Available online at www.sciencedirect.com

ScienceDirect

St. Petersburg Polytechnical University Journal: Physics and Mathematics 1 (2015) 192–198

www.elsevier.com/locate/spjpm

# Fresh approaches to the construction of parameterized neural network solutions of a stiff differential equation

Tatiana V. Lazovskaya*, Dmitry A. Tarkhov

*Peter the Great St. Petersburg Polytechnic University, 29 Politekhnicheskaya St., St. Petersburg 195251, Russian Federation*

Available online 14 August 2015

**Abstract**

A number of new fundamental problems expanding Vasiliev's and Tarkhov's methodology worked out for neural network models constructed on the basis of differential equations and other data has been stated and solved in this paper. The possibility of extending the parameter range in the same neural network model without loss of accuracy was studied. The influence of the new approach to choosing test points and using heterogeneous complementary data on the solution accuracy was analyzed.

The additional conditions in equation form derived from the asymptotic decomposition were used apart from the point data. The classical and non-classical definitions of the problem were compared by entering a parameter into the complementary data. A new sampling scheme of test point choice at different stages of minimization (the procedure of test point regeneration) under various initial conditions was investigated. A way of combining two approaches (classical and neural network) based on the Adams PECE method was considered.



*Keywords:* Neural network; Differential equation; Parameterized mathematical model; Hybrid algorithm.

## 1. Introduction

A methodology of designing neural network models from differential equations or other data (boundary conditions, measurements, etc.) developed by the St. Petersburg Polytechnic University professors Vasiliev and Tarkhov [3] allows solving complex and ill-posed problems of mathematical physics [4–7]. Those showing the most promise are the parameterized neural network models including one or several problem parameters as input variables [6–8] and allowing to simultaneously solve a family of problems with common parameters.

This paper raises and solves some new fundamental questions using a simple modeling task as an example.

First, we studied the possibility of extending the parameter variation range within a single neural network model without loss of accuracy, i.e. without increasing the pool of simultaneously solved tasks.

Second, we investigated how the new approach to choosing test points that we called a special test point regeneration influences solution accuracy.

Third, we continued the study in ref. [3] aimed at refining the solution through the use of heterogeneous complementary data. This is point data of the sought-for

* Corresponding author.
*E-mail addresses:* tatianala@list.ru (T.V. Lazovskaya), dtarkhov@gmail.com (D.A. Tarkhov).

http://dx.doi.org/10.1016/j.spjpm.2015.07.005

function, including the inaccurate data, which is often the case with real models.

The novel nature of the approach we have adopted is, compared with previous studies [3], that the above-mentioned point data was obtained by an intentionally inaccurate numerical method. Additionally, complementary conditions that are equations obtained through asymptotic decompositions are used along with the point data.

To answer the questions listed above, there was a good reason to primarily consider a simplest modeling problem with an analytical solution that the constructed approximate solutions could be compared to, and then objectively estimate the obtained results.

For this modeling task we chose a stiff first-order differential equation [1]. Studies [2–8] give reason to assume that the conclusions from the comparative analysis of the studied methods and algorithms remain valid for more complex tasks, including the problems of mathematical physics; so taking such a simple problem is justified.

Introducing a parameter into the complementary data of the problem (expressed through an equation) allows, in particular, to compare the classical and the non-classical statements of the problem. In the latter case, the conditions are imposed on the sought-for function outside the domain chosen for the solution. The natural asymptotic behavior of the studied problem is used as a starting point for such a condition. An approximate solution of the problem obtained through one of the classical methods serves as the inaccurate complementary point data.

A neural network consolidates the information both in data and equation forms using the minimizing functional reflecting the quality of a model. Additionally, in this paper we studied a new system of choosing test points at different stages of minimization (the test point regeneration procedure) for different types of input conditions.

## 2. Neural network models with complementary data

The problems that are commonly difficult to solve by classical explicit methods or require a lot of iterations are particularly interesting. Among the ordinary differential equations (DEs) these are stiff ones [1].

Ref. [1] deals with a classical example of a stiff equation

$$y' = -50(y - \cos x) \tag{1}$$

with an initial condition $y(0) = 0$.

When this problem is solved by the explicit Euler method, a critical value of the grid step equal to 2/50 occurs, above which the approximate solution becomes unstable with large variations (Fig. 1, *a*). At the same time, the error appears to be too large for a smaller step.

We shall focus on a generalized parameterized problem

$$y' = -\alpha(y - \cos x), \tag{2}$$

$$y(0) = 0,$$

where $\alpha \in [5, 50]$ or$\alpha \in [0.5, 50]$, $x \in [0, 1]$.

The problem is stiff for the variable $x$ in the vicinity of 0, which governs the choice of the proper intervals. Test runs showed that the quality of the neural network solution is also preserved for wider intervals. The problem is solved for all examined values of the parameter $\alpha$. Notice that these intervals of parameter variation are sufficiently wider than those discussed in refs. [6,8].

An approximate solution is sought in the form of an output of an artificial neural network of the given architecture:

$$y(x) = \sum_{i=1}^{n} c_i v(x, \alpha, \mathbf{a}_i),$$

whose weights $\{c_i, \mathbf{a}_i\}_{i=1}^{n}$ are determined when minimizing the error functional

$$\sum_{j=1}^{m} (y'(\xi_j) - F(\xi_j, y(\xi_j), \alpha_j))^2 + \delta y^2(0),$$

and for our case, $F(x, y, \alpha) = -\alpha(y - \cos x)$.

Test points $(\xi_j, \alpha_j)$ are chosen to be random and distributed uniformly over the examined intervals of variation of the value $x$ and the parameter $\alpha$; their choice is repeated after several (3–5) iterations of the optimization algorithm. We shall define a new random choice of test points at some step as test point regeneration.

The quality of the obtained solution is assessed from the exact analytical solution of Eq. (2) with an initial condition $y(0) = 0$, which takes the form

$$y(x, \alpha) = \frac{\alpha^2(\cos x - \exp(-\alpha x)) + \alpha \sin x}{\alpha^2 + 1}. \tag{3}$$

In the present work, we have examined two types of models corresponding to various basic functions with the varying number of neurons in the network. The first case involved choosing universal sigmoids in the form

$$\mathrm{th}[a(x - d)]\mathrm{th}[a_1(\alpha - d_1)],$$

and the second one asymmetric Gaussians in the form

$$x \exp[-a(x - d)^2] \exp[-a_1(\alpha - d_1)^2]$$

that were known to satisfy the initial condition.