

On one numerical algorithm of seismic networks planning

V.N. Beloborodov, O.K. Omelchenko

This paper is dealt with a brief statement of a basis of the theory of optimal planning of seismic networks. Some concepts of such a planning of seismic networks are given. Some specific formulations of problems of planning of seismic networks are presented. A numerical algorithm of constructing discrete optimal plans as applied to problems of planning seismic networks is stated.

Introduction

A source of the primarily observed seismic data in seismology is an observation system (OS), i.e., a network of spatially distributed seismic stations equipped with instruments for recording seismic waves.

The basic OS parameters in seismology are as follows [7]:

- the number of seismic stations, the geometric configuration of a network, and individual station sites;
- the frequency responses of the recording instruments, their dynamic range and amplification.

Definition 1. We shall call a given number of seismic stations deployed at fixed sites a seismograph network (SN).

Definition 2. Seismograph networks are usually divided into several categories by their spatial dimensions:

- *local networks* ranging in size from a few hundred meters to a few tens of kilometers;
- *zonal networks*, from a few tens to a few hundreds of kilometers;
- *regional networks*, from a few hundreds to a few thousands of kilometers;
- *global networks*, which are deployed all over the world or a large part of it.

For the purpose of economy the number of stations in a network should be kept to a minimum without affecting the quality of records; this naturally calls for an optimum network design.

The modern worldwide tendency to optimize seismic networks is to minimize errors in determining basic kinematic hypocenter parameters. This problem known as planning of experiment [1, 3, 9, 10, 12, 16–19], and problem of the SN optimization can be termed the planning of a seismic network. The mathematical planning of experiment uses methods of mathematical statistics and optimization techniques.

The problem of designing a SN has arisen from the parameter estimation problem Hypocenters of earthquakes, setting and methods of solution by which one reduce here briefly.

1. The hypocenter location problem

The basic data for this problem are the coordinates of existing seismic stations and those of possible sites for the new stations to be added to the network, the velocity structure of the region under study, and the positions of the seismic-prone zones which are to be studied using the network in question. This information related to arrival times of the waves excited by an earthquake, is obtained by means of nonlinear equations of condition [3, 8–10]:

$$\vec{T} = \vec{\eta}(X, \vec{\theta}) + \vec{\varepsilon}, \quad (1)$$

where

$\vec{T} = (T_1, T_2, \dots, T_N)^T$ is the arrival time vector,

$\vec{\eta}(X, \vec{\theta})$ is the N -dimensional vector of theoretical arrival times or the regression function,

$\vec{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_N)^T$ is the vector of residuals,

$\vec{\theta} = (\varphi, \lambda, h, t)^T$ is the vector of estimated parameters,

$X = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ is the matrix of stations' coordinates,

N is the number of recorded arrival times,

n is the number of stations.

The estimation of $\vec{\theta}$ is treated by the regression analysis: the solutions are the least squares (LS) estimates

$$\vec{\theta} = \arg \min_{\vec{\theta} \in \Omega} Q(\vec{\theta}), \quad Q(\vec{\theta}) = \sum_{i=1}^N \sigma_i^{-2} (T_i - \eta(\vec{x}_i, \vec{\theta}))^2. \quad (2)$$

The functional $Q(\vec{\theta})$ has usually been minimized in seismology since Geiger's times using the iterative Gauss-Newton method based on a linear fit to the regression function around the point $\vec{\theta}^k$:

$$J(X, \bar{\theta}^k) \Delta \bar{\theta}^k + \bar{\eta}(X, \bar{\theta}^k) - \bar{T} + \bar{\varepsilon} = 0, \quad (3)$$

where

$$J(X, \bar{\theta}) = \left(\frac{\partial \eta(\bar{x}_i, \bar{\theta})}{\partial \theta_1}, \frac{\partial \eta(\bar{x}_i, \bar{\theta})}{\partial \theta_2}, \dots, \frac{\partial \eta(\bar{x}_i, \bar{\theta})}{\partial \theta_m} \right), \quad i = 1, 2, \dots, n. \quad (4)$$

Multiplying both parts of the linearized equations (3), (4) by $J^T(X, \bar{\theta}^k)$ from the left, one obtains the following normal equations:

$$J^T(X, \bar{\theta}^k) J(X, \bar{\theta}^k) \Delta \bar{\theta}^k = J^T(X, \bar{\theta}^k) \bar{y}(X, \bar{\theta}^k), \quad (5)$$

where $\bar{y}(X, \bar{\theta}) = (\bar{T} - \eta(X, \bar{\theta}))^T$.

The estimates of $\bar{\theta}$ are found by iteration ($\bar{\theta} = \lim_{k \rightarrow \infty} \bar{\theta}^k$):

$$\bar{\theta}^{k+1} = \bar{\theta}^k + [J^T(X, \bar{\theta}^k) J(X, \bar{\theta}^k)]^{-1} J^T(X, \bar{\theta}^k) \bar{y}(X, \bar{\theta}^k), \quad k = 0, 1, 2, \dots, \quad (6)$$

where the starting fit $\bar{\theta}^0$ should be chosen as close to the true values.

The matrix

$$M(X, \bar{\theta}) = J^T(X, \bar{\theta}) J(X, \bar{\theta}) \quad (7)$$

is a Fisher data matrix or the design (planning) data matrix. The related matrix

$$D(X, \bar{\theta}) = M^{-1}(X, \bar{\theta}) \quad (8)$$

known as a covariance matrix (of the parameter space) contains estimates of the unknown parameters $\bar{\theta}$ errors.

The matrix $M(X, \bar{\theta})$ is not inverted during the iterative process (6) or its modifications [11]; each step involves transition from (3), (4) to (5), the latter equations being solved by a standard method. For this reason process (6) can be written down as follows:

$$\bar{\theta}^{k+1} = \bar{\theta}^k + \Delta \bar{\theta}^k, \quad M(X, \bar{\theta}^k) \Delta \bar{\theta}^k = J^T(X, \bar{\theta}^k) \bar{y}(X, \bar{\theta}^k), \quad (9)$$

$k = 0, 1, 2, \dots$

The above computational sequence of normal equations scheme has the advantage that many satisfactory techniques exist to solve (3), (4); also, the covariance matrix (8) is easily to be found. Its disadvantage in the fact that $M(X, \bar{\theta})$ is poorly conditioned for certain cases of hypocenter position and network geometry. Such cases can be handled by various regularization techniques [4, 14].

Another approach to solving (1)–(4) is to abstain from using normal equations, but solve the iterative process for (3), (4) directly at each step. The most popular recent method to do this is singular value decomposition (SVD) or the generalized inversion [13, 14]. A standard FORTRAN-IV

procedure is available [13]. The computational scheme of the Gauss–Newton singular value decomposition is to decompose (4) into a product of three matrices at each step of the iterative process:

$$J(X, \vec{\theta}^k) = U_k \Sigma_k V_k^T, \quad (10)$$

where U_k is an orthogonal $n \times n$ matrix, V_k is an orthogonal $m \times m$ matrix, Σ_k is a diagonal $n \times m$ matrix having the structure $\Sigma_k = \begin{pmatrix} S_k \\ 0 \end{pmatrix}$, where $S_k = \text{diag}(\rho_1, \rho_2, \dots, \rho_m)$ is a diagonal matrix of singular values arranged in nonincreasing order $\rho_i \geq \rho_{i+1}$.

The method also provides a so-called singular value analysis, which consists in the elimination of zero singular values and the respective columns in U and V . The iterative process then becomes

$$\vec{\theta}^{k+1} = \vec{\theta}^k + V_k S_k^{-1} \vec{d}^k, \quad k = 0, 1, 2, \dots, \quad (11)$$

where \vec{d}^k is a vector which consists of the first m components of $U_k^T \vec{y}(X, \vec{\theta}^k)$.

It can be shown that, at each step of the iterative process (11), the vector $\Delta \vec{\theta}^k = V_k S_k^{-1} \vec{d}^k$ minimizes not only the functional $Q(\vec{\theta})$, in which the vector $\vec{\eta}(X, \vec{\theta})$ has been replaced by its linear part as given by (4), but also the norm of the parameter vector, which ensures the uniqueness of the solution. The advantage of this process, as compared with (6), (9), is that one easily obtains as a side result not only the covariance matrix of the parameter space but also the matrix of the data space [18].

2. The necessity of planning SN, a concept of plan, and network design

To sum up, using any of the above methods based on the Jacoby matrix (4) of the linearized equations of condition (3), (4), one can estimate earthquake hypocenter parameters and the associated uncertainties. The solution of this problem is discussed, for example, in [8–10, 19]. Note that the elements of the Jacoby matrix are functions of the hypocenter parameters, velocity structure, and station locations, but are independent of the arrival times at the stations. It will be shown below that this allows one to construct theoretical covariance matrices of the parameter and data space, and, hence, estimate possible errors in hypocenter parameters at various points of the region and the contribution of each station into the network performance [9, 10, 15–18].

However good the iterative techniques (6), (11) for hypocenter location may be, they are unsatisfactory for a poorly conditioned matrix (4). The regularization techniques recommended for such cases often fail to give the desired effect in practice. The cause of matrix (4) being poorly conditioned

lies in poor observational arrangements, namely, in poor network geometries with respect to seismic source zones. The necessary conclusion is that observations should be planned beforehand; that is, network geometries should be chosen so that matrix (4) should be as well conditioned as possible to improve in parameter estimation. The design of seismograph networks is thus to remove the cause why matrix (4) is poorly conditioned, rather than trying to mend the matters by using various regularizations.

The founder of the science of the design of experiments, R. Fisher, was the first to see that, whereas the most effective parameter estimation techniques can yield accuracy gains of a few tens of percent at most, the gain from more sensible experimental designs (efficient observation arrangements) may be a few times. It is now generally recognized in the theory of experimental design that well-advised preliminary planning is required for costly experiments (e.g., explosions) or experiments that cannot be reproduced (e.g., natural phenomena such as earthquakes).

All this shows once more that observational arrangements should be planned beforehand. This also concerns seismic networks.

A design in the theory of experiment planning is the set of quantities

$$\xi_n = \left\{ \begin{array}{l} \vec{x}_1, \vec{x}_2, \dots, \vec{x}_n \\ r_1, r_2, \dots, r_n \end{array} \right\}, \quad \sum_{i=1}^n r_i = N,$$

where r_i is the number of measurements at the point $\vec{x}_i \in R$, N is the total number of measurements, and R is a region of \vec{x}_i . The points \vec{x}_i are the reference points of the design, their set being the spectrum of the design ξ_n .

The processing of seismic data is concerned with a spectrum belonging to a wide region R in which the assumption of homogeneous experimental conditions is often inadmissible. In such cases an efficiency function $\lambda(\vec{x})$ should be constructed that would make it possible to compare error variances at the points of the spectrum $\lambda(\vec{x}) = \sigma^{-2}(\vec{x})$.

Definition 3. A design of an optimum seismic network is defined here as a network having a fixed constant list of wave types recorded at each site.

For example,

$$\xi_n = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n; \vec{P}, \vec{S}\}. \quad (12)$$

Definition 4. The SN design problem is to find an optimum design that satisfies one of the criteria of optimal planning [9].

The quality of a design is determined by applying certain criteria to it. These can be classified into statistical and nonstatistical criteria. The former generally incorporate the random errors of the measurements and the model; the latter deal with systematic errors, although there are some statistical

criteria that can incorporate systematic model errors as well. The statistical criteria are usually related to the properties of the covariance matrix $D(\xi, \vec{\theta})$ of unknown parameter estimates. Some of them can be reworded in terms of the data matrix $M(\xi, \vec{\theta})$ by virtue of (8). In [9, 10], three statistical criteria (A-, D-, and E-) and one more nonstatistical C-criterion of an optimum design are described. They are used in seismology [15–17, 19].

In [9, 10], the above criteria are related in one way or another to the matrix of the linearized equations of condition $J(\xi, \vec{\theta})$ (4). It would then be reasonable to carry out a theoretical search for designs that minimize a given function Ψ of matrix (4). As has been mentioned above, this optimization problem is the problem of optimal planning of seismic networks, the results of its solution being optimum designs (optimum seismic networks), and the function Ψ being the optimality criterion.

3. Peculiarities and specific formulations of seismic network design problems

As has been mentioned, the estimation of hypocenter parameters and the designing of seismic networks have difficulties resulting from the fact that the relevant regression function is nonlinear in the parameters to be estimated. In this case [12], optimum designs are functions of estimated hypocenter parameters and can consequently be called locally optimal, e.g., locally D-optimal. They can be much easier arrived at, when the region of unknown parameters can be fixed beforehand. This can be done in most of the seismological problems. These cases can be treated by Bayes or minimax optimum designs [12]. Suppose the estimated parameters $\vec{\theta}$ are known to belong to some set Ω . Then

$$\xi^* = \arg \inf_{\xi} \sup_{\vec{\theta} \in \Omega} \Psi[J(\xi, \vec{\theta})] \quad (13)$$

is called a minimax optimum design. If we also know the distribution function for the earthquake parameters $F(\vec{\theta})$, then we can consider the so-called Bayes optimum design

$$\xi^* = \arg \inf_{\xi} \int_{\Omega} \Psi[J(\xi, \vec{\theta})] dF(\vec{\theta}). \quad (14)$$

Because of a number of factors, such as geological, geographic, and economic, *the SN design problem is also essentially discrete*, the economic factor operating to reduce the number of stations (points in the design spectrum) to a minimum. It is known from the theory of experiment designs that it becomes necessary *in such cases to find exact optimum designs, i.e., designs that are optimal for a fixed number of network stations n* . A search for exact

optimum (discrete) designs is more difficult than a search for so-called continuous optimum designs. The problem is further complicated by the fact that a solution is to be sought for each n [12].

The design of seismic networks in many seismological problems can be stated in either of the following formulations [9, 10]:

1. A region has a network of $k > m$, where m is the number of unknown hypocenter parameters. It is required to select an optimum subnetwork of $k_1 < k$ stations for recording the earthquakes occurring in a given source zone (or zones).
2. A region has a network of $k \geq m$ stations, equipment being available (expected or planned) for k_2 stations more. There are $l > k_2$ sites in the region chosen from geological and geographical considerations for the deployment of this equipment. It is required to supplement the existing network in an optimum manner to have $k + k_2$ stations for recording the earthquakes occurring in a given source zone (zones).
3. A region has no seismic network. It is required to plan (design) an optimum network consisting of a fixed number $k \geq m$ of stations for recording the earthquakes occurring in a fixed source zone (zones). The new network can be designed both using the sites that have been chosen beforehand from geological and geographical considerations and the sites that cover the region uniformly.

Each of the above problems divides in its turn into several subproblems, depending on what set of parameters is planned to be determined using the future network.

4. Numerical algorithm construction of exact optimum designs

In the above discussion of special features inherent in the design of SN, we pointed out that:

- (a) optimum designs are locally optimal;
- (b) when the distribution of hypocenter parameters is known, the Bayes optimum designs are to be considered (14), otherwise minimax optimum designs are used (13);
- (c) the design of SN is the construction of exact optimum designs in the first place.

The construction of an exact optimum design consists in finding the least value (global minima) of the function

$$\sup_{\vec{\theta} \in \Omega} \Psi[J(\xi, \vec{\theta})] \quad (15)$$

or

$$\int_{\Omega} \Psi[J(\xi, \vec{\theta})] dF(\vec{\theta}) \quad (16)$$

in a discrete region of a design spectrum modification for a fixed number of spectrum points.

Various techniques are available for the optimization of functions of several variables. They can be classified into two large sets, those based on the use of derivatives and those without derivatives. Only some of them are designed to find the global minimum, the others are intended for finding local minima. The techniques without derivatives are a straightforward direct and random search, a simplex method, etc. The techniques that use derivatives include various modifications of Newton's method, gradient techniques, etc. Techniques for finding local minima are used to find all of them and then select one global minimum. Although there are many satisfactory general-purpose techniques, the theory and practice of optimization suggest that it is better to use specially designed techniques when dealing with a specific problem rather than to have recourse to good general ones.

In the present paper, the authors propose an algorithm of constructing the precise optimum designs, whose idea belongs to Fedorov [12].

Let a design $\xi_n(\vec{\theta}) = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \dots, \vec{x}_n; \vec{P}\}$ be selected as the initial approximation. Let us substitute instead of one of basic points \vec{x}_i a point $\vec{x} \in R$, then we shall obtain the design $\xi_n^*(\vec{\theta}) = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}, \dots, \vec{x}_n; \vec{P}\}$. Thus, the function $\Psi[J(\xi, \vec{\theta})]$ varies by a certain value

$$\Delta(\vec{x}_i, \vec{x}) = \Psi[J(\xi_n^*, \vec{\theta})] - \Psi[J(\xi_n, \vec{\theta})].$$

If at such a substitution the function Ψ is diminished, i.e., $\Delta(\vec{x}_i, \vec{x}) < 0$, it is reasonable to change the design ξ_n for the design ξ_n^* . The diminution Ψ can be made even larger at the expense of the choice of such a point \vec{x} for which $|\Delta(\vec{x}_i, \vec{x})|$ becomes maximum (at $\Delta < 0$). An additional maximization of $|\Delta(\vec{x}_i, \vec{x})|$ with respect to \vec{x}_i at $\Delta < 0$ will diminish Ψ even greater.

Let us now formulate the algorithm:

- 0) we select an arbitrary design $\xi_n^{(s)} = \{\vec{x}_1^{(s)}, \vec{x}_2^{(s)}, \dots, \vec{x}_n^{(s)}; \vec{P}\}$ as the initial approximation;
- 1) we discover a pair of points $\vec{x}_i^{(s)}, \vec{x}^{(s)}$, for which $\max_i \max_{\vec{x}} |\Delta(\vec{x}_i, \vec{x})|$ is reached at $\Delta(\vec{x}_i, \vec{x}) < 0$;
- 2) a new design $\xi_n^{(s+1)}$ is distinguished from $\xi_n^{(s)}$ by changing the point $\vec{x}_i^{(s)}$ to the point $\vec{x}^{(s)}$ indicated as $\vec{x}_i^{(s+1)}$ in future.

Procedure 1), 2) multiply repeated, generates a monotonically decreasing sequence $\Psi[J(\xi_n^{(s)}, \vec{\theta})]$, whose convergence follows from the existence of the lower boundary. In this case, an indication to attaining a minimum point is the fulfillment of the condition $\Delta(\vec{x}_i, \vec{x}) = 0$.

A minute shortage of the proposed algorithm is the fact that it does not always converge to the optimum design, i.e., the sequence $\Psi[J(\xi_n^{(s)}, \vec{\theta})]$ does not always converge to a global minimum. In this connection, it is recommended to iterate the above described process several times, starting with various designs $\xi_n^{(0)}$. If in all the cases, the value of the criterion for the obtained designs coincide, the latter with high probability are the desired precise optimum designs (they can be various or coincide). If repetitions of the described process result in the various designs with different values of the criterion, it is recommended to continue attempts of searching for the optimum designs until a group of the designs is not formed, for which the values of the criterion are equal to one another and than all the rest (designs of a group can be various or coincide).

An important part of the considered algorithm is the calculation of values of the criterion of optimality $\Psi[J(\xi_n^{(s)}, \vec{\theta})]$, which is in application of singular expansion (10) of the matrix $J(\xi_n^{(s)}, \vec{\theta})$. It enables us to easily and simply calculate values of any of D-, A-, E-, or C-criteria of optimality, and also of any of their combination. A possibility to incidentally calculate, without essential additional time computer costs any combination of the enumerated criteria of optimality will enable us, if necessary, to build the multicriteria optimum designs.

5. Conclusion

The proposed numerical algorithm of designing seismic networks has been realized by the authors as program in the Fortran language. This allows one by the given criterion of optimality to discover optimum networks in any of the three, above described settings. In this case, networks are optimized with respect to an arbitrary 3D hypocenter area concerning arbitrary three-dimensional hypocenter area or several areas at an arbitrary number n of stations in the network.

The program favorably differs in the operation speed from the similar ones, based on Monte Carlo method. The gained experience of the usage of the program demonstrates that in most cases, with various initial designs, the iterations rapidly converge to one of two–three (more often two) different minima of the function Ψ , from which the user selects a global one.

The program has been tested on designing a large number of real and virtual networks, some of which are described in [9, 10]. The authors have carried out the works on the design of seismic networks in the majority of

Siberian and the Far East seismic-prone regions, and also of Global networks of the USSR and Russia.

References

- [1] Burmin V.Yu. The problem the experiment design and condition of a system of linear algebraic equations // *Tekhnicheskaya Kibernetika*. – 1976. – № 2. – P. 195–200 (in Russian).
- [2] Burmin V.Yu. Algorithm of definition of coordinates of the close earthquakes hypocenters and velocity of propagation of seismic waves in a stratum // *Volcanology and Seismology*. – 1985. – № 5 (cover-to-cover translation).
- [3] Ermakov S.M. *The Mathematical Theory of Experimental Designs*. – Moscow: Nauka, 1983 (in Russian).
- [4] Ermakov S.M., Pankratiev I.D. Displaced estimations and the method of regularization // *Vestnik LGU*. – 1976. – Vol. 2, № 7. – P. 27–30 (in Russian).
- [5] Kondorskaya N.V., Aranovich Z.I. Methodical basis of optimization of seismic observation // *Fizika Zemli*. – 1971. – № 7. – P. 14–30 (in Russian).
- [6] Marchuk G.I. *Methods of Computational Mathematics*. – Moscow: Nauka, 1977 (in Russian).
- [7] Nikolaev A.V. Optimization of seismic networks and automation of seismic observations. – Tbilisi: Metsniereba, 1977. – P. 81–89 (in Russian).
- [8] Omelchenko O.K., Beloborodov N.N. Modeling of seismic wave fields // *Mathematical Problems of Geophysics*. – Novosibirsk: VTs SO AN SSSR, 1983. – P. 71–81 (in Russian).
- [9] Omelchenko O.K., Beloborodov V.N. On optimal design of a regional seismic network. – Novosibirsk, 1986. – (Preprint / VTs SO AN SSSR; 668) (in Russian).
- [10] Omelchenko O.K., Gusakov V.K. Planning of seismic networks for tsunami prediction // *Volcanology and Seismology*. – 1996. – Vol. 18. – P. 211–230 (in Russian).
- [11] Ortega J.M., Rheinboldt W.C. *Iterative Solution of Nonlinear Equations in Several Variables*. – New York and London, 1970.
- [12] Fedorov V.V. *The Theory of the Optimum Experiment*. – Moscow: Nauka, 1971 (in Russian).
- [13] Forsythe G.E., Malcolm M.A., Moler C.B. *Computer Methods for Mathematical Computations*. – Englewood Cliffs, N.J., 1977.
- [14] Hatton L., Worthington M., Makin J. *Seismic Data Processing. Theory and Practice*. – Oxford: Blackwell, 1986.

- [15] Flinn E.A. Confidence regions and error determination for seismic event location // *Rev. Geophys.* – 1965. – Vol. 3, № 1. – P. 157–185.
- [16] Kijko A. Methods of the optimal planning of regional seismic networks / Polish Academy of Sciences. Institute of Geophysics. – Warszawa-Lodz: Panstwowe wydawnictwo naukowe, 1978. – A-7(119).
- [17] Kijko A. Optimum seismic network in the Ostravacoal basin // *Acta Montana UGG.* – Praha, Czechoslovakia: CSAV, 1980. – 55. – S. 73–95.
- [18] Uhrhammer R.A. Analysis of small seismographic station networks // *Bull. Seism. Soc. Am.* – 1980. – Vol. 70, № 4, P. 1369–1379.
- [19] Uhrhammer R.A. The optimal estimation of earthquake parameters // *Physics of the Earth and Planetary Interiors.* – 1982. – Vol. 30. – P. 105–118.