projector on any linear subspace with the dimension $k$ that satisfies the condition $U_k \leq \tilde{A} - \tilde{A}$. It follows from the above written inequalities that $\sum_{i=1}^{k} \beta_i^{-2} \leq \sum_{i=1}^{\bar{r}} \tilde{\beta}_i^{-2}$ for any $k \leq \bar{r} \leq r$, where $\beta_i^2$, $i = 1, \ldots, r = \text{rank} \, A$, and $\tilde{\beta}_i^2$, $i = 1, \ldots, \bar{r} = \text{rank} \, \tilde{A}$ are the eigenvalues of problem (8) for the models $[A, \Sigma]$ and $[\tilde{A}, \tilde{\Sigma}]$, respectively, where $\bar{r} \leq r$ since $[A, \Sigma] \prec [\tilde{A}, \tilde{\Sigma}]$. Hence, for any $h \in \mathbb{R}_+$, $\rho[A, \Sigma](h) \geq \rho[\tilde{A}, \tilde{\Sigma}](h)$.

3. Denote the minimum and maximum eigenvalues of the operator $\Sigma$ as $\sigma_{\min}^2$ and $\sigma_{\max}^2$. Let $\Sigma > 0$. Then $\sigma_{\max}^2 \leq \Sigma^{-1} \leq \sigma_{\min}^2$ and for any $e \in \mathbb{R}_m$,

$$\sigma_{\max}^2 ||Ae||^2 \leq (A^* \Sigma^{-1} A e, e) \leq \sigma_{\min}^{-2} ||Ae||^2. \quad (11)$$

Let $e = \epsilon_r$ be the eigenvector of problem (8) that corresponds to the minimum nonzero eigenvalue $\beta_r^2$. In this case inequalities (11) can be rewritten in the form

$$\sigma_{\max}^{-2} ||Ae_r||^2 \leq \beta_r^2 \leq \sigma_{\min}^{-2} ||Ae_r||^2.$$
Mean square error of estimation of the vertical distribution profile of ozone versus the dimension of its representation for the measurement model of the scattered ultraviolet radiation spectrum at fixed zenith angles of the Sun $\theta_s = 70^\circ$ (1) and $\theta_s = 45^\circ$ (2).

Diagram of the effective rank of the measurement model of the scattered ultraviolet radiation spectrum at fixed zenith angles of the Sun $\theta_s$ of $70^\circ$ (1) and $45^\circ$ (2) in the vertical ozone distribution problem.

The condition $||\Sigma||_2 \to 0$ is equivalent to the condition $\sigma^2_{\text{max}} \to 0$. Hence, $\beta_r^2 \to \infty$ and for any $k \leq r$ we have $\sum_{i=1}^{k} \beta_i^{-2} \to 0$. From here we obtain that for any $h \in \mathcal{R}_+$ $\lim_{||\Sigma||_2 \to 0} \rho[A, \Sigma](h) = r$.

The functions $h = \sum_{i=1}^{r} \beta_i^{-2} = h(k)$ and $\rho[P, \Sigma](h)$, $h \in \mathcal{R}_+$, $k = 1, 2, \ldots, r$, ($r = 10$) in the problem of reconstruction of the vertical profile of ozone from measurements of the ultraviolet radiation at a given position of the Sun ($m = 15$, $n = 10$) are given in Figs. 1 and 2. One can see that $\rho[A(\theta_s=45^\circ), \Sigma](h) \leq \rho[A(\theta_s=70^\circ), \Sigma](h)$, and therefore measurements at the angle $\theta_s = 70^\circ$ enable the vertical ozone profile to be reconstructed with greater accuracy.

REFERENCES


21 November 1994