Bull. Nov. Comp. Center, Math. Model. in Geoph., 7 (2002), 69–75

Using of nonlinear regression with fuzzy input data for analysis of seismisity*

K.V. Simonov, An.G. Marchuk, V.A. Okhonin, A.L. Shchemel

A variant of modified training functional that allows considering fuzzy input data is suggested. A limiting case when a part of input data is completely undefined, and, therefore, a problem of reconstruction of hidden parameters should be solved, is also considered. Some numerical experiments are presented. The suggested approach yields satisfactory results, which has been already used in a number of applications.

1. Propounding of the problem

It is assumed that a dependence of known input variables upon output ones should be found in the classic problem definition, which is widely used in the majority of neural nets algorithms. The quality of approximation is evaluated as a performance function

$$H = \sum_t h_t(x, A_t), \tag{1}$$

where h_t is the error of the task number t, x is a set of tunable regression parameters, and A_t is an input data of the task t. Often the error of the task t is evaluated as squared distance between known input data $\tilde{\alpha}_t$ and predicted data $\alpha_t(x, A_t)$:

$$h_t = (\alpha_t(x, A_t) - \tilde{\alpha}_t)\varepsilon(\alpha_t(x, A_t) - \tilde{\alpha}_t). \tag{2}$$

The matrix ε in expression (2) consists of the coefficients of positive determined squared form. If this matrix is diagonal, then the coefficients are reciprocal to permissible mean squared deviation between predicted and known outputs. Coefficients ε may be named "precision coefficients". When the outputs are not known exactly, natural generalization of (1) is as follows:

$$H' = H + \sum_t (A_t - \tilde{A}_t)\mu(A_t - \tilde{A}_t), \tag{3}$$

where known input precision coefficients μ and input parameters $\tilde{A}_t$ are introduced, which do not coincide with the best predicted ones A_t defined via optimization of performance function H' with these parameters.

*Supported by the Russian Foundation for Basic Research under Grant 99-05-65372.

The limiting case when some or all inputs are undefined is to be considered separately.

2. Regularization and definition of the realization of the neural net

It is desirable that the set of tunable parameters was deliberately compact for training to be converging. In problem (2), it is possible to choose variants of demands of a priori compactness, which allow meaningful interpretation in the smoothness of the dependence α_t upon A_t. For example, two variants of constraints can be used:

$$\langle (\nabla a_t)^2 \rangle \le \text{const}; \tag{4}$$

$$\frac{\langle (\nabla a_t)^2 \rangle}{\langle a_t^2 \rangle} \le \text{const}. \tag{5}$$

These inequations limit mean squares of derivatives of input variables (4) or relations of these mean squares to mean squared inputs (5) in inputs' space.

The following form of neural net approximation of dependence between outputs and inputs was used:

$$a_t^i = b^i + c^i \sum_q \sin\Big(\varphi^{iq} + \sum_j k^{qj} A_t^j\Big). \tag{6}$$

If such approximations were chosen, the case of compactness condition (4) would be expressed as

$$\Big(\sum_i (c^i)^2\Big) \times \sum_{q,j} (k^{qj})^2 \le \text{const}, \tag{7}$$

and the case (5) as

$$\sum_{q,j} (k^{qj})^2 \le \text{const}. \tag{8}$$

Asymptotic universality of neural net (ability to approximate various smooth functions with any accuracy by increase of the number of tunable parameters) is often the base for selecting a type of neural net approximation [1, 2]. It is possible to show that neural net (6) will approach to the Fourie integral transform (whose approximate abilities are known) with increasing of the number of tunable parameters.

Coded neural net (6) demonstrates high performance; it can be trained quicker than many others. The ability to calculate outputs together with their derivatives of different ranks along inputs fast and easy is significant for some applications.

A case of conjugate gradient algorithm was used to solve problem (2) modified with taking into account one of conditions (7) or (8). Programs of data processing were coded as additional functions for "Matlab" and "Excel". It has been used to solve a set of applied problems.

According to obtained results, both ways of regularization (7) and (8) are satisfied. The solution of problem (3) was obtained in two ways: a case when conjugate gradients algorithm optimized all tunable and input parameters simultaneously and a case when optimization of tunable parameters and inputs was performed separately in odd and even iterations.

The second case can be used to adapt every task step-by-step, which could reduce waste of memory. According to the results of the tests, it is possible to reconstruct corrupted inputs in some cases. The results of correction depend on the number of samples.

3. Reconstruction of hidden parameters

In the limiting case, when input data is set with zero precision, the problem of reconstruction of hidden parameters with observed output data appears. The problem can be formally propounded this way:

A distribution $P(a)$ of various combinations of observed values should be estimated. Totality of the combinations is represented by the set of variables a. The results of observations determine the excerpt $\{a_t\}$. An approximation of the distribution is chosen as

$$P(a) = \int \delta(a - a'(A))\rho(A)\, dA, \tag{9}$$

where $\delta(\cdot)$ is the Dirak delta function, $a'(A)$ is the neural net approximating function, $\rho(A)$ is a positive measure in the space of hidden parameters. To avoid a problem of approximation $\rho(A)$ it is natural to assume $\rho(A)$ constant and $a'(A)$ variable.

Expression (9) corresponds to the approximation of the distribution $P(a)$ using special Jacobean of the reflection $a'(A)$ in the case when a dimensionality of the observed parameters a coincides with the dimensionality of hidden ones A and the reflection $a'(A)$ is turnable:

$$P(a') = \frac{\rho(A(a'))}{|Da'(A)/DA|}. \tag{10}$$

In the scope of the propounded problem, a continuous (along with its derivations) homomorphic reflection of the space of hidden parameters to the space of observed parameters should be found. It allows to reconstruct lack information of the inputs when the number of the inputs is not less than the number of hidden parameters and to estimate the distribution if information for synonymous prediction of unknown inputs is not sufficient.

The following approach to build approximation based on the excerpt $\{a_t\}$ is suggested: the excerpt $\{a_t\}$ is complemented with the hidden parameters $\{A_{t'}\}$, which are distributed according to the measure $\rho(A)$.

Then one should find the correspondence of t' and t which provides that the best approximation of a_t with $a'(A_{t'})$ is the most accurate. In the odd iterations, the dependence $a'(A)$ is being optimized (like problem (2) is solved). The correspondence between t' and t is changing in the case when error (2) is reducing and the distribution $\rho(A)$ remains intact.

It is possible to leave the form $\rho(A)$ intact when the Jacobean of the transforms $A \to A'(A)$ satisfies the following condition:

$$\rho(A'(A) \times DA'(A)/DA) = \rho(A). \tag{11}$$

Therefore, a special transform that obeys (11) is applied to reduce error (2) at every iteration. If the measure $\rho(A)$ is constant, then condition (11) is simplified to

$$DA'(A)/DA = 1. \tag{12}$$

Such transforms (12) form the bases of the classic theoretical physics [3], where they are named "canonical" or "volume invariant transforms" and, therefore, are well-known. We can offer one more interpretation of the suggested approach.

Let us consider that we have a set of records and want to number them with a set of rows of indexes and then to create neural net which would continuously reflect indexes to records. Obviously, the problem could be solved only when similar indexes correspond to similar records. The suggested approach builds both neural net and the required numeration.

4. Numerical algorithm of the reconstruction of hidden parameters

The following variant of the problem of the reconstruction of hidden parameters solution was coded (in "Matlab"):

Orthogonal periodic mesh on multidimensional map was considered. Periodic functions "sin" and "cos" with the period of the number of the nodes of the mesh were calculated with the arguments equal to node coordinates. Their values formed a multitude of permissible values of inputs.

Condition (11) was performed due to the transforms were combinations of shifts of the lines formed mesh parallel to itself with the step multiply to the distance between nodes. Neural net approximation is based on (6).

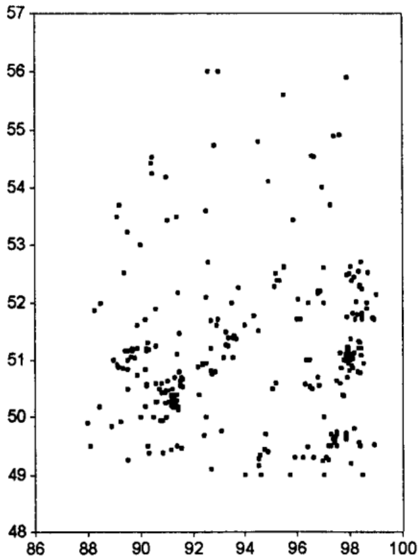

Figure 1. Coordinates of the main 315 earthquakes in Krasnoyarsk region

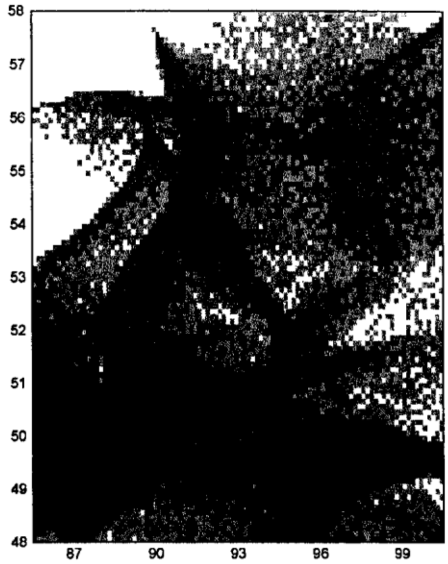

Figure 2. Distribution of logarithms of earthquakes' probabilities

Results of the numerical experiments on the reconstruction of distribution were satisfactory. As an illustration, such example can be quoted: Starting from 1761 year there is a log of earthquakes in Krasnoyarsk region. These 315 earthquakes are shown in Figure 1.

The grade of gray in Figure 2 corresponds to increasing of logarithm of probability in this place (i.e., to increasing probability by 2 times). The location of Krasnoyarsk city on this map is determined by coordinates 56°N, 93°E.

To produce this distribution, the space of hidden parameters was chosen as a cubic mesh constrained with sphere, reflection of this space to the space of real data (which included coordinates and energy class of earthquakes) was made by neural net of 300 neurons. Approximation's accuracy of 0.987% of dispersion was achieved.

Numerical realization brings the following problems, in comparison with the formal propounding:

Discrete analog of condition (11) leads to strict constraints on a number of samples. Training excerpt is conditioned by mesh parameters and is difficult to be modified by inserting or removing some records. Perhaps it is necessary to use other variants of discrete approximations of essentially continuous condition (11) to overcome these difficulties.

Generally, not natural values of dimension of the space of hidden parameters are achieved by estimating the Hauzdorf dimension of the excerpt $\{a_t\}$, which cause difficulties in applying suggested algorithm.

Different limitations of discrete mathematics and logic, geometry with natural dimensions arise during work with complicated information, which is not adapted for our theoretical paradigm. Encountered hardships in the developing of the algorithms show insufficiency of discrete approach.

5. Graphic interface for visualization and analysis of regional seismicity

Now a lot of digital seismic data is collected for different areas of Russia. The graphic interface with digital geographic mapping is needed for effective using of this data for seismic zoning and earthquake prediction. Technology for creating of GIS applications for geophysical investigations was developed in the Institute of Computational Mathematics and Mathematical Geophysics. For instance, on the base of this technology systems for analysis and prognosis of seismic situation of the Krasnoyarsk region.

This system uses vector and raster graphics and geophysical data in specified format. The program of creating pseudo 3D and full 3D images of selected area is developed for visualization of the Earth relief. Then several layers with vector geographic data are drawn on this raster background. Colors and sizes of the elements of geographical map are defined by the

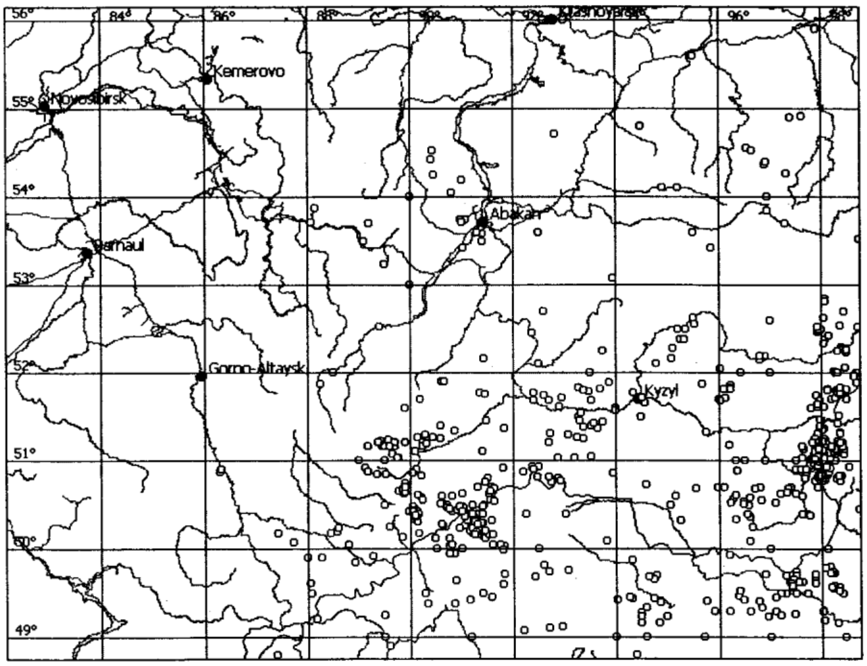

Figure 3. Vector map of the region with epicenters location

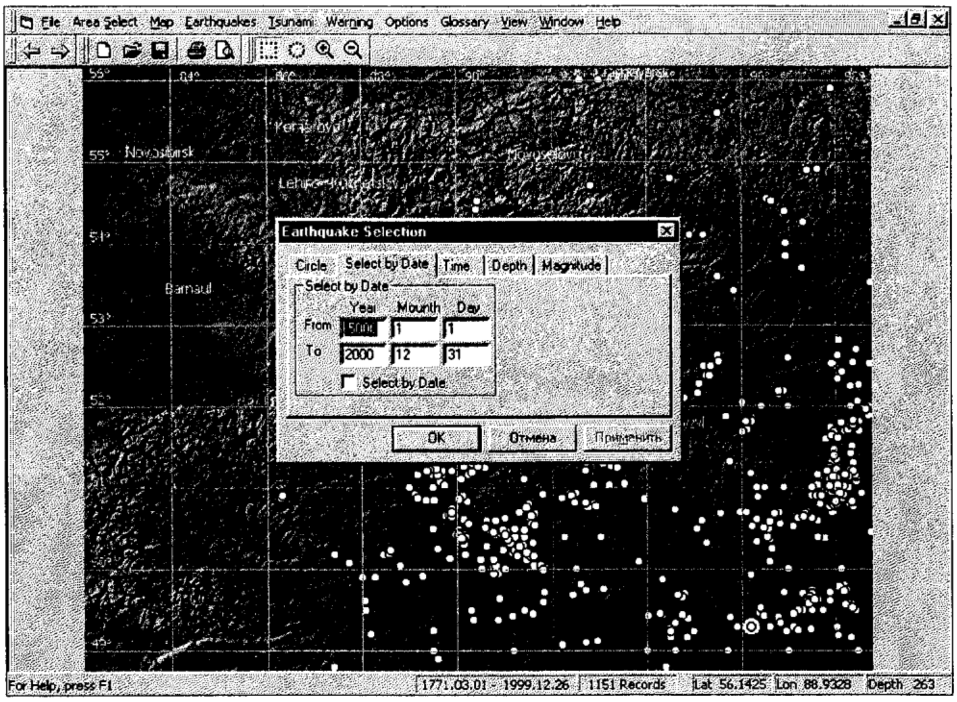

Figure 4. 3D background relief with the event selection window

user. In Figure 3, the vector map (without relief) of South Krasnoyarsk region is shown.

On this background map the epicenters of earthquakes, which were recorded in this region, can be plotted (see Figure 3). From this large region the smaller area for detailed analysis of seismicity can be chosen. The possibility of the land relief visualization and retrieving of seismic events can be very useful for analysis of seismic situation and earthquake prediction in the area of interest (Figure 4). In the center of Figure 4, the event selection window is shown.

References

[1] Hornik K., Stinchombe M., White H. Multilayer feedforward networks are universal approximators // Neural Networks. – 1989. – P. 359–366.

[2] Gilev S.E., Gorban A.N. On completeness of the class of functions computation by neural networks // Proc. of the World Congress on neural networks (WCNN'96). Sept. 15–18, 1996. – San Diego, CA, Lawrens Erlbaum Accotiates, 1966. – P. 984–991.

[3] Synge J.L. Classical dynamics / Handbuch der Physic, hrsd. V.S. Flugge, Bd 3/1. – Prinzipien der Klassischen Mechanik und Feldtheorie, B, 1960.