ABOUT CRITERIA FOR AN ESTIMATION OF NONLINEAR PARAMETERS IN MODELS OF MONITORING

Sergii Mostovyi, Vasilii Mostovyi

***Abstract**: In the problem of optimal estimation of model parameters using risk criteria we propose an approach of separation of linear parameters from the nonlinear ones. In the problem of finding a global minimum of risk criteria our approach leads to a decrease of the dimension of the space of free variables up to the dimension of the space of the nonlinear parameters. This allows one to obtain a simpler minimization problem, which can be solved more efficiently via Monte Carlo methods. Such an improvement is very significant in estimation of the models of the object “aging” while investigation of geophysical objects, models of which typically have high dimensionality. We illustrate the proposed method with processing and analysis of data obtained during the field observations in the regime of monitoring.*

Keywords: math model, active monitoring, risk criteria, linear and nonlinear parameters.

ACM Classification Keywords: G. 1. 6. Mathematics of Computing, Numerical Analysis, Optimization.

Introduction

Often monitoring systems are described by complicated models with the big size of dimension of the free parameters vector of these models. And only the part of these parameters are linear ones, and the biggest part of parameters are nonlinear ones.

In such case the essential role belongs to an aprioristic level of processes scrutiny. It allows to simplify an estimation process of free parameters. In such cases it is natural to build such procedure of an estimation of these parameters which is based on Bayesian estimations. It considers the aprioristic information which is available for the researcher.

As a result of monitoring there is a chance to analyze the changes in time of set of free parameters of model. It allows to draw conclusions on a status of object of research and to predict its behavior in the future. Such approach is fruitful in an estimation of a status of historically valuable architectural monuments and their possible reaction to earthquakes and creations concerning an optimum level of resistance to seismic events [Tassios, 2010].

Mathematical model of an estimation of parameters

Dynamic systems are modeled by superposition periodic or are quasi periodic processes which represent a subject of the analysis as enable to predict on background behavior of system in the future.

The essential role in creation of analysis algorithms is played with that fact, that in model of dynamic process there is even a part of free parameters which enter into model linearly, It allows to accelerate process of processing and the analysis of the data received as a result of monitoring. Model of process we build in the form of manifold [Виноградов, 1977]. It displays model in a point N - dimensional space $\mathbb{R}_N$ by means of free parameters of model.

Here N is a quantity of these parameters. $M(\mathbf{\Lambda},t)$ is a model of dynamic process which depends from time, but also from set of free parameters $\mathbf{\Lambda}$. The result of monitoring of process $y(t,\mathbf{\Lambda},\alpha)$ can be presented in the form of equality:

$$y(t,\mathbf{\Lambda},\alpha)=M(\mathbf{\Lambda},\alpha,t)+n(t)\ ,t\in\left[0,T\right] \tag{1}$$

In (1) $\mathbf{\Lambda}$ is a set of parameters which are reflected in model of process, are ordered in a rectangular matrix as the free parameters which are a subject an estimation, $\alpha(t)$ is a source of stochasticity- the fluctuating disturbance which are not allowing precisely to reproduce set $\mathbf{\Lambda}$. T is an area of supervision in time. A source of stochasticity except for an additive noise are also fluctuations of parameters of model.

We choose such model, at which set of its free parameters is contained by a subset of nonlinear parameters. We will put in order it to on and, putting in order it in a vector $\mathbf{h}=\left\{h_k\right\}, k=\overline{1,n}$ and selecting from the set of parameters $\mathbf{\Lambda}$ it subset. In $\mathbf{\Lambda}$ remain only nonlinear parameters, so come to the model of such kind:

$$M(\mathbf{\Lambda},t)=\sum_{k=1}^{n} h_k \exp\left\{-\lambda_{k,1} t\right\}\sin(\lambda_{k,2}t-\lambda_{k,3}). \tag{2}$$

$\mathbf{\Lambda}^{\langle k\rangle}=\left\{\lambda_{k,1},\ldots,\lambda_{k,3}\right\}^T, \quad k=\overline{1,n},$ -columns are in the matrix of nonlinear parameters $\mathbf{\Lambda}$.

The purpose of analysis is a construction of optimum estimations $\tilde{\tilde{\mathbf{\Lambda}}}$ by a set $\tilde{\tilde{\mathbf{\Lambda}}}=\left\{\tilde{\mathbf{\Lambda}}_r(y(t,\alpha(t))\right\}$ decision rules. At the choice of decision rule as a criterion of optimality on some set of R decision rules and observed data $y(t,\alpha(t)$ choose the average risk criterion $R(\tilde{\mathbf{\Lambda}}_r(y(t,\alpha))$ [3]

$$R(\tilde{\mathbf{\Lambda}}_r(y(t,\alpha))=\sum_{r=1}^{R} L(\tilde{\mathbf{\Lambda}}_r(y(t,\alpha),\mathbf{\Lambda}))P(\mathbf{\Lambda}/y(t,\alpha)). \tag{3}$$

Here $L(\tilde{\mathbf{\Lambda}}_r(y(t,\alpha),\mathbf{\Lambda}))$ is a loss function. Loss function depends on the estimation of free parameters which are accepted as a solution by decision rule $\tilde{\mathbf{\Lambda}}_r(y(t,\alpha)$ and truth value of these parameters $\mathbf{\Lambda}$. Here r is a rule number.

The type of function $L(\tilde{\mathbf{\Lambda}}_r(y(t,\alpha),\mathbf{\Lambda}))$ is not determined by a theory, it gets out from the engineering or intuitional considering. If accepted decision closer to the «truth», the value of loss function have to less, or even no more. $\tilde{\tilde{\mathbf{\Lambda}}}(y(t,\alpha)$ is the set of R decision rules. In formula (3) $P(\mathbf{\Lambda}/y(t,\alpha)$ is conditional probability of value of nonlinear free parameters on condition of realization $y(t,\alpha)$.

According to Bayes formula [Большаков, 1969] average risk, accurate within constant, appears in a kind:

$$R(\tilde{\mathbf{\Lambda}}_r(y(t,\alpha)) = C\sum_{\mathbf{\Lambda}\in\tilde{\mathbf{\Lambda}}} L(\tilde{\mathbf{\Lambda}}_r(y(t,\alpha),\mathbf{\Lambda}))P(y(t,\alpha)/\mathbf{\Lambda})P(\mathbf{\Lambda})) \tag{4}$$

Here $C=(P(y(t,\alpha)))^{-1}$ is the inverse to the total probability $P(y(t,\alpha)$ value. This value depends only on the realization $y(t,\alpha)$. The optimal estimation $\tilde{\mathbf{\Lambda}}^*$ of model parameters is obtained by minimization of average risk at the set of decision rules $\tilde{\mathbf{\Lambda}}_r(y(t,\alpha)$:

$$\tilde{\mathbf{\Lambda}}^* = \min_{\tilde{\mathbf{\Lambda}}_r(y(t,\alpha)\in\tilde{\tilde{\mathbf{\Lambda}}}} R(\tilde{\mathbf{\Lambda}}_r(y(t,\alpha)); \quad r=\overline{1,R} \tag{5}$$

Loss function of next kind was used in a calculations:

$$L(\tilde{\mathbf{\Lambda}}_r(y(t,\alpha),\mathbf{\Lambda})) = \begin{cases} 0, & \text{при} \quad \left\|\tilde{\mathbf{\Lambda}}_r - \mathbf{\Lambda}\right\| \le \Omega \\ \dfrac{1}{\Omega}, & \text{при} \quad \left\|\tilde{\mathbf{\Lambda}}_r - \mathbf{\Lambda}\right\| > \Omega \end{cases} \tag{6}$$

If the norm of difference of vector of estimations of model parameters does not excel the fixed value Ω, the observer of losses does not carry, otherwise the losses is Ω^{-1}.

Model of the research object

Such approach was applied by authors for data processing and analysis, got in the field of supervisions of monitoring of object with it spectral characteristics in seismic frequencies band (0,1 Hz – 8 Hz). The feature of experiment was that in monitoring duration the object characteristics were changed by the certain program, and the purpose of experiment was to estimate these changes in space of parameters of object model. The dimension of vector of free parameters changed depending on a selectable model from 12 to 20. Object was described a great number of linear and nonlinear model parameters, and the search of global a minimum of criterion of optimality of estimations for each of decision rules was carried out depending on a hypothesis about an additive background noise $\alpha(t_i)$. This background noise was designed, in general case, by non-stationary Gaussian noise. A decision rule was following: an estimation was accepted as minimum deviation of model from observed data in a norm of Hilbert spaces. Dot product for two functions $\langle y(t), f(t)\rangle_r$ in their discrete presentation got out for two vectors thus:

$$\langle y(t), f(t)\rangle_r = k(\boldsymbol{\sigma}_r)\sum_{i,j} y(t_i)f(t_j)W_r(t_i,t_j), \quad i,j=\overline{1,M},\ r=\overline{1,R} \tag{7}$$

Here - $W_r(t_i,t_j)$ is inverse matrix to the covariance matrix of additive background $n(t)$, $\boldsymbol{\sigma}_r$ - parameters of this matrix. $k(\boldsymbol{\sigma}_r)$ are coefficient, depending on the determinant of covariance matrix $\boldsymbol{\sigma}_r$. $M\Delta=T$ here

M is dimension of vector of discrete presentation of functions on a time interval with a length T, Δ - quantization interval, thus $M\Delta = T$.

The considered algorithm in calculations was described by two types of decision rules. One of them was based on supposition about a high-frequency background noise with a fixed number of values of power of such process. For example - white noise [Виноградов, 1977] in this case kernel in (7) will be:

$$k(\boldsymbol{\sigma}_r)W_r(t_i,t_j) = \frac{1}{\sigma_r}\delta(t_i - t_j) \tag{8}$$

Norma of deviation of one function from other (norm of difference) will be:

$$\sqrt{\langle y(t) - f(t), y(t) - f(t)\rangle_r} = \frac{1}{\sigma_r}\sqrt{\sum_{i=1}^{M}(y(t_i) - f(t_i))^2} \tag{9}$$

The second decision rule is based on a hypothesis about low-frequency stationary normal noise with fixed a number of determining parameters of this noise [Амиантов,1971]. Calculations of the data of monitoring of object, show that the hypothesis about a high-frequency background noise gave the best goodness of fit of model of object with observed data.

Model of monitoring object

We will consider hypothesis that an object is described by linear combination from n oscillators, each of which is submodel (case (2)):

$$M_k(\boldsymbol{\Lambda}^{\langle k\rangle},t) = h_k \exp\left\{-\lambda_{k,1} t\right\}\sin\left(\lambda_{k,2}t - \lambda_{k,3}\right). \tag{10}$$

In this case the model (2) appears in a kind:

$$M(\boldsymbol{\Lambda},t) = \sum_{k=1}^{n} h_k \exp\left\{-\lambda_{k,1} t\right\}\sin\left(\lambda_{k,2} t - \lambda_{k,3}\right) \tag{11}$$

Using the metrics of decision rule with the number of *n*, we get the system of linear equalizations regard to a vector $\mathbf{h}$ and observed date $y(t)$, at the fixed values of matrix $\boldsymbol{\Lambda}$:

$$\boldsymbol{\Psi}_r \mathbf{h}_r = \mathbf{I}_r;\ r = \overline{1,R} \tag{12}$$

Here a matrix with the index of *r* has elements:

$$\Psi^r_{k,s} = \frac{1}{\sigma_r}\sum_{i=0}^{M}\sum_{k=1}^{n} M_k(\boldsymbol{\Lambda}^{\langle k\rangle},t_i)M_s(\boldsymbol{\Lambda}^{\langle s\rangle},t_i),\ s = \overline{1,n}. \tag{13}$$

The vector of right side of equation $\mathbf{I}_r$ was calculated by the following formula:

$$\mathbf{I}_r = \{I^r_s\};\ I^r_s = \frac{1}{\sigma_r}\sum_{i=0}^{M} y(t_i)M_s(\boldsymbol{\Lambda}^{\langle s\rangle},t_i). \tag{14}$$

It is necessary to mark that Bayesian approach to the decision-making is used widely in geophysics. It is enough to mark recent published work [Imoto, 2010].

Processing of observed data

Processing of observed data was carried out by a next algorithm: by the method of Monte Carlo, on a prior distribution, a pseudorandom point was thrown out in space of nonlinear parameters. In the model of object (formula (10)) for these point and observed data the root of the system of linear equations (D12) for linear parameters was calculate a point in subspace of linear parameters. The set of nonlinear and linear free parameters of model, got thus used for the estimation of local minimum in a nonlinear problem as an initial point for an algorithm of Livenberga-Makvardta [Levenberg, 1944; Marquardt, 1963]. This procedure was executed recursively, the estimation of a minimum of minimoruma was as a result calculated on the set of pseudorandom points. Such estimation approach stochastically to the global minimum. In 12 dimensional space of free parameters of model an object was reflect in to a point with such components:

$\lambda^T =$

0	1	2	3	4	5	6	7	8	9	10
$7.313 \cdot 10^{-3}$	2.962	0.698	0.94	0.114	9.028	0.452	-0.203	-0.231	11.284	$6.07 \cdot 10^{-3-}$

Here 0, 4 and 8 components are an index of decrement of exponent to the first, the second, and the third harmonics. First, fifth and ninth components is circular frequency according to the first the second and the third harmonic (given in radians). Second, sixth and tenth components is amplitudes of harmonics) given in relative units). Third, seventh and eleventh harmonic is a phase change in radians. Third, seventh and eleventh components is a phase change (given in radians).

Conclusions

Proposed an approach of division of linear and nonlinear parameters of model, at their optimum evaluation by the criteria of risk. It allows to decrease the dimension of space of the estimated parameters in the Monte Carlo method at the search of global a minimum of risk criteria to the dimension of space only of nonlinear parameters. It is specially important at the estimation of models with the large dimension of free parameters, that typically for the tasks of geophysics. Data processing of object monitoring with changing characteristics by the offered algorithm allowed to estimate even insignificant changes in a state of object in space of free parameters of model, i.e. position of point in space of it characteristics.

Bibliography

[Tassios, 2010] Tassios T. P. Seismic engineering of monuments, Bulletin of earthquake engineering.//Volume 8, Number 6 / December 2010 - P.1231-1265.

[Виноградов, 1977] Математическая энциклопедия (н.р. Виноградов И.М.) -М: Сов. энциклопедия, т.3, 1977.- 742 с.

[Большаков, 1969] Большаков И.А. Статистические проблемы выделения потока сигналов из шума.-М: Совесткое радио, 1969 - 453 с.

[Виноградов, 1977] Математическая энциклопедия (н.р. Виноградов И.М.).-М:Сов. энциклопедия, т. 1, 1977.- 402 с.

[Амиантов, 1971] Амиантов И.Н. Избранные вопросы статистической теории связи.-М: Советское радио, 1971.- 423 с.

[Imoto, 2010] Imoto M. Hypocenter determination with prior information for clustering earthquakes// Geophysical Journal International, 2010, Volume 182, Issue 3 - P. 1374-1382.

[Levenberg, 1944] Levenberg K. A. Method for the Solution of Certain Non-Linear Problems in Least Squares// The Quarterly of Applied Mathematics, N2, 1944 -P.164–168.

[Marquardt, 1963] Marquardt D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. // SIAM Journal on Applied Mathematics, N11, 1963.-P. 431–441.

Authors information

Mostovyi Vasilii *– Institute of Geophysics of the National Academy of Sciences, Kiev, Ukraine. E-mail: vasmost@mail.ru ;*

Mostovyi Sergii *– National Aviation University, Kiev, Ukraine. E-mail: smost@volia.ua*