UDC 550.834

V. S. Mostovoy, S. V. Mostovyi

Estimation of the parameters of seismic waves

(Presented by Academician of the NAS of Ukraine V. I. Starostenko)

We consider a problem of estimation of the parameters of a seismic signal based on real observed data. For this purpose, we propose a new mathematical model that reduces the task to a nonlinear, nonsmooth, nonconvex minimization problem. Using a special structure of the signal, we also propose a numerical algorithm of finding the solution of the optimization problem. Our method is based on the combination of the Levenberg–Marquardt algorithm and a simulated annealing type approach. We show the convergence of the algorithm and discuss the practical implications, which lie in a good compatibility with real seismic data, and the applicability to experiments.

Introduction. One of the central problems in the seismic signal processing is the determination of parameters of a signal such as the damping decrement, the principal eigenfrequency, the moment of arrival, etc. Usually, seismic signals such as compression plane waves, distortional waves, surface waves, Rayleigh waves, ground roll modes, and Love waves are modeled via physically realizable signals, see [2, 14, 15], that are characterized by the spectral band. It is natural to approximate such signals by a superposition of Berlage impulses, as a generalization of the model of simple oscillators. Such kind of modeling is widely used in seismology. With the goal to find an optimal approach to the estimation of the "key" seismic parameters, it might be reasonable to restrict ourself to a set of mathematical models of the seismic signals that are used in practice. One way of doing this is to use Berlage impulses that have enough degrees of freedom for the approximation of a wide class of seismic signal forms in comparison to other types of signals, see, e. g., [2]. In this case, the form of a signal is defined by a 5-dimensional vector of free parameters. In Section 2, we propose a novel approach that is based on an enlargement of the vector of free parameters to a 6-dimensional vector. This gives more flexibility and allows one to consider different models in one formulation. By choosing some of the free parameters to be zero or one, we can include almost every practically interesting case in our model and, in particular, to recover the original 5-dimensional models. Our approach leads to a minimization problem, a solution to which gives the optimal set of "key" parameters in the sense explained below. We also propose a numerical algorithm for solving this optimization problem, convergence properties of which are established and are discussed as well.

One of the merits of our method is in the balance between the ability to well approximate the desired set of parameters of the signals (from the practical side) and the analytical tractability (from the theoretical viewpoint). In practice, our model was successfully implemented in multiple experiments such as the oil and gas detection and the monitoring of natural and man-made objects, see [10, 11, 12] for details. From the theoretical point of view, first, it it proven in [12] that the solution to our central optimization problem (5) exists. Second, since the objective function in the minimization problem (5) is nonlinear, nonsmooth, nonconvex, and multidimensional, the delicate optimization techniques are required to construct an approximation to the solution to (5). Thus, we propose a numerical algorithm of solving (5), the efficiency of which is based on *local* continuity, boundedness, and differentiability of the objective. Behind our numerical method are the well-known ideas of nonlinear optimization such as the simulated annealing and the Levenberg–Marquardt algorithm, see [7, 8, 13]. Our approach turns out to be convergent to a true solution (in probability with respect to a certain probability measure introduced in Section 4), as the running time increases. Convergence results were established in [12]. Finally, it should be mentioned that our method is well-suited for parallel computing, and the running time of numerical evaluations is of order of minutes on a standard lap top, which is convenient in experiments.

The remainder of the paper is organized as follows: in Section 2, we describe the model of signal, Section 3 contains the formulation of the central optimization problem (5), a numerical algorithm of solution to which is presented in Section 4. Conclusion is stated in Section 5.

2. Mathematical model of seismic activity process. Let us consider a general representation of the impulses used to model many seismic signals, which are often used in seismology. Free parameters of such impulses define different models. Let us define the function

$$S\colon \mathbb{R}\times[-A,A]\times\mathbb{R}_+\times\mathbb{N}\bigcup\{0\}\times\mathbb{R}_+\times\mathbb{R}_+\times[\Omega_1,\Omega_2]\times\mathbb{R}_+\to\mathbb{R}$$

given by

$$S(t,a,\alpha,\beta,\omega,\tau,T)=aI_{\tau,\tau+T}(t)\exp\{-\alpha(t-\tau)\}(t-\tau)^{\beta}\sin(\omega(t-\tau)), \tag{1}$$

where the meaning of all the parameters is described below. From the physical point of view, all the parameters in (1) are physically substantial ones. The meaning of the symbols in the previous expression is the following: t stands for the time argument, the other variables are treated as the free parameters of the model. The parameter a denotes the amplitude of the signal. It is the only parameter that enters model (1) linearly, whereas all other parameters enter (1) nonlinearly. The first nonlinear parameter α is the damping characteristic of oscillations. The next parameter β is used for a correction of the impulse front. This parameter gives us the possibility to regulate the steepness of the pulse edge. Parameter τ is used for physically realizable impulses as the time characteristic of the signal appearance (the left endpoint of the indicator interval of the signal). Parameter ω characterizes the angular frequency of the impulse oscillation. Parameter T characterizes the length of the interval, where the signal exists ($\tau+T$ is the right endpoint of the indicator interval). We restrict the admissible set of amplitudes and frequencies to the intervals $[-A,A]$ and $[\Omega_1,\Omega_2]$, respectively, where A, Ω_1, and Ω_2 are some positive constants. In (1), $I_{\tau,\tau+T}(t)$ denotes the indicator function of the interval $[\tau,\tau+T]$, i. e.

$$I_{\tau,\tau+T}(t)=\begin{cases}1, & \text{if}\quad t\in[\tau,\tau+T],\\ 0, & \text{if}\quad t\in(-\infty,\tau)\bigcup(\tau+T,+\infty).\end{cases}$$

By changing the parameters τ and T, we have a possibility to restrict the signal to the interval between its appearance and disappearance. We will call such impulse (1) *a generalized Berlage impulse* with the linear parameter a and five nonlinear other parameters, as oppose to the Berlage impulse defined, e. g., in [2], which is a particular case of the generalized Berlage impulse when the parameters $\beta=0$ and $T=\infty$. We also need to introduce the vector $\mathbf{P}$ of the free parameters of model (1). Its representation is given by

$$\mathbf{P}=\{P_k\};\qquad k=\overline{1,6};\qquad \{P_1=a,P_2=\alpha,P_3=\beta,P_4=\tau,P_5=\omega,P_6=T\}. \tag{2}$$

In view of (2), we can rewrite (1) as follows:

$$S(t, \mathbf{P}) = P_1 I_{P_5, P_5+P_6}(t) \exp\{-P_2(t - P_5)\}(t - P_5)^{P_3} \sin(P_4(t - P_5)). \tag{3}$$

Fixing certain components of the vector $\mathbf{P}$ of the seismic signal model, one can obtain a restriction of the model of seismic signal. Restricted signals will be determined by a smaller number of free parameters and will have more specific properties. In order to set up the optimization problem (5) and to construct a numerical algorithm of its solution, it is convenient to arrange such kind of vectors into a matrix. This matrix will determine the set of different models of seismic signals. The index of every model will be the same as the column number of this matrix. In such a case, we will get a possibility to model the recorded data via a *superposition* of signals, see (4) below.

Let us demonstrate how certain signals widely used in seismic practice can be obtained from our model (1) by changing the parameters of the model that are encapsulated in vector $\mathbf{P}$. For example, when $P_3 = 1$, we get a model of approximated Berlage impulse in the interval $[P_5, P_5 + P_6]$:

$$S(t, \mathbf{P}) = P_1 I_{P_5, P_5+P_6}(t) \exp\{-P_2(t - P_5)\} \sin(P_4(t - P_5)).$$

When $P_6 = \infty$, the indicator of the interval $I_{P_5, P_5+P_6}(t)$ becomes the Heaviside step function $\eta(t - \tau)$, where

$$\eta(x) = \begin{cases} 1, & \text{if} \quad x \geqslant 0, \\ 0, & \text{if} \quad x < 0, \end{cases}$$

and we get a Berlage impulse with the linear parameter P_1 denoting its amplitude. If P_3 is equal zero and $P_6 = \infty$, expression (3) models a fading sinusoid

$$S(t, \mathbf{P}) = P_1 \eta(t - P_5) \exp\{-P_2(t - P_5)\} \sin(P_4(t - P_5)),$$

which can be represented in terms of widely used symbols in physics as follows:

$$S(t, a, \alpha, \omega, \tau) = a\eta(t - \tau) \exp\{-\alpha(t - \tau)\} \sin(\omega(t - \tau)).$$

If $P_2 = P_3 = 0$ and $t \in [P_5, P_5 + P_6]$, the impulse will turn into the first Fourier harmonic on the interval $[P_5, P_5 + P_6]$ with a period P_6:

$$S(t, \mathbf{P}) = P_1 I_{P_5, P_5+P_6}(t) \sin(P_4(t - P_5)).$$

The natural generalization, which allows us to consider more complicated signals, is based on a superposition of signals (3). In this case, we can define a (more general) signal $\mathbb{S}(t, \mathbf{M})$ as

$$\mathbb{S}(t, \mathbf{M}) = \sum_{q=1}^{Q} S(t, \mathbf{P}_q), \tag{4}$$

where the matrix $\mathbf{M}$ consists of the column vectors $\mathbf{P}_q$, $q = \overline{1, Q}$, and contains all free parameters of our model. The signal $S(t, \mathbf{P}_q)$ with a specified vector $\mathbf{P}_q$ is called the q-th submodel. The set of the admissible values for $\mathbf{M}$ is denoted by $\mathcal{A}$, which is uniquely defined by the domain of S. Finally, we are ready to state the model of signal that we used for applications and that is considered in the following sections. For every seismic record $y(t)$, $t \in [0, \mathcal{T}]$, where $\mathcal{T}$ is the length of the seismic record, we suppose that $y(t)$ is the sum of a seismic signal of the form (4) and an additive background noise $n(t)$,

$$y(t) = \mathbb{S}(t, \mathbf{M}) + n(t), \qquad t \in [0, \mathcal{T}].$$

3. Optimization problem. The mathematical problem consists of an estimation of the matrix of free parameters $\mathbf{M}$ corresponding to a recorded seismogram $y(t)$, $t \in [0, \mathcal{T}]$, and some statistical characteristics of the background noise n. In the simplest model of the noise n, one can consider an additional condition that the background noise is uncorrelated, i. e. its autocorrelation is the δ-function. We assume that the *a priori* distribution of $\mathbf{M}$ is a uniform distribution over $\mathcal{A}$ and use the goodness-of-fit test for $\mathbb{F}(\mathbf{M})$ to set up an optimization problem, a solution to which we will call the optimal value of $\mathbf{M}$, where $\mathbb{F}$ is defined as

$$\mathbb{F}(\mathbf{M}) = \int_0^{\mathcal{T}} (y(t) - \mathbb{S}(t, \mathbf{M}))^2 dt, \qquad \mathbf{M} \in \mathcal{A}.$$

This leads to the optimization problem

$$\inf_{\mathbf{M} \in \mathcal{A}} \mathbb{F}(\mathbf{M}), \tag{5}$$

where $\mathcal{A}$ is the set of admissible values of $\mathbf{M}$ defined above. When the objective is continuous, one can see that problem (5) is the Mayer problem in the calculus of variations, see [3, 5]. Therefore, variational methods are natural candidates for the analysis of (5).

Since the primal goal of this work is to construct a model and a mathematical approach for the *practical* purposes, it is natural to satisfy ourself with an *approximate* solution to (5) with a given tolerance $\varepsilon > 0$, i. e. by calculating a value $\widehat{\mathbf{M}}$ such that

$$\mathbb{F}(\widehat{\mathbf{M}}) - \inf_{\mathbf{M} \in \mathcal{A}} \mathbb{F}(\mathbf{M}) \leqslant \varepsilon.$$

Such a value $\widehat{\mathbf{M}}$ will be called a solution to (5).

One drawback of the variational approach is in the requirement of continuity (or even smoothness) of the objective $\mathbb{F}$ that is often not available in practice. In particular, in our case, impulse (1) includes the discontinuous term I corresponding to the step function, see (3). Therefore one has to use the alternative techniques of nonlinear optimization to study (5). Additional challenges lie in the nonconvexity and the high nonlinearity of $\mathbb{F}$. Of course, one can approximate $\mathbb{F}$ with a smooth function, e. g. via mollifiers, see [4], and to study the corresponding smooth problem using variational techniques. However, since mollification does not remedy nonlinearity, one still has to apply some delicate methods of nonlinear optimization, see, e. g., [1].

Solution to the optimization problem (5). The main difficulty in the construction of an effective numerical algorithm for finding a solution to (5) is the *discontinuity* of the objective $\mathbb{F}$ explained above. Instead of mollifying the objective, we propose a different approach that is convergent (to the true solution of (5)) and analytically tractable in the sense explained below.

Even though our objective $\mathbb{F}$ is not continuous everywhere in its domain, it is (infinitely many times) differentiable by Lebesgue almost everywhere in $\mathcal{A}$, except for some set of discontinuities of (Lebesgue) measure zero. Moreover, $\mathbb{F}$ does not show a pathological behavior at discontinuities, outside of which it stays locally smooth and uniformly bounded on compact sets. Therefore, we can use optimization techniques to find the **local** minima. Numerically, we use the Levenberg–Marquardt algorithm for this purpose, see [7, 8, 13]. Having the procedure of finding the local minima, we can hope to recover the global minimum by repeating the Levenberg–Marquardt procedure many times starting from different (randomly chosen) initial points. The reader can see the intimate relation to the *simulated annealing* and *Metropolis–Hastings* algorithms, see [6, 9]. In our settings, the problem is studied in [12], and the convergence in probability (for a given tolerance, with respect to a probability measure on $\mathcal{A}$ that is used for the generation of starting points) is proven there. Moreover, we can even show the almost sure convergence if the set $\mathcal{A}$ is assumed to be compact (this is a reasonable and non-restrictive assumption that often holds in practice, if we have some *a priori* information about the possible ranges of the components of $\mathbf{M}$). Thus, if we run our algorithm for a sufficiently long time, the estimate very close to the true value of a minimizer to (5) will be found. The description of the applications of our method to various practical problems is contained in [12].

To guarantee the convergence of the algorithm in probability, *any* probability distribution over $\mathcal{A}$, such that the Lebesgue measure on $\mathcal{A}$ is absolutely continuous in the probability measure, i. e. every subset of $\mathcal{A}$ with a positive Lebesgue measure has positive probability, could be used for the generation of starting points. From the practical point of view, our algorithm is tailor made for parallel computing. It converges fast enough to be used on a usual computer, since the Levenberg–Marquardt procedure is extremely efficient for minimizing twice differentiable functions, which is almost our case, except for the fact that $\mathbb{F}$ is differentiable (infinitely many times) locally, not globally.

Conclusion. We proposed a new model of seismic data and an algorithm of finding the optimal values of the seismic parameters of special interest, such as the damping decrement, the principal eigenfrequency, and the moment of arrival. Our model is a generalization of the classical seismic models in [2]. The determination of the optimal values of free parameters of the model is based on a criterion, which we aim to minimize. This leads to a nontrivial optimization problem (5), a solution to which can be obtained via a numerical algorithm we constructed as well. In turn, our numerical algorithm is studied both theoretically (here, the existence of a solution and the convergence of the algorithm are proven) and practically in the numerical experiments (see examples in [10]).

1. *Bertsekas D. P.* Nonlinear programming. – Belmont, MA: Athena Sci., 1999. – 201 p.
2. *Berzon I. S., Epinat'eva A. M., Pariiskaya G. N., Starodubrovskaya S. P.* Dynamical characteristics of seismic waves in real media. – Moscow: Izd. AN SSSR, 1962. – (in Russian).
3. *Bliss G. A.* Lectures on the calculus of variations. – Chicago: Univ. of Chicago, 1947. – 291 p.
4. *Evans L. C.* Partial differential equations. – Providence, RI: Amer. Math. Soc., 1997. – 664 p.
5. *Fleming W. R., Rieshel R. W.* Deterministic and stochastic optimal control. – New York: Springer, 1975. – 222 p.
6. *Kirkpatrick S., Gelatt C., Vecchi M.* Optimization by simulated annealing // Science. – 1983. – **220**. – P. 671–680.
7. *Levenberg K.* A method for the solution of certain non-linear problems in least squares // Quart. Appl. Math. – 1944. – **2**. – P. 164–168.
8. *Marquardt D. W.* An algorithm for least-squares estimation of nonlinear parameters // J. Soc. Indust. Appl. Math. – 1963. – **2**, No 2. – P. 431–441.
9. *Metropolis N., Rosenbluth A., Teller A., Teller E.* Equations of state calculations by fast computing machines // J. Chem. Phys. – 1953. – **21**, No 6. – P. 1087–1092.
10. *Mostovyi S. V., Mostovoy V. S.* A variational method to the solution of an inverse problem at the accumulation of seismic signals in the active monitoring // Dopov. NAN Ukr. – 2008. – No 8. – P. 113–116.

11. *Mostovyi S. V., Starostenko V. I.* Interpretation of geophysical data under uncertainty information // Izv. AN SSSR. Fiz. Zem. – 1987. – No 5. – P. 31–40.
12. *Mostovoy V. S.* Models of systems of monitoring of geophysical fields (Doctoral Dissertation, Phys.-Math. Sci.). – Kyiv: Kyiv Univ., 2013 (in Russian).
13. *Pujol J.* The solution of nonlinear inverse problems and the Levenberg–Marquardt method // Geophys. – 2007. – **72**, No 4. – P. W1-W16.
14. *Ricker N.* The form and laws of propagation of seismic wavelets // Ibid. – 1953. – **18**. – P. 10–40.
15. *Robinson E.* Predictive decomposition of time series with application to seismic exploration // Ibid. – 1967. – **32**, No 3. – P. 418–484.

Institute of Geophysics of the NAS of Ukraine, Kiev
National Aviation University, Kiev

Received 19.08.2013

В. С. Мостовий, С. В. Мостовий

Оцінка параметрів сейсмічних хвиль

Розглядається задача оцінки параметрів сейсмічного сигналу, що базується на реальних спостережених даних. Для цього пропонується нова математична модель, яка зводить задачу до нелінійної негладкої задачі неопуклої мінімізації. Пропонується чисельний алгоритм знаходження розв'язку задачі оптимизації, що враховує специфіку структури сигналу. Метод заснований на комбінації алгоритму Левенберга–Марквардта и метода симуляції аннілінга. Показано збіжність алгоритму, його практичне застосування та хорошу сумісність моделі з сейсмічними экспериментальными даними.

В. С. Мостовой, С. В. Мостовой

Оценка параметров сейсмических волн

Рассматривается задача оценки параметров сейсмического сигнала, основанная на реальных наблюденных данных. Для этого предлагается новая математическая модель, которая сводит задачу к нелинейной, негладкой задаче невыпуклой минимизации. Предлагается численный алгоритм нахождения решения задачи оптимизации, учитывающий специфику структуры сигнала. Метод основан на комбинации алгоритма Левенберга–Марквардта и метода симуляции аннилинга. Показаны сходимость алгоритма, его практическое применение и хорошая совместимость модели с сейсмическими экспериментальными данными.