MSC 74A45

Algorithm of neural network method for creep model parameter identification problem

E.B. Kuznetsov [1], S.S. Leonov [1], A.N. Vasilyev [2]

[1] Moscow Aviation Institute (National Research University),
[2] Peter the Great St. Petersburg Polytechnical University

Abstract: The paper deals with a parameter identification problem for creep and fracture model that describe a process of metal structures deformation. The system of ordinary differential equations of Rabotnov's structural parameters kinetic creep theory is applied for describing this model. For solving the parameter identification problem, we proposed to use principals and techniques of neural network modeling. The procedure of neural network modeling application we are going to use for finding of uniaxial tension model parameters for isotropic steel 45 specimens at creep conditions. The obtained results of neural network modeling will be compared with theoretical strain-damage characteristics, experimental data and results of other authors.

Keywords: creep, fracture, damage parameter, artificial neural network, parameter identification problem, ordinary differential equation, initial value problem

1. Introduction

In recent years, there is an increasing need for the description of deformation and fracture processes in complex temperature-power regimes for materials with complicated rheological characteristics including viscosity. These problems find wide application in different fields of science and technology, e. g. mechanical engineering and aerospace industry. Special attention is paid to the possibility of creep accounting at high and moderate temperatures for metal and composite structures. However, up to now there has been no common approach to the description of this phenomenon, and there are dozens various creep theories and their modifications, e. g. the aging theory, the hardening theory, the heredity theory and the Rabotnov theory of structural parameters. It is not usually possible to reliably determine which of the theories is better to use in a particular case. Applying equations of any theory may be a very complex process, as equations used usually contain several material constants (creep characteristics), which complicated to obtain. In other words, these parameters can be determined by using information about a deformation process. A primary source of information is an experiment. Creep characteristics may depend on the type of used material and its condition, regime of loading, temperature, type of anisotropy, and other factors. Problems of parameter identification are very complicated. All these reasons indicate the need for a common approach to determining parameters of models for various equation types. This paper provides a unified method for identifying parameters of models describing creep and fracture processes of structures. The method considered involves the use of experimental data. The principles and the techniques of neural network modeling are adopted as a basis for the approach developed in the work.

2. Neural network methodology

By using the results of the monograph [1], we consider the construction of a neural network for the system of m ordinary differential equations (ODEs) of order r with p unknown scalar parameters given by a vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p)^T$

$$\mathbf{F}(t, \mathbf{y}, \mathbf{y}', \ldots, \mathbf{y}^{(r)}, \alpha) = 0,\ t \in [t_0, t_*] \tag{1}$$

with initial conditions

$$\begin{cases} \mathbf{y}(t_0) = \mathbf{y}_0, \\ \vdots \\ \mathbf{y}^{(r-1)}(t_0) = \mathbf{y}_{r-1}. \end{cases} \tag{2}$$

The following notation is introduced here

- $\mathbf{F}\left(t, \mathbf{y}, \mathbf{y}', \ldots, \mathbf{y}^{(r)}, \alpha\right) = \left(f_1(t, \mathbf{y}, \mathbf{y}', \ldots,\ \mathbf{y}^{(r)}, \alpha), \ldots, f_m(t, \mathbf{y}, \mathbf{y}', \ldots, \mathbf{y}^{(r)}, \alpha)\right)^T$ is a vector function of vector arguments;
- $f_i(t, \mathbf{y}, \mathbf{y}', \ldots \mathbf{y}^{(r)}, \alpha)$, $i = 1, \ldots, m$ are scalar functions of vector arguments;
- $\mathbf{y}(t) = (y_1(t), y_2(t), \ldots, y_m(t))^T$ is the required solution, which is a vector function of argument t and also implicitly depends on the parameters $\alpha_1, \alpha_2, \ldots, \alpha_p$;
- $\mathbf{y}_j = (y_{j1}, y_{j2}, \ldots, y_{jm})^T$, $j = 0, \ldots, r-1$ are vectors of $\mathbf{y}(t)$ and its $r-1$ derivatives values at the point t_0;
- $\mathbf{y}^{(j)}(t) = \dfrac{d^j \mathbf{y}(t)}{dt^j}$;
- $\mathbf{y}^{(0)}(t) = \mathbf{y}(t)$.

While solving physical problems certain restrictions must be imposed on ranges of values of the parameters $\alpha_1, \alpha_2, \ldots, \alpha_p$

$$\alpha_s \in A_s \subseteq \mathbb{R}, \quad s = 1, \ldots, p. \tag{3}$$

Suppose that the problem (1)-(2) also satisfies the Cauchy existence theorem. Moreover, there is a set of additional data on behavior of function $\mathbf{y}(t)$ at points $t_1,\ t_2,\ \cdots,\ t_l$

$$\mathbf{y}(t_q) = \mathbf{y}_q^e, \quad t_q \in (t_0, t_*], \quad q = 1, \ldots, l, \tag{4}$$

where $\mathbf{y}_q^e = \left(y_{q1}^e, y_{q2}^e, \ldots, y_{qm}^e\right)^T$ are vectors of $\mathbf{y}(t)$ values at points $t_1, t_2, \ldots, t_l$.

We use the approach to the solution of the initial value problem (1)-(2) based on neural network modeling technique. We are going to approximate each component of the vector function $\mathbf{y}(t)$ via neural networks, i. e. we represent them in the form

$$\hat{y}_i(t,\ \mathbf{w}_i) = \sum_{j=1}^{N_i} c_{ij} \nu_j(t, \mathbf{a}_{ij}), \quad i = 1, \ldots, m. \tag{5}$$

Here:

- $\mathbf{w}_i = (\mathbf{w}_{i1}, \mathbf{w}_{i2}, \ldots, \mathbf{w}_{iN_i})$ are matrices of parameters to be determined (neural network coefficients);
- $\mathbf{w}_{ij} = (c_{ij}, \mathbf{a}_{ij})$;
- c_{ij} are linear input parameters;
- $\mathbf{a}_{ij} = \left(a_{ij}^1, a_{ij}^2\right)$ are nonlinear input parameters;
- N_i are numbers of neuron units in (5).

The type of a neural network basis element $c_{ij}\nu_j(t,\mathbf{a}_{ij})$ is determined by a scalar function of the scalar argument, so-called activation function

$$\nu_j(t,\mathbf{a}_{ij}) = \varphi_j(x),\ x = \psi(t,\mathbf{a}_{ij}),$$

where $\psi(\cdot)$ is some given function (for example, $\psi(t,\mathbf{a}_{ij}) = a_{ij}^1 \cdot t + a_{ij}^2$). The activation function may be given in the form of a hyperbolic tangent, radial basis function (e. g. in form of Gaussian function $\varphi(x) = \exp\{-x^2\}$) or in another form as it is presented in the book [1].

Using relations (1)-(2) as well as the neural network approximation (5) and additional data (4), we obtain a normalized error functional in the integral (continuous) form

$$\begin{gathered} J(\alpha,\mathbf{w}_1,\dots\mathbf{w}_m) = R\cdot\sum_{i=1}^{m}\left(\beta_i\int\limits_{t_0}^{t_*}\left|f_i(\xi,\hat{\mathbf{y}},\hat{\mathbf{y}}',\dots\hat{\mathbf{y}}^{(n)},\alpha)\right|^2 d\xi + \right. \\ \left. +\gamma_i\sum_{j=0}^{r-1}\left|\hat{y}_i^{(j)}(t_0,\mathbf{w}_i) - y_{ji}\right|^2 + \delta_i\sum_{q=1}^{l}\left|\hat{y}_i(t_q,\mathbf{w}_i) - y_{qi}^e\right|^2\right), \end{gathered} \tag{6}$$

where

- $\hat{\mathbf{y}}(t,\mathbf{w}) = (\hat{y}_1(t,\mathbf{w}_1),\dots,\hat{y}_m(t,\mathbf{w}_m))^T$ is a vector of the neural network approximation (5);
- $R = \left(\sum\limits_{i=1}^{m}(\beta_i\cdot(t_*-t_0)+\gamma_i\cdot r+\delta_i\cdot l)\right)^{-1}$;
- $\beta_i,\ \gamma_i,\ \delta_i,\ i=1,\dots,m$ are penalty coefficients.

Also we can get the error functional (6) in the discrete form

$$\begin{gathered} J(\alpha,\mathbf{w}_1,\dots\mathbf{w}_m) = \frac{1}{\sum\limits_{i=1}^{m}(\beta_i\cdot M+\gamma_i\cdot r+\delta_i\cdot l)}\times \\ \times\sum_{i=1}^{m}\left(\beta_i\sum_{j=1}^{M}\left|f_i(\xi_j,\hat{\mathbf{y}},\hat{\mathbf{y}}',\dots\hat{\mathbf{y}}^{(n)},\alpha)\right|^2 + \right. \\ \left. +\gamma_i\sum_{j=0}^{r-1}\left|\hat{y}_i^{(j)}(t_0,\mathbf{w}_i) - y_{ji}\right|^2 + \delta_i\sum_{q=1}^{l}\left|\hat{y}_i(t_q,\mathbf{w}_i) - y_{qi}^e\right|^2\right). \end{gathered} \tag{7}$$

The components of the vector function $\mathbf{F}(t,\mathbf{y},\mathbf{y}',\dots,\mathbf{y}^{(r)},\alpha)$ are calculated at test point (point of reference) ensemble $\{\xi_j\}_{j=1}^{M}$; the random components of the ensemble are supposed to be uniformly distributed on the segment $[t_0,t_*]$.

We obtain parameters and neural network coefficients while solving the error functional (7) (or (6)) minimization problem within the constraints (3)

$$J(\alpha,\mathbf{w}) \xrightarrow[\mathbf{w}_1,\mathbf{w}_2,\dots,\mathbf{w}_m]{\alpha_1,\alpha_2,\dots,\alpha_p} \min. \tag{8}$$

So we get the parameters $\alpha_1^*,\alpha_2^*,\dots,\alpha_p^*$ and the matrices of the neural network coefficients $\mathbf{w}_1^*,\mathbf{w}_2^*,\dots,\mathbf{w}_m^*$ that minimize the error functional (7) (or (6)). Also we can get the neural network approximation to the solution in the form

$$\hat{\mathbf{y}}(t) = \hat{\mathbf{y}}(t,\mathbf{w}_1^*,\dots,\mathbf{w}_m^*),\ t\in[t_0,t_*]. \tag{9}$$

Note that the process of the error functional (7) (or (6)) minimization is being conducted not up to a global minimum, but up to the moment when the functional value is less than the value of a permissible error η, i. e. $J < \eta$. And this functional value J_* is taken as an approximation to the error functional global minimum. In order to avoid interruption of the minimization process at a local minimum we periodically make some regeneration of $\{\xi_j\}_{j=1}^{M}$ after a few iterations of the minimization algorithm.

This approach we are going to used for solution of the creep model parameter identification problem for uniaxial tension of isotropic aviation steel 45 cylindrical specimens at constant temperature $T = 850\,^{\circ}$C.

3. Conclusion

The process of the Cauchy problem solution constructing for ODEs system of order r with unknown parameter set is described with the use of the neural network method. This approach will be used to solve the creep model parameter identification problem for uniaxial tension of steel 45 specimens at the constant stress and constant temperature.

In further research we are going to develop the application of neural network technique to the solution of parameter identification problems for creep and viscoelasticity models using method of solution continuation with respect to the best parameter [6]. Special attention will be paid to problems of determining strain-strength characteristics of structural members at a complex stress state. Problems of this type often encountered in engineering practice. At the moment, there is a lack of experimental data for them. This fact can complicate a process of constructing an adequate model. So, before going on to these problems, it is also important to investigate the use of other forms of neural network basis elements and error functionals for the problem considered in the paper.

4. Acknowlegments

The work was supported by the Russian Foundation for Basic Research, project number 16-08-00943.

References

1. Vasilyev A. N., Tarkhov D. A. Neirosetevoe modelirovanie. Printsipy. Algoritmy. Prilozheniya [Neural Network Modeling. Principles. Algorithms. Applications.] Saint-Petersburg. SPbSPU Publishing House, 2009. 528 p.

2. Rabotnov Yu. N. Creep Problems in Structural Members. Amsterdam/London. North-Holland Publishing Company, 1969. 822 p.

3. Gorev B. V., Zakharova T. E., Klopotov I. D., On description of creep and fracture of the materials with non-monotonic variation of deformation-strength properties // Physical Mesomechanics Journal. 2002. V. 5, No. 2. P. 17-22.

4. Gorev B. V., Klopotov I. D., Lyubashevskaya I. V., Creep and damage behavior of AK4-1T and VT-9 alloys under different stress states // Theor. and Appl. Fract. Mech. 1998. No. 29. P. 1-10.

5. Polak E., Computational Methods in Optimization: A Unified Approach. New York. Academic Press, 1971. 329 p.

6. Shalashilin V. I., Kuznetsov E. B. Parametric Continuation and Optimal Parametrization in Applied Mathematics and Mechanics. Dordrecht / Boston / London. Kluwer Academic Publishers, 2003. 236 p.