Application of Neural Network Approach to Numerical Integration

Gregory A. Shipunov[1,*], Oksana I. Streltsova[1,2] and Yuri L. Kalinovski[1,2]

[1]Dubna State University, 19 Universitetskaya St, Dubna, 141980, Russian Federation
[2]Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation

Abstract

This paper is dedicated to the description of the application of neural network approach to numerical integration of functions of one and multiple variables. The essence of the approach is to train a neural network model to approximate the integral function and than use the parameters of the model to numerically calculate the value of the integral using the formulae based on those parameters. The usage of the approach will reduce the amount of calculations (and time) required to get a numerical integration result when the number of integral function's variables is big. Where the common numerical methods become too complex, the numerical approach allows calculations to be less demanding of the computational time and resources. This approach is being tested within the framework of a physics problem of modeling of the particles formation and their properties in the NICA experiment. In this experiment the key problem is to calculate integrals of functions of multiple variables. Currently the author of this paper is developing the framework for integration of functions of two variables. The main goal of the project though is to develop a Python library for numerical integration based on neural network approach.

Keywords

neural networks, numerical integration, meson, NICA, python library

1. Introduction

The problem of numerical integration appears in both theoretical and applied calculations and potentially raises two obstacles. The first one is the complexity of the analytical form of the anti-derivative of the integral function (which is required to get the value of the definite integral). Sometimes it is just hard to produce the anti-derivative of the function and sometimes there are none. In general case there is no general algorithm to get a given function's anti-derivative. Methods of numerical integration exist (e.g. Simpson's method) to solve this problem by providing a formulae to get the value of an integral with some error without knowing the anti-derivative. But as the dimensions of the integral function's grow, the complexity of those methods also grows as well as the error bound. In this case the neural network approach to the numerical integration may help reduce the complexity as well as increase accuracy of the calculation.

Currently, there were several works done in the field of application of neural networks to the numerical integration problem. We here will highlight the paper "Using neural networks for fast numerical integration and optimization" by Lloyd et al.[1]. This work was served as an inspiration for the development of the neural network approach. Another source was the physics problem of modeling of the particles formation and their properties in the NICA experiment which involves solving the equations containing double and triple integrals. The development of this neural network approach to the numerical integration has the main goal of

Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems 2025 (ITTMM 2025), Moscow, April 07–11, 2025
*Corresponding author.
shgregory3@gmail.com (G. A. Shipunov); strel@jinr.ru (O. I. Streltsova); kalinov@jinr.ru (Y. L. Kalinovski)
0009-0007-7819-641X (G. A. Shipunov); 0000-0003-4522-6735 (O. I. Streltsova); 0000-0002-7596-5531 (Y. L. Kalinovski)

the development of a software library for the Python programming language with functionality of numerical integration based on neural network approach.

This paper contains several parts. Firstly, we will give the description of one of the possible neural network approaches to the numerical integration as well as highlight other potential approaches. Secondly, the physics problem of modeling of the of the particles formation and their properties in the NICA experiment will be depicted. Thirdly, we will give the observation on the results already reached in the development of the neural network approach to the numerical integration. Lastly, the insight on the future of this project will be given.

2. Neural Network Approach to Numerical Integration

The problem we are investigating in this paper is the application of neural network approach to the calculation of approximate value of integral of a function, i.e. numerical integral with usage of neural network approach.

Let a continuous real function $f(\mathbf{x})$ be defined as $f : \mathbb{R}^n \to \mathbb{R}$. Let Ω be a compact subset of $\mathbb{R}^n$ and let G be a bounded convex subset of Ω. Let, also, $\mathbf{x}$ be a vector of n dimensions in Ω. Than

$$I(f) = \int_G f(\mathbf{x}) d\mathbf{x}, \tag{1}$$

is a definite integral of f across set G. Therefore, the problem is to get the $I(f)$ value calculated with use of neural network approach.

The work of Lloyd et al.[1] proposes, that function $f(\mathbf{x})$ can be approximated using an MLP (multilayer perceptron) neural network. Let this network contain of 3 layers with sizes n, k and 1 accordingly. n is defined above as a number of dimensions of the function $f(\mathbf{x})$, k is the hidden-layer size which is set to minimize the approximation error and 1 neuron of the output layer accumulates the approximated $f(\mathbf{x})$ value. Because this value is not equal to the value $f(\mathbf{x})$ with given $\mathbf{x}$, let us denote it as $\hat{f}(\mathbf{x})$ and further on we will refer to the MLP network as $\hat{f}(\mathbf{x})$. Figure 1 depicts the structure of the $\hat{f}(\mathbf{x})$ network.

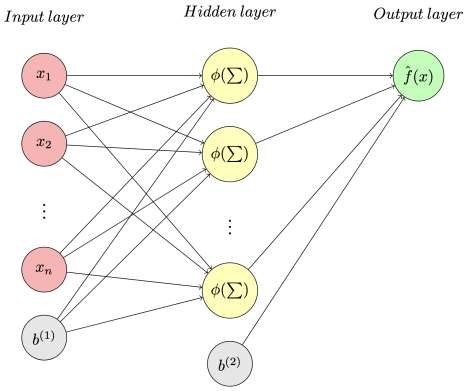


Figure 1: The MLP structure used in the neural network approach

MLP network can have activation functions applied to each neuron's value before this value propagates forward. In our case both input and output layer has *linear activation function* applied and the hidden layer has *logistic sigmoid function* applied. Last has the following definition:

$$\phi(z) = \frac{1}{1 + \exp(-z)}. \tag{2}$$

With $\phi(z)$ defined, network structure's mathematical form is:

$$\hat{f}(x) = b^{(2)} + \sum_{j=1}^{k} w_j^{(2)} \phi(b_j^{(1)} + \sum_{i=1}^{n} w_{ji}^{(1)} x_i). \tag{3}$$

So here Lloyd's et al. work puts the proposition to use the trained neural network $\hat{f}(\mathbf{x})$ to get the approximate value of integral (1):

$$\hat{I}(f) = \int_G \hat{f}(\mathbf{x}) d\mathbf{x}. \tag{4}$$

To integrate the (3) and get the value of $\hat{I}(f)$ let us apply following substitution:

$$-Li_0(-\exp(z)) = \frac{1}{1+\exp(-z)} = \phi(z), \tag{5}$$

where $Li_0(u(z))$ is a Jonquière's function or the polylogarithm of order 0:

$$Li_0(u) = \frac{u}{1-u}. \tag{6}$$

This substitution is proved to be numerically accurate in the Lloyd's et al. work[1]. With substitution (5) network (3) can be integrated over given $G : [\alpha_1, \beta_1] \times ... \times [\alpha_n, \beta_n]$ in $\mathbb{R}^n$. to produce the following numerical integration formulae:

$$\hat{I}(f) = b^{(2)} \prod_{i=1}^{n} (\beta_i - \alpha_i) + \sum_{j=1}^{k} w_j^{(2)} [\prod_{i=1}^{n} (\beta_i - \alpha_i) + \frac{\Phi_j}{\prod_{i=1}^{n} w_{ij}^{(1)}}], \tag{7}$$

$$\Phi_j = \sum_{r=1}^{2^n} \xi_r Li_n(-exp[-b_j^{(1)} - \sum_{i=1}^{n} w_{ij}^{(1)} l_{i,r}]), \tag{8}$$

$$\xi_r = \prod_{d=1}^{n} (-1)^{[r/2^{n-d}]}, \tag{9}$$

$$l_{i,r} = \begin{cases} \alpha_i, & \text{if } [r/2^{n-d}]\%2 = 0 \\ \beta_i, & \text{if } [r/2^{n-d}]\%2 \neq 0, \end{cases} \tag{10}$$

where $b^{(1)}, b^{(2)}, w^{(1)}, w^{(2)}$ – neural network's parameters, $\alpha_i, \beta_i, 1 \leq i \leq n$ – integration boundaries for each dimension, n – number of $f(x)$ dimensions.

The described neural network approach to numerical integration concludes to following steps:

1. generate the dataset to train, validate and test the neural network model (3) to approximate given integral function;
2. train, validate and test (3) to approximate given integral function;
3. extract parameters of the network (3) and use them with (7-10) to calculate the numerical integral.

3. Modeling of the formation of particles and their properties for the NICA experiment

The main task is to build a model for the unified description of both light and heavy mesons formed in heavy ion collisions.

This will be a nonlocal interaction model [2, 3]. The model concludes to the fact that meson (quark-antiquark pair) formation proceeds with a gluon exchange. But the gluon exchange cannot be described as a diagram sum, because perturbation theory cannot be applied here.

Mesons can be described as $q\bar{q}$ bound states by the Bethe-Salpeter equation

$$\Gamma(q,P) = -\frac{4}{3}\int \frac{d^4p}{(2\pi)^4} D(q-p)\gamma_\alpha S_1 \Gamma(q,P) S_2 \gamma_\alpha \,. \tag{11}$$

The vertex function $\Gamma(p,P)$ depends on the relative momentum p, and the total momentum of the bound state P. $S_i(p)$ are the dressed quark propagator in the Euclidean space

$$S_i(p_i) = \frac{1}{i(p_i\cdot\gamma) + m_i} \tag{12}$$

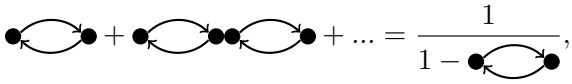
The momenta $p_i = p + q_i$, $q_i = b_i P$, $i = 1,2$ with $b_1 = -m_1/(m_1+m_2)$, $b_2 = m_2/(m_1+m_2)$, m_i are the constituent quark masses.

Equation (11) contains the interaction kernel $D(q-p)$ which describes the effective gluon interaction within a meson. We consider the rank -1 separable model

$$D(q-p) = D_0 F(q^2) F(p^2)\,, \tag{13}$$

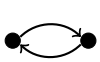
where D_0 is the coupling constant and the function $F(p^2)$ is related to scalar part of Bethe - Salpeter vertex function. We employ $F(q^2)$ in the Gaussian form $F(p^2) = e^{-p^2/\lambda^2}$ with the parameter λ which characterizes the finite size of the meson.

This separable Anzatz of the interaction kernel (13) allows to write the meson observables in the term of the polarization operator (the buble diagram):



$$= \frac{1}{1-}, \tag{14}$$

where



$$\to \int_0^\infty \frac{dp}{2\pi^4} F(p^2)\frac{1}{u_1 u_2} = \int_0^\infty \frac{dp}{2\pi^4} F(p^2)\frac{1}{[(p+q_1)^2+m_1^2][(p+q_2)^2+m_2^2]}. \tag{15}$$

To calculate these integrals we use the equations:

$$\frac{1}{(p+q_i)^2+m_i^2} = \int_0^\infty dt\{\exp(-t[(p+q_i)^2+m_i^2])\}, \tag{16}$$

$$F(p^2) = \int_0^\infty ds\{\exp(-sp^2)F(s)\}. \tag{17}$$

The case of more than two mesons is described using triple and quadruple integrals.

The amplitudes of particular decays and the squared amplitudes of particular decays are calculated using FORM (software for analytical calculation). But the key problem is to calculate integrals of functions of multiple variables.

In the framework of this problem there is a particular iterated integral:

$$\int_0^1 d\alpha\{\alpha^a(1-\alpha)^b\}\int_0^\infty dt\{\frac{t^m}{(1+t)^n}F[z_0]\} \equiv I(a,b,m,n;F[z_0]), \tag{18}$$

$$F[z_0] = \exp[-2z_0], \tag{19}$$

$$z_0 = tD + \frac{t}{1+t}R^2, \tag{20}$$

$$D = \alpha_1(b_1^2P^2+m_1^2) + \alpha_2(b_2^2P^2+m_2^2), \tag{21}$$

$$R^2 = (\alpha_1^2 b_1^2 + \alpha_2^2 b_2^2 + 2\alpha_1\alpha_2 b_1 b_2)P^2, \tag{22}$$

$$b_1 = -\frac{m_1}{m_1+m_2}, \tag{23}$$

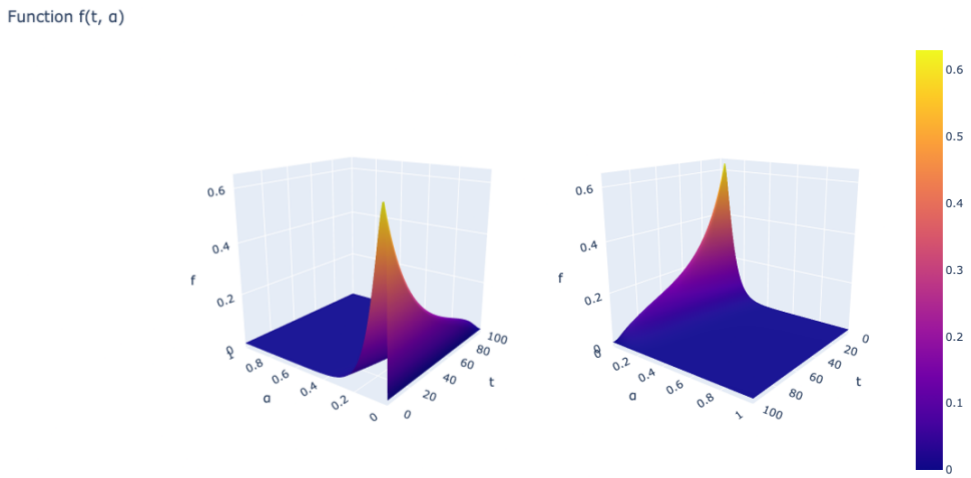


Figure 2: Integrand function (18) plot - view from (0, 0, 0) and (100, 0, 0) corners

$$b_2 = \frac{m_2}{m_1 + m_2}, \tag{24}$$

$$\alpha_1 = \alpha, \alpha_2 = 1 - \alpha. \tag{25}$$

This integral is used in multiple calculations in the problem. Shape of its 2-dimensional integral is sown on Figure 2. As it can be seen in (18) it is denoted as $I(a, b, m, n; F[z_0])$, and with different combination of parameters it is used multiple times in the calculations. The integral (18) also is interesting in that case, that it depends on two variables only, so it is can be used to test the integration framework we've been developing early in the development cycle.

4. Approach testing

TEXT

5. Usage of Neural Network Approach

As stated in the previous part, the integral (18) has been already used to test the neural network approach.

It was firstly interpreted as a product of two integrals of single-variable functions ($\alpha_{1,2}$ were treated as constants). The first integral of function $\alpha^a(1-\alpha)^b$ has a very simple analytical anti-derivative with any integer parameter a and b so it was easy to check the result of the integration. In most cases the neural network approach provided error bound $\varepsilon < 10^{-5}$. On the other hand, the second integral of function $\frac{t^m}{(1+t)^n}F[z_0]$ has a much more complex anti-derivative, so in this case the goal was to make the approximation bound of neural network $\hat{f}(x)$ was $\epsilon < 10^{-5}$ and this requirement was met.

Secondly, the iterated integral (18) was interpreted as a double integral:

$$\int_0^1 \alpha^a(1-\alpha)^b d\alpha \int_0^\infty \frac{t^m}{(1+t)^n} F[z_0] dt = \int_0^1 \int_0^\infty \frac{t^m}{(1+t)^n} F[z_0] \alpha^a (1-\alpha)^b dt d\alpha, \tag{26}$$

and this double integral function was used to train the neural network model to approximate it. The approximation error bound of $\epsilon < 10^{-5}$ was also met, but currently the software implementation of the formulae (7-10) requires further improvement and testing.

Also both 1-dimensional ans 2-dimensional functionality was tested on simple arbitrary functions, e.g. transcendental functions.

6. Future development

Currently, the further development of the neural network approach to the numerical integration continues. The resulting software will be tested on the integrals from the meson interaction modeling problem, which provides plenty of cases where numerical integration of high-dimensional functions is required. Also, we hope, the final product - the software library for numerical integration with use of numerical integration approach called *Skuld* will help to solve difficult calculation problems in the meson interaction modeling problem domain as well, as in other fields, both theoretical and applied.

7. Acknowledgments

Authors of this work express their acknowledgments to Sara Shadmehri and Tatevik Bezhanyan from Meshcheryakov Laboratory of Information Technologies, JINR for their advices both in the field of neural networks and English language.

References

[1] S. Lloyd, R. A. Irani, M. Ahmadi, Using neural networks for fast numerical integration and optimization, IEEE Access 8 (2020) 84519–84531.

[2] S. Schmidt, D. Blaschke, Y. L. Kalinovsky, Scalar-pseudoscalar meson masses in nonlocal effective qcd at finite temperature, Physical Review C 50 (1994) 435.

[3] P. Costa, M. C. Ruivo, Y. L. Kalinovsky, Pseudoscalar neutral mesons in hot and dense matter, Physics Letters B 560 (2003) 171–177.