INFORMATION THEORY OF THE INTERPRETATION OF GEOPHYSICAL INVESTIGATIONS

1958

SovietRxiv

Abstract

Full Text

GEOPHYSICS

L. A. KHALFIN

INFORMATION THEORY OF THE INTERPRETATION OF GEOPHYSICAL INVESTIGATIONS

(Presented by Academician V. I. Smirnov, 19 V 1958)

In the present work an information theory of the interpretation of geophysical investigations is considered. It differs from the usual theory of interpretation in that it is based on the idea of describing geophysical methods of investigation as systems of informational observation ([1],[2]). In addition to the information-theoretic approach, the problems considered may be investigated on the basis of the classical theory of statistical estimation*. The question of the connection between the information-theoretic approach and the theory of statistical estimation has not yet been fully resolved ([3–5]).

1. As is known, the problem of interpreting the data of a geophysical method consists in determining, from the geophysical field $\varphi_\rho(\mathbf{r})$ (the signal), the corresponding distribution of the sources of this field ρ (the message)**. In this case the obvious requirement is that, for a given geophysical field $\varphi_\rho(\mathbf{r})$, only such a characteristic of the sources of this field may be called the message ρ as is in a one-to-one correspondence (uniqueness theorem) with the given geophysical field (signal) $\varphi_\rho(\mathbf{r})$:

$$\varphi_\rho(\mathbf{r}) \rightleftarrows \rho. \tag{1}$$

In such a formulation, the division into direct and inverse problems is natural.

It is obvious that interpretation may then be either palette-based or analytical. By palette interpretation is meant the following: assigning various ρ_i, one finds, by solving direct problems, the corresponding $\varphi_{\rho_i}(\mathbf{r})$, so that as a result one obtains tables—palettes:

$$\rho_i \to \varphi_{\rho_i}(\mathbf{r}); \tag{2}$$

the interpretation then consists in finding, for the measured field $\psi(\mathbf{r})$, the corresponding $\varphi_{\rho_j}(\mathbf{r})$ in the tables (2) (it is assumed that this can always be done if the palettes are sufficiently detailed), and consequently the desired ρ_j as well. Recently the so-called analytical method of interpretation has also been developing intensively. The essence of this method consists in solving inverse problems (in particular, in potential theory).

As we see, neither in the palette nor (still more) in the analytical methods of interpretation is there any place at all for so-called complex geological-geophysical interpretation, although it is precisely this that is applied in practice. This contradiction in the usual theory of interpretation must evidently be caused by some fundamental shortcoming of it. This shortcoming consists in the fact that, in the usual approach to the theory of interpretation, the presence of noise in the broadest sense of the word is completely ignored.

* Yu. V. Linnik drew my attention to this during a discussion of the work.
** In writing $\varphi_\rho(\mathbf{r})$ we denote by $\mathbf{r}$ the variable coordinates of the geophysical field (for example, spatial coordinates).

2. From the very beginning we shall assume that in the measured field (signal), which we denote by $\psi(\mathbf{r})$, there is a field (signal) with useful ρ-information and noise $n(\mathbf{r})$ with interfering n-information:

$$\psi(\mathbf{r}) = \varphi_\rho(\mathbf{r}) + n(\mathbf{r}). \tag{3}$$

Let us dwell somewhat more on the concept of "noise." For this it is necessary to specify the formulation of the problem confronting the geophysical method. It is quite natural that it can be formulated as follows: from a given class of messages ρ_i, where the number of different messages is finite ($i = 1, 2, \dots, I$), determine the message ρ_j that corresponds to the measured field (signal) $\psi(\mathbf{r})$. The mere fact that we have restricted the set (class) of messages to be interpreted (and this is fundamentally necessary) leads to the necessity of including in the noise the uncontrolled influence of all the remaining non-interpreted sources of messages. In addition, we include in the noise the uncontrolled part due to instrumental measurement errors, as well as the errors that arise in processing $\psi(\mathbf{r})$.

Since the measured field $\psi(\mathbf{r})$ differs from $\varphi_\rho(\mathbf{r})$, it is no longer possible, from the given field $\psi(\mathbf{r})$, to indicate unambiguously the corresponding message ρ. The most that can be done, without resorting to guesses (which destroy information), is to determine the probability of the message ρ for the measured field $\psi(\mathbf{r})$. Since one-to-one transformations do not change the amount of information (are informationally void), the problem of interpretation obviously reduces to the possible preservation of ρ-information and the destruction of n-information by means of an irreversible transformation of the field $\psi(\mathbf{r})$.

3. Suppose that the noises are homogeneous, i.e., the values $n(\mathbf{r}_k)$ for different k* are statistically independent and the distribution $p_1(n(\mathbf{r}_k))$ is the same for different k**. Then, on the basis of Bayes' theorem and definition (3), we obtain the desired result ([6],[7]):

$$p_\psi(\rho_i) = ap(\rho_i)p_1(n) = ap(\rho_i)\prod_k p_1(\psi(\mathbf{r}_k) - \varphi_{\rho_i}(\mathbf{r}_k)), \qquad \sum_i p_\psi(\rho_i) = 1, \quad (4)$$

where $p_\psi(\rho_i)$ is the posterior probability of the message ρ_i for the measured field $\psi(\mathbf{r})$ (i.e., the desired result of the interpretation); $p(\rho_i)$ is the prior probability of the message ρ_i; $p_1(n) = p_{\rho_i}(\psi)$ is the likelihood function; a is a normalizing constant. The principal result (4) consists in isolating the factor $p(\rho_i)$, which does not depend on $\varphi_{\rho_i}(\mathbf{r})$ and $\psi(\mathbf{r})$, thereby determining the probability of the message ρ_i before the application of the given geophysical method, so that, for example, the factor $p(\rho_i)$ reflects preliminary geological conceptions or the results of interpretation obtained on the basis of other geophysical methods.

4. As was shown in the preceding point, if the noise distribution $p_1(n)$ were known, the result of the interpretation would be easy to find according to (4). However, usually the noise distribution $p_1(n)$ is completely unknown; only some of its characteristics are known.

Suppose that we know the first m moments of the distribution $p_1(n)$:

$$\int_{N_1}^{N_2} p_1(n)\,dn = 1, \qquad \int_{N_1}^{N_2} np_1(n)\,dn = \varepsilon_1, \qquad \int_{N_1}^{N_2} n^m p_1(n)\,dn = \varepsilon_m. \quad (5)$$

Then it is easy to find such a distribution $p(n)$ which, being normalized to m moments according to (5), makes maximal the amount

* We thereby explicitly take into account that measurements are made at a finite number of points.

** This assumption is by no means restrictive, since homogeneous noises are the most dangerous. Naturally, in a more exact approach it is necessary to take into account the correlation characteristics of the noise.

n-information (8):

$$I\{n\} = -\int_{N_1}^{N_2} p(n)\log p(n)\,dn. \quad (6)$$

Solving the corresponding variational problem, we obtain*:

$$p(n) = \exp\{\lambda_0 + \lambda_1 n + ... + \lambda_m n^m\}. \quad (7)$$

Then choosing in (4) $p_1(n)$ in the form (7), we obtain

$$p_\psi(\rho_i) = ap(\rho_i)\exp\left\{\sum_k \Big(\lambda_1[\psi(\mathbf{r}_k) - \varphi_{\rho_i}(\mathbf{r}_k)] + \ldots + \lambda_m[\psi(\mathbf{r}_k) - \varphi_{\rho_i}(\mathbf{r}_k)]^m\Big)\right\}. \tag{8}$$

The result (8) determines the minimum amount of information in $\psi(\mathbf{r})$ relative to ρ for the specified characteristics (5) of the distribution $p_1(n)$.

5. Let us consider several basic cases.

1) Suppose that only the normalization of $p_1(n)$ is known:

$$\int_{N_1}^{N_2} p_2(n)\,dn = 1, \tag{9}$$

i.e., practically nothing is known about the noise. In this case, from (7), (8) we obtain

$$p_\psi(\rho_i) = \frac{a}{N_2 - N_1}\,p(\rho_i) = a_1 p(\rho_i), \tag{10}$$

i.e., a completely natural result—the geophysical experiment adds nothing to the a priori (geological) concepts.

2) Suppose that the mean value of the noise is also known:

$$\int_{N_1}^{N_2} p_1(n)\,dn = 1; \qquad \int_{N_1}^{N_2} np_1(n)\,dn = \varepsilon. \tag{11}$$

Then, on the basis of (7), (8), we obviously obtain:

$$p_\psi(\rho_i) = a_1 p(\rho_i)\exp\left\{-\lambda_1 \sum_k \varphi_{\rho_i}(\mathbf{r}_k)\right\}, \tag{12}$$

i.e., in this case the result of the interpretation refines the a priori concepts, but only on the basis of theoretical concepts ($\varphi_{\rho_i}(\mathbf{r})$), and not on the basis of the experimental results ($\psi(\mathbf{r})$).

3) Finally, suppose that the variance is also known for the noise:

$$\int_{N_1}^{N_2} p_1(n)\,dn = 1; \qquad \int_{N_1}^{N_2} np_1(n)\,dn = \varepsilon; \qquad \int_{N_1}^{N_2} n^2 p_1(n)\,dn = \sigma^2. \tag{13}$$

In this case, on the basis of (7), (8) we easily obtain

$$p_{\psi}(\rho_i) = a_1 p(\rho_i) \exp\left\{-\lambda_1 \sum_k \varphi_{\rho_i}(\mathbf{r}_k) + \lambda_2 \sum_k \varphi^2_{\rho_i}(\mathbf{r}_k)\right\} \times$$

$$\times \exp\left\{-2\lambda_2 \sum_k \psi(\mathbf{r}_k)\varphi_{\rho_i}(\mathbf{r}_k)\right\}. \qquad (14)$$

* $\lambda_0, \lambda_1, \dots, \lambda_m$ are constants determined from (5).

Interpretation in this case consists of: a) geological (the term $p(\rho_i)$); b) theoretical (the first exponent, depending only on $\varphi_{\rho_i}(\mathbf{r})$); c) theoretical processing of the experimental results (the second exponent, depending both on $\psi(\mathbf{r})$ and on $\varphi_{\rho_i}(\mathbf{r})$). Thus, we have indeed shown that the interpretation of geophysical investigations is a combined geological-geophysical one.

It is evident from (14) that knowledge of at least the second moment of the noise distribution is necessary in order for the result of the experiment ($\psi(\mathbf{r})$) to refine the a priori (geological) concepts in the interpretation.

6. Let us include in the consideration of the information theory of interpretation the combining of geophysical methods. The essence of combining is manifested in the fact that the result of interpreting one geophysical method ($p_{\psi_1}(\rho_i)$) can be used as the a priori probability for another method:

$$p_{\psi_2}(\rho_i) = ap(\rho_i) \exp\left\{\sum_k \left[(\lambda_1^{(1)}[\psi_1(\mathbf{r}_k) - \varphi^{(1)}_{\rho_i}(\mathbf{r}_k)] + \lambda_1^{(2)}[\psi_2(\mathbf{r}_k) - \varphi^{(2)}_{\rho_i}(\mathbf{r}_k)]) + \dots\right.\right.$$

$$\left.\left.\dots + (\lambda_m^{(1)}[\psi_1(\mathbf{r}_k) - \varphi^{(1)}_{\rho_i}(\mathbf{r}_k)]^m + \lambda_m^{(2)}[\psi_2(\mathbf{r}_k) - \varphi^{(2)}_{\rho_i}(\mathbf{r}_k)]^m)\right]\right\}, \qquad (15)$$

i.e., the results of interpreting one method refine the results of interpreting another geophysical method.

It is precisely the combining of geophysical methods that makes it possible to obtain reliable interpretation results when the reliability of the interpretation results of the individual methods is low.

7. Let us now single out the accounting, in interpretation, for the physical properties of geological objects. Namely, since the geophysical field is proportional to the concentration of sources C, formulas (3), (8) become the formulas

$$\psi(\mathbf{r}) = C\varphi_{\rho}(\mathbf{r}) + n(\mathbf{r});$$

$$p_\psi(\rho_i) = ap(\rho_i) \int p_{\rho_i}(C) \exp\left\{\sum_k \Big(\lambda_1[\psi(\mathbf{r}_k) - C\varphi_{\rho_i}(\mathbf{r}_k)] + ... \right.$$
$$\left. ... + \lambda_m[\psi(\mathbf{r}_k) - C\varphi_{\rho_i}(\mathbf{r}_k)]^m\Big)\right\} dC, \quad (16)$$

where the function $p_{\rho_i}(C)$ determines the so-called physical properties of geological objects.

8. The information theory of interpretation of geophysical investigations set forth above also makes it possible to compare various (ordinary) methods of interpretation with one another, taking into account that, in the information scheme, different methods of interpretation correspond to different signals $\varphi_\rho(\mathbf{r})$ for one and the same source of a message ρ.
9. We have briefly presented above a scheme of the information theory of interpretation of geophysical investigations. The principal result is the obtaining of an algorithm of the information theory of interpretation, which can be realized by a computing device with memory $(\varphi_\rho(\mathbf{r}))$.

The results actually obtained are results of the general information theory of observation.

In conclusion, I express my gratitude to Prof. Yu. V. Linnik and Prof. A. S. Semenov for discussing the work and for valuable comments.

All-Union Scientific-Research Institute
of Exploration Geophysics

Received
14 V 1958

CITED LITERATURE

1. L. A. Khalfin, Course of lectures at Leningrad State University in 1956–1957 (in press).
2. L. A. Khalfin, Reports at the section on information theory of the All-Union scientific session devoted to Radio Day, Moscow, 1957.
3. S. Kullback, *Ann. Math. Stat.*, **23**, 88 (1952).
4. S. Kullback, *Ann. Math. Stat.*, **25**, 745 (1954).
5. D. V. Lindley, *Ann. Math. Stat.*, **28**, 986 (1957).
6. B. V. Gnedenko, *Course in Probability Theory*, Moscow, 1954.

7. F. M. Woodward, *Probability Theory and Information Theory, with Applications to Radar*, Moscow, 1955.

8. R. E. Shannon, W. Weaver, *The Mathematical Theory of Communication*, Illinois, 1949.

Note: Figure translations are in progress. See original paper for figures.