Asian Research Journal of Mathematics

7(3): 1-9, 2017; Article no.ARJOM.37856
ISSN: 2456-477X

Nonlinear Inverse Problems for Von Karman Equations: A Neural Network Approximation

Natalia I. Obodan[1], Oleksii S. Mahas[1*] and Vasilii A. Gromov[1]

[1]Oles Honchar Dnipro National University, Dnipro, Ukraine.

Authors' contributions

This work was carried out in collaboration between all authors. Author NIO designed the study. Authors OSM and VAG performed the method developing, wrote the protocol and wrote the first draft of the manuscript. Author OSM implemented the method. All authors managed the analyses of the study. Authors NIO and OSM managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/ARJOM/2017/37856
Editor(s):
(1) Dr. Junjie Chen, University of Texas at Arlington, USA.
Reviewers:
(1) Santhosh George, National Institute of Technology, India.
(2) Fayyaz Ahmad, Polytechnic University of Catalonia, Spain.
(3) Ioannis K. Argyros, Cameron University, USA.
(4) Teodoro Lara, University of Los Andes, Venezuela.
Complete Peer review History: http://www.sciencedomain.org/review-history/22077

Original Research Article

Received: 30th October 2017
Accepted: 23rd November 2017
Published: 28th November 2017

Abstract

This paper considers the coefficient inverse problem for the nonlinear boundary problem of von Karman equations. The Fréchet differentiability of the inverse operator is proved and its neural network approximation is constructed with neuroevolution augmented topology model. The model used proves efficient to solve the coefficient inverse problem even for the parameters values close to those corresponding to singular solutions of the direct problem.

Keywords: The coefficient inverse problem; nonlinear boundary problem; von Karman equations; the inverse operator; the Fréchet differentiability; neuro-evolution augmented topologies.

1 Introduction

The coefficient inverse problem for nonlinear boundary problem of PDEs (particularly, of von Karman equations) is conventionally solved with the employment of various optimization methods like the Newton,

**Corresponding author: E-mail: olexiy.magas@gmail.com;*

the Gauss-Newton, the gradient descent and others [1]. This approach implies that some regularization technique is applied to iterative process: Engl and Kügler [2] present a brilliant review of various inverse problems and regularization techniques with a particular emphasis on nonlinear ones.

Alternatively, one can employ a neural network approximation for the inverse operator that maps traces of the direct (forward) problem solution onto unknown functions of the inverse one. To this end, these solutions are discretized and a learning sample for a neural network is selected in such a way that the values of the inverse problem function corresponding to its vectors form a compact set. Such approach guarantee that the inverse problem is regularized provided the direct problem does not possess singular points on its definition domain and that the traces of the direct problem solution corresponding to the inverse problem solution are close enough to the specified ones. Both formulations require that the Fréchet differentiability of the inverse operator be proved [3]. The paper [4] considers the inverse problem for non-linear time-dependent Schrödinger equation; the problem is formulated using variational approach. The author investigates the existence, uniqueness, and Fréchet differentiability of its solution. The paper [5] deals with inverse problems in elasticity with application to cancer identification.

The coefficient inverse problem for von Karman equations forms the subject of the present paper. The respective direct problem – the nonlinear boundary problem of von Karman type – belongs to nonlinear elliptic problems for partial differential equations; the problem features – multiple solutions. This fact explains the complexity to solve the inverse problem, namely, one should, on the one hand, explore the solvability and Fréchet differentiability of the inverse problem and, on the other hand, apply methods of artificial intelligence to approximate its solution. The present paper employs the constructive neural network (neuroevolution of augmenting topologies) to this end.

2 Materials and Methods

2.1 Problem statement

The subject of the present study is the coefficient inverse problem for the nonlinear elliptic equations of von Karman-type. The respective direct problem reads as follows:

$$Q_1 \equiv \nabla_{ij}\left(A_1^{ijkl}(H_1)\nabla_{kl}u_1\right) - 1^{ik}1^{jl}\nabla_{kl}u_2\left(B_{ij}(H_1) + \nabla_{ij}u_1\right) = \lambda H_2\,, \tag{1}$$

$$Q_2 \equiv \nabla_{ij}\left(A_2^{ijkl}(H_1)\nabla_{kl}u_2\right) - 1^{ik}1^{jl}\nabla_{kl}u_1\left(B_{ij}(H_1) + \nabla_{ij}u_1\right) = 0\,, \tag{2}$$

$$Q = \{Q_1, Q_2\}^T,\ H = \{H_1, H_2\}^T,\ u = \{u_1, u_2\}^T$$

The definition domain $\Omega = \left\{X \middle| X = (x_1, x_2) \in R^2\right\}$ for solution vector-function $u(X, H(X))$ is assumed to be a bounded Lipchitz domain in R^2 with a boundary Γ, which is piecewise smooth. Here $A_1^{ijkl}(H_1) > 0$, $A_2^{ijkl}(H_1) > 0$, $B_{ij}(H_1) > 0$ are specified functionals of the known function H_1; H_2 is the right-hand member; $\nabla_i = \dfrac{\partial}{\partial x_i}$; $\nabla_{ij} = \dfrac{\partial^2}{\partial x_i \partial x_j}$; $i, j = 1,2$, λ is the parameter. The solution should satisfy the following boundary conditions

$$u_\Gamma = 0\,;\ \frac{\partial u_\Gamma}{\partial n} = u_n = 0\,, \tag{3}$$

where n is the unit normal vector to the surface Ω. Of fundamental importance is a set of admissible solutions for the direct and inverse problems defined as

$$\tilde{U}:\left|\begin{array}{l}\tilde{u}(X)\in V_{\Omega}^{12};\, Q(\tilde{u})=0\\ \tilde{u}\big|_{\Gamma}=0,\, \tilde{u}_n\big|_{\Gamma}=0\end{array}\right. ;\ \tilde{H}:\left|\begin{array}{l}H_{\min}\le H\le H_{\max};\, H\in W_{2\Omega}^{1}\\ a\le \dfrac{\partial H}{\partial u}\le b;\, \dfrac{\partial^2 H}{\partial u^2}\ge 0\end{array}\right. . \tag{4}$$

Hereinafter $W_{2\Omega}^{1}$ is the Sobolev space; V_{Ω}^{1}, V_{Ω}^{2}, V_{Ω}^{12} are the Hilbert spaces with norms and scalar products defined as

$$\begin{array}{l}
(u_1, v_1)_{V_{\Omega}^{1}} = \int\limits_{\Omega} A_1^{ijkl}(H)\nabla_{ij}u_1\nabla_{kl}v_1 d\Omega\\
\|u_1\|_{V_{\Omega}^{1}}^{2} = \int\limits_{\Omega} A_1^{ijkl}(H)\nabla_{ij}u_1\nabla_{kl}u_1 d\Omega\\
(u_2, v_2)_{V_{\Omega}^{2}} = \int\limits_{\Omega} A_2^{ijkl}(H)\nabla_{ij}u_2\nabla_{kl}v_2 d\Omega\\
\|u_2\|_{V_{\Omega}^{2}}^{2} = \int\limits_{\Omega} A_2^{ijkl}(H)\nabla_{ij}u_2\nabla_{kl}u_2 d\Omega\\
(u,v)_{V_{\Omega}^{12}} = [(u_1,u_2)(v_1,v_2)]_{V_{\Omega}^{12}} = (u_1,v_1)_{V_{\Omega}^{1}} + (u_2,v_2)_{V_{\Omega}^{2}}\\
\|u_2\|_{V_{\Omega}^{12}}^{2} = \|u_1\|_{V_{\Omega}^{1}}^{2} + \|u_2\|_{V_{\Omega}^{2}}^{2}
\end{array} \tag{5}$$

respectively. The functions that belong to the set $\tilde{H}$ are equibounded, monotonic, and convex, and, consequently, the set $\tilde{H}$ is a compact set. Input data for the inverse problem are known traces of the function u at the specified points γ_r

$$u(\gamma_r, H) = u^{*},\ r = 1, R\ . \tag{6}$$

In the framework of variational approach, this allows the following equivalent statement of the inverse problem in question

$$\rho_{V_{\Omega}^{12}} = \left(u(\gamma_r, H), u^{*}\right)_{V_{\Omega}^{12}},\ H\in\tilde{H},\ u^{*}\in V_{\Omega}^{12}. \tag{7}$$

The discrepancy functional (7) represents the distance in the space V_{Ω}^{12} between the calculated $u(\gamma_r, H)$ and specified u^{*} solutions of the direct problem. Thus, the solution of the inverse problem is defined as

$$H = \arg\min \rho_{V_{\Omega}^{12}}\left(u(\gamma_r, H), u^{*}\right),\ H\in\tilde{H},\ u\in\tilde{u}\ . \tag{8}$$

2.2 Properties of the inverse problem model

Of primary interest is the Fréchet differentiability of the vector-functions u with respect to H ; the respective proof utilizes definition of the generalized solution, namely: the vector-function $u = (u_1, u_2)^T \in V_{\Omega}^{12}$ is the generalized solution of the direct problem, if the following holds

$$(u_1, v_1)_{V_\Omega^1} = \int_\Omega \left[1^{ik}1^{jl}\left(B_{ij}v_1 - \nabla_i v_1 \nabla_j u_1\right)\nabla_{kl}u_2 - H_2 v_1\right]d\Omega,$$
$$(u_2, v_2)_{V_\Omega^2} = -\int_\Omega \left[1^{ik}1^{jk}\left(B_{ij}u_1 - \nabla_i u_1 \nabla_j u_1\right)\nabla_{kl}v_2\right]d\Omega, \tag{9}$$

for the arbitrary vector-function $v = (v_1, v_2)^T$, $v \in V_\Omega^{12}$. The Riesz theorem guarantees the existence of operators M, N that enables the operator representation of (9):

$$u_1 = N(u(H_1, H_2)),$$
$$u_2 = M(u(H_1, H_2)). \tag{10}$$

Vorovich [6] proves that these operators are strongly continuous. To prove their Fréchet differentiability, we consider a function H_0 (such that $u(H_0)$ is non-singular solution) and its perturbation $H = H_0 + \Delta H$ with $\|\Delta H\| \le \varepsilon$ (ε is small) and examine the respective solutions $\overset{(1)}{u}(H_0)$ and $\overset{(2)}{u}(H_0 + \Delta H) = \overset{(1)}{u}(H_0) + \Delta u$. The non-singularity ensures that $\mu = 1$ is not an eigenvalue of the following equations

$$\Delta u - \mu\, gradG(u)\big|_{u=u_0} \Delta u = 0, \tag{11}$$

where $\Delta u = \{\Delta u_1, \Delta u_2\}^T$, $G(u) = \{N(u), M(u)\}^T$. Since ε is small, the following holds true

$$\overset{(2)}{A}{}_p^{ijkl} = \overset{(1)}{A}{}_p^{ijkl} + \left.\frac{\partial A_p^{ijkl}}{\partial H}\right|_{H=H_0} \Delta H = \overset{(1)}{A}{}_p^{ijkl} + \Delta A^{ijkl}\Delta H;$$
$$\overset{(2)}{B}_{ij} = \overset{(1)}{B}_{ij} + \left.\frac{\partial B_{ij}}{\partial H}\right|_{H=H_0} \Delta H = \overset{(1)}{B}_{ij} + \Delta B_{ij}\Delta H. \tag{12}$$

Furthermore, spaces $\overset{(1)}{V}{}_\Omega^1$, $\overset{(1)}{V}{}_\Omega^2$ and $\overset{(2)}{V}{}_\Omega^1$, $\overset{(2)}{V}{}_\Omega^2$ symbolize the spaces associated with points like (u_0, H_0) and $(u_0 + \Delta u, H_0 + \Delta H)$, respectively. It is straightforward to prove that these spaces are equivalent and thereby each element of $u \in \overset{(1)}{V}{}_\Omega^p$ belongs to $\overset{(2)}{V}{}_\Omega^p$, and vice versa. Therefore, expressions (9) result in

$$\left(\overset{(2)}{u}_p, v_p\right)_{\overset{(2)}{V}{}_\Omega^p} = \left(\overset{(1)}{u}_p, v_p\right)_{\overset{(1)}{V}{}_\Omega^p} + \int_\Omega \left(\Delta A_p^{ijkl}\Delta H \nabla_{ij}u_p \nabla_{kl}v_p\right)d\Omega \tag{13}$$

and in

$$\left(\overset{(2)}{u}_p, v_p\right)_{\overset{(2)}{V}{}_\Omega^p} - \left(\overset{(1)}{u}_p, v_p\right)_{\overset{(1)}{V}{}_\Omega^p} \le m\|\Delta H\| \cdot \|u_p\|_{V_\Omega^p} \cdot \|v_p\|_{V_\Omega^p}, \quad p = 1,2. \tag{14}$$

With operators D_p [6] defined as

$$\left(D_p u_p, v_p\right)_{V_\Omega^p}^{(1)} = \left(u_p, v_p\right)_{V_\Omega^p}^{(2)}, \quad p = 1,2, \tag{15}$$

the inequality (14) yields to

$$1 - m\|\Delta H\| \le \left\|D_p\right\| \le 1 + m\|\Delta H\|, \quad p = 1,2. \tag{16}$$

If one estimates each term of the generalized solution using the embedding theorems, one finally obtains

$$\begin{aligned} \left\|\overset{(1)}{M} - \overset{(2)}{M}\right\|_{\overset{(1)}{V}{}^2_\Omega} &\le m_1\|\Delta H\|\left(1 + \|\varDelta u_1\|_{V_\Omega^1} + \|\varDelta u_1\|^2_{V_\Omega^1}\right); \\ \left\|\overset{(1)}{N} - \overset{(2)}{N}\right\|_{\overset{(1)}{V}{}^1_\Omega} &\le m_2\|\Delta H\|\left(1 + \|\varDelta u_1\|_{V_\Omega^1} + \|\varDelta u_1\|^2_{V_\Omega^1} + \|\varDelta u_1\|^3_{V_\Omega^1}\right), \end{aligned} \tag{17}$$

where m_1, m_2 are constants. Thus, the principal part of the increment of operators M and N are linearly proportional to $\|\Delta H\|$ and finite, and thereby the operators are the Fréchet differentiable.

2.3 Methods to solve the direct and inverse problems

To represent the unknown functions of the direct and inverse problems, we introduce two nets with nodes X_p, $p = \overline{1,P}$ and X_k, $k = \overline{1,K}$, respectively; specified values U^* are determined at the points X_r, $r = \overline{1,R}$.

To solve the direct problem, we use the finite element methods [7]. The definition domain is meshed, and solutions of the direct and inverse problems are approximated in the local coordinate system using their nodal values $\{U_p\}$, $\{U_p^*\}$, $\{H_k^*\}$, $p = \overline{1,P}$, $k = \overline{1,K}$. Thus, for the specified vector $H_0 = \{H_{0k}\}$, a solution of the direct problem (1)-(3) satisfies the finite system of nonlinear algebraic equations

$$K(U(H)) = Q. \tag{18}$$

Hereinafter, $U = (U_1, U_2,)^T$ is the vector of nodal values for the vector function u; $K = K_L + K_N$ is the transformation operator with the linear $K_L = [B_L]^T[D][B_L]$ and nonlinear $K_N = [B_N]^T[D][B_N]$ terms; B_L and B_N are shape functions matrices; Q is the vector of discretized right-hand terms.

In turn, a solution of the inverse problem satisfies (8) and, consequently, its discrete counterpart

$$H^* = \arg\min_{H \in \tilde{H}} \rho\left(U, U^*\right), \tag{19}$$

with $\rho = \sum_{r=1}^{R}\left(U_r(H_k) - U_r^*(H_k)\right)^T\left(U_r(H_k) - U_r^*(H_k)\right)$. This determines the function $H^* = Z\left(U^*\right)$, where $Z\left(U^*\right)$ is the approximate nonlinear mapping $U^* \to H^*$.

The Kolmogorov superposition theorem [8,9] states that any function of many variables can be represented as a cascade superposition of linear operations and nonlinear functions of a single variable. Hecht-Nielsen

[10] proposes to use multilayer perceptron (*MP*) to implement the structure of the Kolmogorov theorem (see [11] for a detailed description of MP and an extended discussion of its ability to implement this structure and approximate nonlinear mappings). For the problem under study, one clamps the known vector-function U^* to perceptron inputs and uses H^* as the corresponding desirable output values. The trained net approximates the discretized inverse operator $Z(U^*)$ with an error, which depends on perceptron structure, type of an activation function, training algorithm, representativeness of a sample, etc [11].

The principal problem of such approximation is that whether the discretized solution $H(U^*)$ is correctly determined at points U^* that do not belong to the learning sample used. To this end, one considers distance between elements U_1^* and U_2^* of the sample defined as $\rho(U_1^*,U_2^*)=\left\|Z(U_1^*)-Z(U_2^*)\right\|$. Vorovich proves [6] that if $\rho(U_1^*,U_2^*)\le\varepsilon$ and the non-singular solution $u^{(1)}$ corresponding to U_1^* exists for $H_1=Z(U_1^*)$ then the non-singular solution $u^{(2)}$ exists for $H_2=Z(U_2^*)$ likewise. Moreover, $\left\|u^{(1)}-u^{(2)}\right\|\le\delta(\varepsilon)$ and $\delta(\varepsilon)\to 0$ as $\varepsilon\to 0$.

The topological lemma states [6] that if the one-to-one correspondence $\Im:\tilde{H}\to\tilde{U}$ is continuous and the subset $H^*\subset\tilde{H}$ is dense in $\tilde{H}$, then the inverse one-to-one correspondence is also continuous $\Im^{-1}:U^*\to H^*$ on U^*. Therefore, if the learning sample H^* is compact and respective solutions u are non-singular then the lemma holds and one can correctly determine $H(u)$. Consequently, the inverse operator $H(U)$ can be determined for all non-singular nonlinear solutions $U(H)$ and it is possible to construct its neural network approximation.

It is worth stressing that a fully-connected perceptron is prone to redundant representation of the function to be approximated: connections between neurons that do not match those of real-world system tend to impair the efficiency of training algorithm, degrade solution quality and make it logically non-transparent. To overcome the problem, we employ neuro-evolution augmented topology (*NEAT*) [12] to generate the optimal network structure.

This method uses specific genetic algorithm framework to establish neural networks evolution. Thus, each chromosome corresponds to a neural network, while its gen describes a separate connection (its input and output neurons, its weight and activity). The network features an innovation list that is a tool to store information about newly-created neurons and connections. This allows unified description of all neurons and connections for all networks over all generations and thereby reasonable way to mutate the networks and cross them over. The generalization error (19) calculated for a specified epoch is used as a value of fitness function for a respective chromosome; a conventional back-propagation algorithm is used to train networks.

Of fundamental interest is an ability of the approach described above to approximate the inverse operator $H=H(U)$ for a strongly nonlinear function of the direct one $U(H)$, even in the neighbourhood of its limit points; of course, this suggests that the network has been trained using a sample selected over the representative domain of direct problem solution. The representativeness means that the domain should reflect variation of the model at hand.

3 Results and Discussion

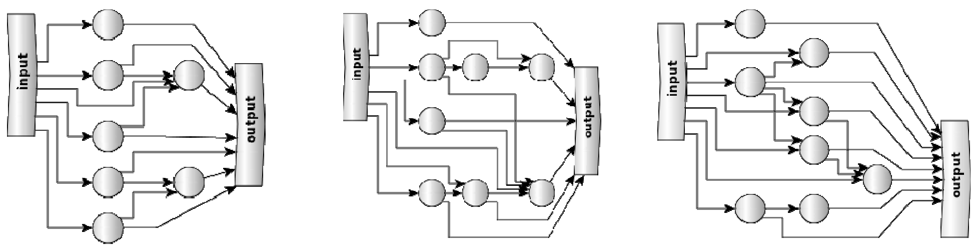
The method described above was applied to build neural-network approximations to determine unknown parameters of the right-hand terms for the rectangular definition domain $\Omega:\{X\,|\,X=(x_1,x_2),-a\le x_1\le a,-b\le x_2\le b\}$.

We considered three types of functions H_2:

$$H_A = \begin{cases} \lambda, & -\varphi \le x_2 \le \varphi, \\ 0, & otherwise, \end{cases},$$

$$H_B(X) = \lambda(0{,}5 + 0{,}5\cos x_2)^m, \quad m = 1{,}2\ldots,$$

$$H_C(X) = \lambda(0{,}5 + 0{,}5\cos m x_2), \quad m = 1{,}2\ldots.$$

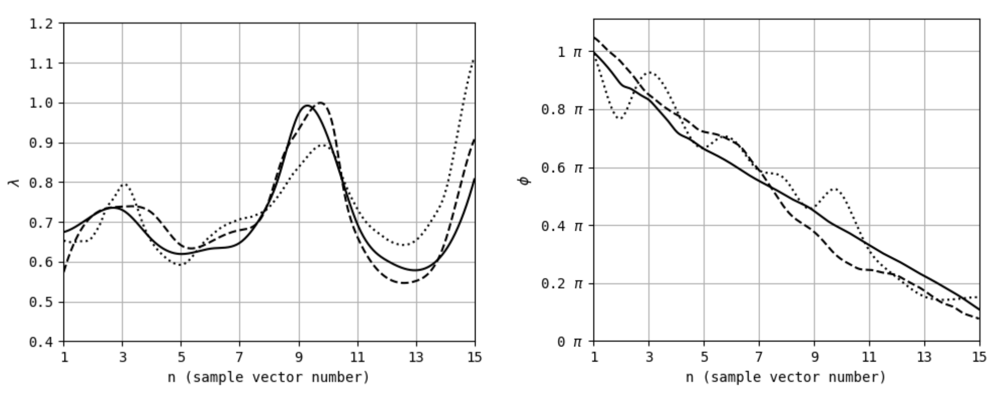


$H_A = \begin{cases} \lambda, & -\varphi \le x_2 \le \varphi, \\ 0, & otherwise, \end{cases}$ $\quad H_B(X) = \lambda(0{,}5 + 0{,}5\cos x_2)^m,$ $\quad H_C(X) = \lambda(0{,}5 + 0{,}5\cos m x_2),$

(a) **(b)** **(c)**

Fig. 1. The optimal structures of neural networks restoring the right-hand term of a certain class.
a - $H_A(X)$, *b* - $H_B(X)$, *c* - $H_C(X)$.

The wide-ranging simulation allows concluding that variability of the direct problem solution strongly affects the optimal network structure. Fig. 1 exhibits the optimal structures corresponding to the above right-hand terms. Fig. 2 shows actual and recovered values of parameters, namely, amplitudes of the right-hand member λ (Fig. 2a) and the spanning angles φ (Fig. 2b) for $H_A(X)$. The abscissa values correspond to various numbers of testing set; solid curves stand for actual values, whereas the values restored using MP and NEAT are designated by dotted and dashed curves, respectively.



(a) **(b)**

Fig. 2. True and recovered values of parameters for various amplitudes of the right-hand member
The abscissa values correspond to various numbers of testing set; solid, dotted and dashed curves stand for actual values and those restored by MP and NEAT, respectively. a - amplitudes of the right-hand member λ *;*
b - the spanning angles φ

Fig. 3 shows average errors (difference between actual and recovered values) ε for noisy direct problems vs. noise amplitude σ for the same parameters (Figs. 3a and b correspond to the restoration of λ and φ). The noise is uniformly distributed and perturbs $\{U_p^*\}$; its values (abscissa axis) represent its relative amplitude (with respect to amplitude of the perturbing quantity). Dotted and dashed curves denote the same characteristics. One may conclude that NEAT model performs better than MP; moreover, it performs two times better for the parameters values close to those corresponding to singular solutions – this means that the plain MP is inapplicable to such cases.

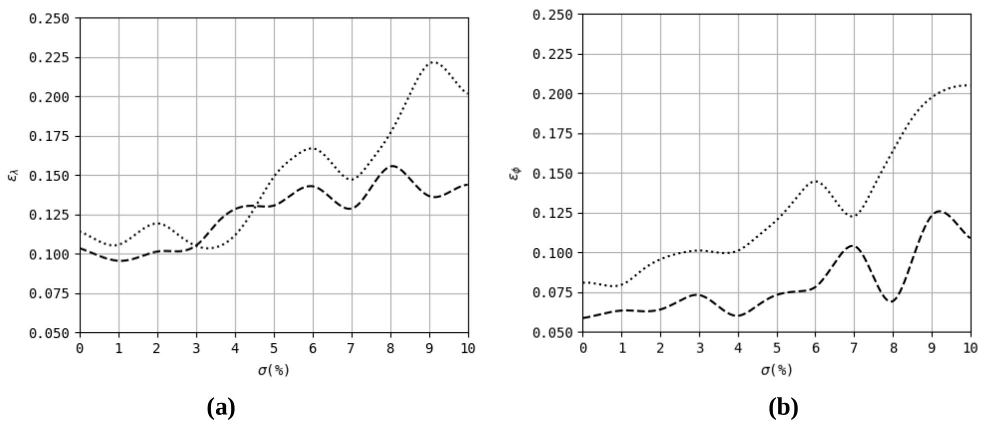


Fig. 3. Average errors (difference between actual and recovered values) ε for noisy direct problems vs. noise amplitude σ.

Dotted and dashed curves stand for values restored by MP and NEAT, respectively. a - amplitudes of the right-hand member λ; b - the spanning angles φ

4 Conclusion

The variational statement of the coefficient inverse problem for von Karman equations is introduced; the Fréchet differentiability of the inverse operator is proved.

In turn, the Fréchet differentiability allows neural network approximation of the inverse operator. To construct the optimal approximation, we apply neuro-evolution augmented topology model.

In the framework of wide-ranging simulation, NEAT model proves more efficient in comparison with conventional multilayer perceptron MP, and, more importantly, it allows approximation for the parameters values close to those corresponding to singular solutions of the direct problem.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Smołka M. Differentiability of the objective in a class of coefficient inverse problems. Comput. Math. Appl. 2017;73:2375–2387.

[2] Engl HW, Kügler P. Nonlinear inverse problems: Theoretical aspects and some industrial applications, in: Capasso and Periaux (Eds.), Multi-disciplinary Methods for Analysis, Optimization and Control of Complex Systems, Springer Heidelberg, Series Mathematics in Industry. 2005;3–48.

[3] Dierkes T, Dorn O, Natterer F, Palamodov V, Sielschott H. Fréchet derivatives for some bilinear inverse problems. SIAM J. Appl. Math. 2002;62(6):2092–2113.

[4] Yildirim Aksoy N. Variational method for the solution of an inverse problem. J. Comput. Appl. Math. 2017;312:82–93.

[5] Babaniyi OA, Oberai AA, Barbone PE. Direct error in constitutive equation formulation for plane stress inverse elasticity problem. Comput. Methods Appl. Mech. Engrg. 2017;314:3–18.

[6] Vorovich II. Nonlinear theory of shallow shells. Springer, New York; 1999.

[7] Wriggers P. Nonlinear finite elements method. Springer, New York; 2008.

[8] Kolmogorov AN. On representation of continuous functions of several variables as superpositions of continuous functions of one variable. Dokl. AN SSSR. 1957;114(5):953-956. (In Russian)

[9] Braun J, Griebel M. On a constructive proof of Kolmogorov's superposition theorem. Constr. Approx. 2009;30(3):653-675.

[10] Hecht-Nielsen R. Theory of back-propagation neural networks. Proceedings of IJCNN. 1989;1:583–604.

[11] Gupta M, Jin L, Homma N. Stability of continuous-time dynamic neural networks, in static and dynamic neural networks: From fundamentals to advanced theory. John Wiley & Sons, Inc., Hoboken, NJ, USA; 2003.

[12] Stanley KO, Miikkulainen R. Evolving neural networks through augmenting topologies. Evol. Comput. 2002;10(2):99–127.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/22077