Research of properties of conditionality of task to optimization of processes of concentrating technology is on the basis of application of neural networks

Andrey Kupin

Dean of IT faculty, Doctor of Science, Professor of Computer Science, Computer Systems and Networks department, Krivyi Rih National University

Abstract
The paper describes research of properties of conditionality of task to optimization of regime parameters of technological processes of concentrating of iron-stone is on the basis of forming of goal function with application of multi-layered neural networks.
Key words: CONDITIONALITY OF TASK TO OPTIMIZATION, GOAL FUNCTION, NEURAL NETWORK, CONCENTRATING OF IRON ORE.

Complexity indexes of optimization of multidimensional objective functionals in order to fulfil intellectual control in the conditions of technological processes (TP) of iron ore concentration are considered [1, 2]. It is necessary to define whether the task is badly conditioned. The result of such analysis determines efficiency of application of selected method of global optimization. The following factors influence the conditionality of multivariable task: presence of local extremums in search areas, their amount and dimension, bulge and smoothness of objective function, etc. [3]

Let us estimate the conditionality of task of global optimization of multivariable goal function for the conditions of TP section of concentrating complex. Granting the certain stage of concentrating of magnetite quartzites is approximated by a neural network (NN) with one hidden layer (fig. 1).

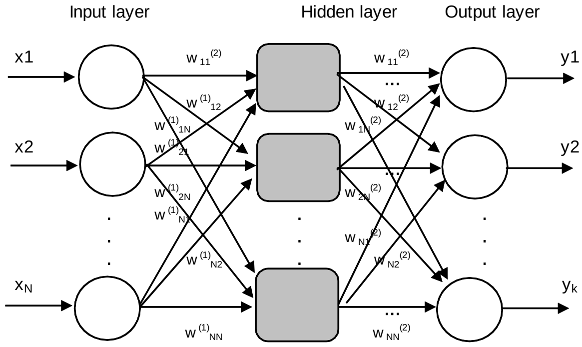

Figure 1. An example of multi-layered neutral network of direct distribution for approximation of TP of one stage of concentrating

If the function of activating as sigmoid for all types of layers is used, then mathematical expression of neural network appears as:

$$y_i(\theta) = F_i\left(\sum_{j=1}^{n_h} W_{ij} f_j\left[\sum_{l=1}^{n} x_l w_{ij} + w_{j0}\right] + W_{i0}\right) =$$

$$= \left(1 + Exp\left(-a\left(\sum_{j=1}^{n_h} W_{ij}\left(1 + Exp\left(-a\left[\sum_{l=1}^{n} x_l w_{ij} + w_{j0}\right]\right)\right)^{-1} + W_{i0}\right)\right)\right)^{-1},$$

(1)

where $F_i(s) = f_j(s) = \varphi(s) = 1/(1 + Exp(-as))$ is common expression for the activation function of sigmoid type for the output hidden and input glowed networks; y_i – a value of output signal of network; n_h – amount of neurons of the hidden layer; n – dimension of input vector; θ – vector of parameters of neural network, which require tuning (contains weight coefficients and neuron changes w_{jl}, W_{ij}); $f_j(x)$ – the value of activation function of neurons for an input and hidden glowed; $F_i(x)$ – value of function of activating of neurons of initial layer.

Then on the basis of expression of technological criterion from [1] for one certain stage one can form objective functional for further optimization as follows:

$$J(Q_i, \beta_i, \beta_{XBi}) = \begin{cases} Q_i \to \max \\ \beta_i^{\min} \le \beta_i \le \beta_i^{\max} \\ \beta_{Xi}^{\min} \le \beta_{Xi} \le \beta_{Xi}^{\max} \end{cases} \Rightarrow J(y_1, y_2, y_3) = \begin{cases} y_1 \to \max \\ b_1 \le y_2 \le b_2 \\ c_1 \le y_3 \le c_2 \end{cases}, \qquad (5.2)$$

where i is the number of chosen stage of concentrating; Q_i, y_1 - productivity of the corresponding stage; $\beta_i, \beta_{Xi}, y_2, y_3$ - quality of output product and tails of the selected stage; $\beta_i^{\min}, \beta_i^{\max}, \beta_{Xi}^{\min}, \beta_{Xi}^{\max}, b_1, b_2, c_1, c_2$ - minimum and maximum limits on the parameters of quality of industrial product and loss in tails after the stages.

Taking into account the expression (1) a goal function may be presented as optimization task with four conditions:

$$J(y_i) = \begin{cases} y_i = \left(1 + Exp\left(-a\left(\sum_{j=1}^{n_h} W_{ij}\left(1 + Exp\left(-a\left[\sum_{l=1}^{n} x_l w_{ij} + w_{j0}\right]\right)\right)^{-1} + W_{i0}\right)\right)\right)^{-1} \\ y_1 \to \max \\ y_2 - b_1 \ge 0,\ y_2 - b_2 \le 0 \\ y_3 - c_1 \ge 0,\ y_3 - c_2 \le 0. \end{cases} \qquad (3)$$

For the removal of the marked limitations and decision of such optimization task it is necessary to apply the method of penalty functions or method of the modified Langrangians [3-5]. At the same time the use of the marked approaches lead to worsening of properties of conditionality. Therefore for the estimation of conditionality expression (3) on the first stage can be simplified (to analyse case-insensitive limitations)

$$J(y_1) = \left(1 + Exp\left(-a\left(\sum_{j=1}^{n_h} W_{ij}\left(1 + Exp\left(-a\left[\sum_{l=1}^{n} x_l w_{ij} + w_{j0}\right]\right)\right)^{-1} + W_{i0}\right)\right)\right)^{-1} =$$

$$= \left(1 + Exp\left(-a\left(\sum_{j=1}^{n_h} W_{1j}\left(1 + Exp\left(-a\left[\sum_{l=1}^{n} x_l w_{1j} + w_{j0}\right]\right)\right)^{-1} + W_{10}\right)\right)\right)^{-1} \to \max. \qquad (4)$$

In case if task of search of global extremum as (4) is badly conditioned, then, accordingly, concerning more general and difficult (due to the necessity of taking into account the limitations) task (3) it is necessary to do analogical conclusions. Otherwise it is necessary additionally to check the conditionality of complete optimization task (3).

According to [3] direct method of calculation of conditionality of optimization task is based on the calculation of matrix of Hesse for the goal functional (of flexon $J''(x)$) and for finding the complete great number of own values. Herein, if $J''(x) > 0$, then the following expression is true:

$$\eta(x) = cond[J''(x)] = \frac{\max \lambda_i(x)}{\min \lambda_i(x)}, \quad i = \overline{1, N_\lambda}, \quad (5)$$

where $\eta(x)$ - local degree of multi-modality (the presence of many extremums) in a point (x); $cond[J''(x)]$ is a spectral number of conditionality of matrix of Hesse $J''(x)$; $\lambda_i(x)$ is a value of own numbers of matrix is the point (x); N_λ - complete amount of own numbers.

The criterion of conditionality of optimization task is the expression

$$\log_2 \eta > \varepsilon_m, \quad (6)$$

where ε_m is the length of bit net of calculable machine (for the most modern programming language of high level the maximal exactness of calculations may be realized by two types of Double/Extended: range of values of bit net within the limits of 5,0e-324...1,7e+308/3,4e-4932.1,1e+4932 respectively; amount of numbers of mantissa is 15/19 signs; a volume of necessary memory is 8/10 byte [6, 7]).

Fulfilment of condition (6) allows to classify optimization task as badly conditioned. Vector of gradient for goal function (4) equals:

$$J'(x) = \left[\frac{\partial J}{\partial x_1}, ..., \frac{\partial J}{\partial x_n}\right] = \left[J_0 J_1, ..., J_0 J_n\right] = \left[J_0 J_i\right], \quad (7)$$

where $J_0 = -\left(1 + Exp\left(-a\left(\sum_{j=1}^{n_h} W_{1j}\left(1 + Exp\left(-a\left[\sum_{l=1}^{n} x_l w_{1j} + w_{j0}\right]\right)\right)^{-1} + W_{10}\right)\right)\right)^{-2}$;

$$J_i = \frac{\partial}{\partial x_i}\left[1 + Exp\left(-a\left(\sum_{j=1}^{n_h} W_{1j}\left(1 + Exp\left(-a\left[\sum_{l=1}^{n} x_l w_{1j} + w_{j0}\right]\right)\right)^{-1} + W_{10}\right)\right)\right]' =$$

$$= Exp\left(-a\left(\sum_{j=1}^{n_h} W_{1j}\left(1 + Exp\left(-a\left[\sum_{l=1}^{n} x_l w_{1j} + w_{j0}\right]\right)\right)^{-1} + W_{10}\right)\right) \times$$

$$\times\left(-a^2 \sum_{j=1}^{n_h} W_{1j}\left(1 + Exp\left(-a\left[\sum_{l=1}^{n} x_l w_{1j} + w_{j0}\right]\right)\right)^{-2} \cdot x_i w_{1j}\right) \times Exp\left(-a\left[\sum_{l=1}^{n} x_l w_{1j} + w_{j0}\right]\right)$$

Hesse matrix for goal function of (4) type is determined on the basis of vector of gradient

$$J''(x) = \left[\frac{\partial J}{\partial x_1}, ..., \frac{\partial J}{\partial x_n}\right]' = \begin{bmatrix} \frac{\partial^2 J}{\partial^2 x_1}, \frac{\partial^2 J}{\partial x_1 \partial x_2}, ..., \frac{\partial^2 J}{\partial x_1 \partial x_n} \\ \frac{\partial^2 J}{\partial x_2 \partial x_1}, \frac{\partial^2 J}{\partial^2 x_2}, ..., \frac{\partial^2 J}{\partial x_2 \partial x_n} \\ ... \\ \frac{\partial^2 J}{\partial x_n \partial x_1}, \frac{\partial^2 J}{\partial x_n \partial x_2}, ..., \frac{\partial^2 J}{\partial x_n^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial[J_0 J_1]}{\partial x_1}, \frac{\partial[J_0 J_1]}{\partial x_2}, ..., \frac{\partial[J_0 J_1]}{\partial x_n} \\ \frac{\partial[J_0 J_2]}{\partial x_1}, \frac{\partial[J_0 J_2]}{\partial x_2}, ..., \frac{\partial[J_0 J_2]}{\partial x_n} \\ ... \\ \frac{\partial[J_0 J_n]}{\partial x_1}, \frac{\partial[J_0 J_n]}{\partial x_2}, ..., \frac{\partial[J_0 J_n]}{\partial x_n} \end{bmatrix} = \begin{bmatrix} J_{11}, J_{12}, ..., J_{1n} \\ J_{21}, J_{22}, ..., J_{2n} \\ ... \\ J_{n1}, J_{n2}, ..., J_{nn} \end{bmatrix} \quad (8)$$

where $J_{ij} = \dfrac{\partial[J_0 J_i]}{\partial x_j}, \quad i = \overline{1,n};\ j = \overline{1,n}.$

By convention the own numbers of matrix of (8) type are determined as roots of equation:

$$\left| J''(x) - \lambda E \right| = 0,$$

where E is a single diagonal matrix of dimension $n \times n$.

With taking (8) into account we will obtain:

$$\left| \begin{bmatrix} J_{11}, J_{12},..., J_{1n} \\ J_{21}, J_{22},..., J_{2n} \\ ... \\ J_{n1}, J_{n2},..., J_{nn} \end{bmatrix} - \lambda \begin{bmatrix} 1,0,...,0 \\ 0,1,...,0 \\ ... \\ 0,0,...,1 \end{bmatrix} \right| = \begin{vmatrix} (J_{11} - \lambda), J_{12},..., J_{1n} \\ J_{21}, (J_{22} - \lambda),..., J_{2n} \\ ... \\ J_{n1}, J_{n2},..., (J_{nn} - \lambda) \end{vmatrix} = 0$$

(9)

After opening of determinant (9), we will get characteristic equalization of the following type:

$$a_1\lambda^n + a_2\lambda^{n-1} + ... + a_n\lambda + a_{n+1} = 0, \qquad (10)$$

where a_i, $i = \overline{1, n+1}$ are certain coefficients, which are calculated on the basis of determinant (10).

Hereafter the numeral solution of characteristic equation (10) with selected exactness ε_λ is made. Herein the roots $\lambda > 0$ are selected. The maximum and minimum values of such roots are placed into expression of criterion (6). After calculation of the value of criterion there made a conclusion about the conditionality of task of global optimization.

Methodology of determination of degree of conditionality of optimization task of goal function of (4) type can be represented as an algorithm (fig. 2).

Methodology of determination of degree of conditionality of optimization task of goal function of (4) type can be represented as an algorithm (fig. 2).

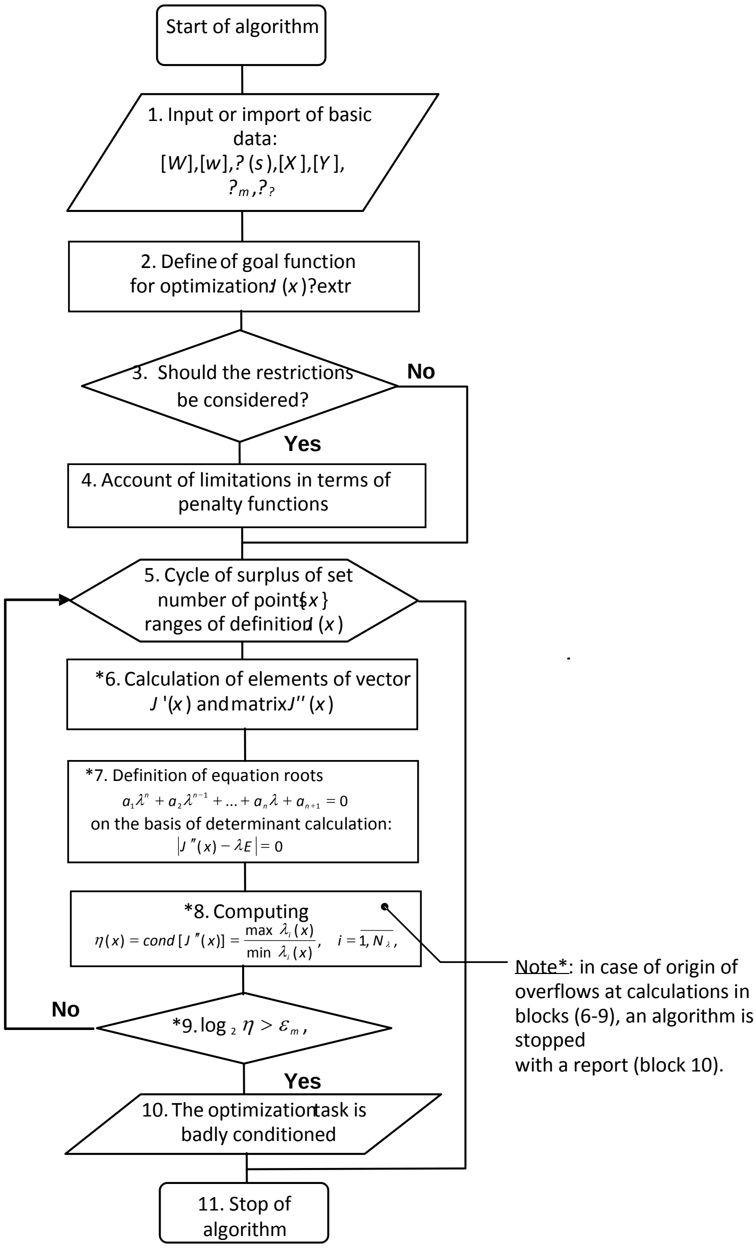

Figure 2. The general algorithm of determination of conditionality of goal function on the basis of NN

Calculations executed according to the algorithm (fig. 2) with parameterization only of the first stage of TP concentration allow to draw conclusion about bad conditionality of task of global optimization even for a goal function of (4) type. Separate results over and estimation of such calculations are shown in the figure 3.

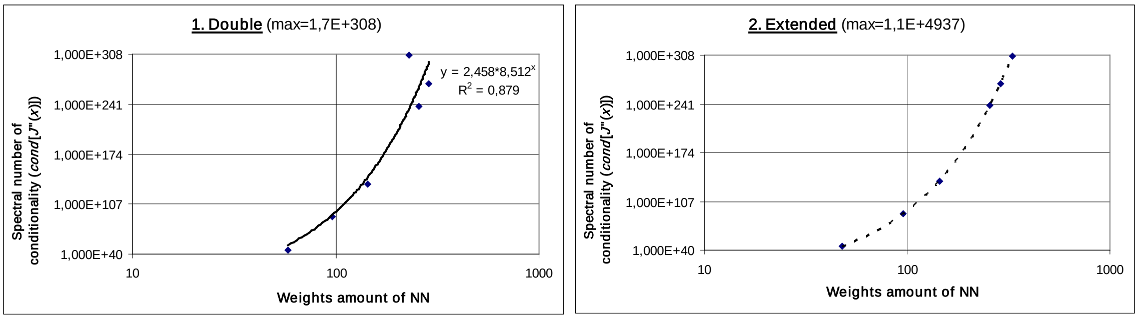

Figure 3. Indexes of conditionality of optimization task of NN goal functions for TP concentrating at different maximum parameters of computer bit grid of (1 is a trend on the base of application of data types Double; 2 is extrapolation with the use of Extended type).

Analysis of fig 3 shows that at application of 8-byte of Double data type with mantissa of 15 signs the maximum value of spectral number of conditionality (1.7E+308) is achieved at already at the amount of weights of neural network of less then 1000. At the further increase of weights of NN from 1000 to 10000 a maximum value is achieved for the type of Extern (1.1e+4937).

Conclusions

Undertaken studies allow to confirm that task of global optimization of parameters of TP of concentrating for one or a few stages on the base of NN are goal (with the use of limitations and without them) functions it badly conditioned. The above mentioned shows the limit possibilities of application of traditional methods of multivariable optimization (in particular, gradient) [3]. For the successful decision of the task it is necessary to use the methods, which are proof to the conditions of multi-modality of goal functions.

References

1. Kupin, A. I. Intellectual identification and controls in the conditions of processes of concentrating technology. The monography. *Kyiv, Korneychuk's Publishing house*, 2008.
2. Kupin, A. I. Neural identification of technological process of iron ore beneficiation. – Proceedings of 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems Technology and Applications (IDAACS’2007). Dortmund, Germany, 2007. P.225–227.
3. Chernodutsky I.G. Optimization methods into control theory. *SPb*, *Piter*, 2004. 256 p.
4. Karr C.L. (200). Optimization of hydrocyclone operation using a geno-fuzzy algorithm. *Computer Methods in Applied Mechanics and Engineering*, No 186. P.517–530.
5. Morkun, V., Goncharov, S., Pikilnyak, A., Krivenko, A. (2012) Iron ore benefication processes optimization. *TEKA. Commision of Motorization and Energetics in Agriculture 12*, No.4, P.162-166.
6. Jones M.T. AI Application Programming. Charles River Media, 2003.
7. Popovic D., Bhatkar V.P. Distributed Computer Control Systems in Industrial Automation. CRC Press, 1990.