



ИТОГИ НАУКИ И ТЕХНИКИ.
Современная математика и ее приложения.
Тематические обзоры.
Том 237 (2024). С. 76–86
DOI: 10.36535/2782-4438-2024-237-76-86

УДК 517.972.7, 004.032.26

ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ ГИПЕРБОЛИЧЕСКОГО УРАВНЕНИЯ ПРИ ПОМОЩИ КВАЗИКЛАССИЧЕСКОГО ФУНКЦИОНАЛА

© 2024 г. С. Г. ШОРОХОВ

Аннотация. Рассматривается задача построения функционала потерь на основе квазиклассического вариационного принципа для обучения нейронной сети, аппроксимирующей решения гиперболического уравнения. При помощи метода симметризирующего оператора В. М. Шалова построен вариационный функционал краевой задачи для гиперболического уравнения второго порядка, содержащий интегралы по области краевой задачи и фрагменту ее границы, зависящие от производных первого порядка неизвестной функции. Показано, что нейронная сеть, аппроксимирующая решение рассматриваемой краевой задачи, может быть обучена с применением построенного вариационного функционала.

Ключевые слова: вариационный принцип, гиперболическое уравнение, нейронная сеть, функционал потерь.

TRAINING A NEURAL NETWORK FOR A HYPERBOLIC EQUATION BY USING A QUASICLASSICAL FUNCTIONAL

© 2024 S. G. SHOROKHOV

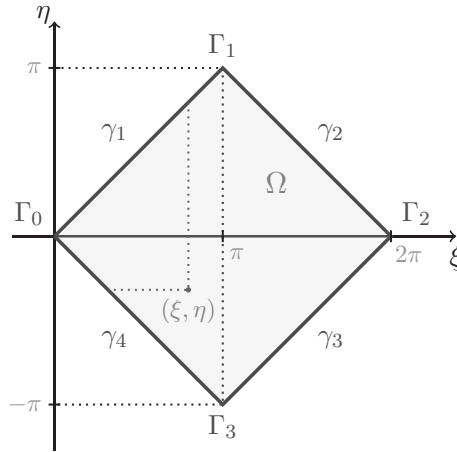
ABSTRACT. We study the problem of constructing a loss functional based on the quasiclassical variational principle for training a neural network, which approximates solutions of a hyperbolic equation. Using the method of symmetrizing operator proposed by V. M. Shalov, for the second-order hyperbolic equation, we construct a variational functional of the boundary-value problem, which involves integrals over the domain of the boundary-value problem and a segment of the boundary, depending on first-order derivatives of the unknown function. We demonstrate that the neural network approximating the solution of the boundary-value problem considered can be trained by using the constructed variational functional.

Keywords and phrases: variational principle, hyperbolic equation, neural network, loss functional.

AMS Subject Classification: 35A15, 68T07

1. Введение. Вариационные принципы для непотенциальных операторов (см. [3]) применяются для решения широкого круга теоретических и прикладных задач, моделируемых уравнениями с частными производными. Современным подходом к решению краевых задач для уравнений с частными производными является аппроксимация решений при помощи искусственных нейронных сетей различной архитектуры (см. [9]). Для обучения нейронных сетей требуются функционалы потерь (ошибки), в качестве которых для нейронных сетей, аппроксимирующих решения дифференциальных уравнений, обычно применяются функционалы невязки (см. [9, 10]), реже

Работа выполнена в рамках проекта № 002092-0-000 Российского университета дружбы народов имени Патриса Лумумбы.

Рис. 1. Область Ω краевой задачи (1)–(2)

энергетические функционалы (потенциалы; см. [7]). Однако, как отмечалось в [8, 12], для обучения нейронных сетей также могут быть использованы квазиклассические функционалы теории вариационных принципов для непотенциальных операторов (см. [3]).

Квазиклассический функционал краевой задачи для гиперболического уравнения с частными производными был получен В. М. Шаловым (см. [6]), альтернативный квазиклассический функционал краевой задачи для гиперболического уравнения с другим набором граничных условий был построен В. М. Филипповым (см. [2]).

2. Построение вариационной формулировки гиперболического уравнения. Метод построения вариационных формулировок для краевых задач, предложенный В. М. Шаловым (см. [4, 5]), состоит в построении таких двух векторных операторов $\mathbf{A}$ и $\mathbf{B}$ в гильбертовом пространстве H , что оператор $\mathbf{A}$ является $\mathbf{B}$ -симметричным:

$$\langle \mathbf{A}u, \mathbf{B}v \rangle = \langle \mathbf{B}u, \mathbf{A}v \rangle, \quad u \in H, \quad v \in H,$$

и $\mathbf{B}$ -положительным:

$$\langle \mathbf{A}u, \mathbf{B}u \rangle > 0, \quad u \in H, \quad u \neq 0,$$

$$\langle \mathbf{A}u_n, \mathbf{B}u_n \rangle \xrightarrow{n \rightarrow \infty} 0, \quad u_n \in H \quad \Rightarrow \quad \|u_n\|_H \xrightarrow{n \rightarrow \infty} 0,$$

причем краевая задача (уравнение с частными производными и граничные условия) должна быть представлена в виде

$$\mathbf{A}u = \mathbf{f},$$

где $\mathbf{f}$ — вектор-функция. Тогда вариационный функционал для рассматриваемой краевой задачи записывается в виде

$$D[u] = \langle \mathbf{A}u, \mathbf{B}u \rangle - 2 \langle \mathbf{f}, \mathbf{B}u \rangle.$$

В. М. Шалов сформулировал и исследовал вариационную задачу для неоднородного уравнения гиперболического типа, заданного в области достаточно общего вида со смешанными граничными условиями, заданными на всей границе, установил существование и единственность решения вариационной и краевой задач (см. [6]).

Рассмотрим область Ω в форме ромба с вершинами в точках $\Gamma_0(0, 0)$, $\Gamma_1(\pi, \pi)$, $\Gamma_2(2\pi, 0)$, $\Gamma_3(\pi, -\pi)$, где отрезок γ_1 соединяет точки $\Gamma_0(0, 0)$ и $\Gamma_1(\pi, \pi)$, отрезок γ_2 соединяет точки $\Gamma_1(\pi, \pi)$ и $\Gamma_2(2\pi, 0)$, отрезок γ_3 соединяет точки $\Gamma_2(2\pi, 0)$ и $\Gamma_3(\pi, -\pi)$, отрезок γ_4 соединяет точки $\Gamma_3(\pi, -\pi)$ и $\Gamma_0(0, 0)$ (см. рис. 1). Область Ω не соответствует ограничениям на область краевой задачи, установленным в [6], однако подход В. М. Шалова к построению вариационного функционала применим и для области Ω на рис. 1.

Рассмотрим построение вариационной формулировки для однородного гиперболического уравнения

$$u_{\xi\eta} = 0, \quad (\xi, \eta) \in \Omega, \quad (1)$$

в области Ω с граничными условиями

$$u|_{\gamma_1} = \chi_1(\xi), \quad u_\eta|_{\gamma_2} = \varphi_2(\eta), \quad u_\eta|_{\gamma_3} = \varphi_3(\eta), \quad u_\xi|_{\gamma_3} = \psi_3(\xi), \quad u_\xi|_{\gamma_4} = \psi_4(\xi). \quad (2)$$

Здесь и в дальнейшем использование греческих символов в качестве нижнего индекса означает дифференцирование по соответствующей переменной.

Решение задачи (1)–(2) ищется в классе функций $W_2^1(\Omega)$, при этом предполагается, что функции в правых частях (2) измеримы, т.е. $(\chi_1, \varphi_2, \varphi_3, \psi_3, \psi_4) \in L_2(\gamma_1 \times \gamma_2 \times \gamma_3 \times \gamma_3 \times \gamma_4)$, и граничные условия (2) совместны.

Компоненты внешней нормали $\mathbf{n}$ к границе $\partial\Omega$ на различных участках границы $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ вычисляются по формулам:

$$\mathbf{n}(\gamma_1) = \frac{1}{\sqrt{2}}(-1, 1), \quad \mathbf{n}(\gamma_2) = \frac{1}{\sqrt{2}}(1, 1), \quad \mathbf{n}(\gamma_3) = \frac{1}{\sqrt{2}}(1, -1), \quad \mathbf{n}(\gamma_4) = \frac{1}{\sqrt{2}}(-1, -1).$$

Следуя [6], построим векторные операторы

$$\mathbf{A}, \mathbf{B} : W_2^1(\Omega) \rightarrow L_2(\Omega \times \gamma_1 \times \gamma_2 \times \gamma_3 \times \gamma_3 \times \gamma_4)$$

следующим образом. Первую компоненту векторного оператора $\mathbf{A}$, действующую на Ω , определим формулой

$$(\mathbf{A}u)_1 = u_{\xi\eta}, \quad (\xi, \eta) \in \Omega.$$

Соответствующую первую компоненту вспомогательного (симметризирующего) векторного оператора $\mathbf{B}$ на Ω определим равенством

$$(\mathbf{B}v)_1 = \int_{\xi}^{\gamma_1 \cup \gamma_4} v_\eta(\zeta, \eta) d\zeta + \int_{\eta}^{\gamma_1 \cup \gamma_2} v_\xi(\xi, \tau) d\tau, \quad (\xi, \eta) \in \Omega.$$

Здесь интегрирование по переменной ζ ведется при фиксированном значении η от значения ξ до соответствующего значения на одной из кривых γ_1 или γ_4 (в зависимости от значения η), а интегрирование по переменной τ ведется при фиксированном значении ξ от значения η до соответствующего значения на одной из кривых γ_1 или γ_2 .

Последующие компоненты векторных операторов $\mathbf{A}$ и $\mathbf{B}$, действующие на участках границы $\gamma_1, \gamma_2, \gamma_3, \gamma_3, \gamma_4$ соответственно, определим формулами:

$$\begin{aligned} (\mathbf{A}u)_2 &= u, \quad (\mathbf{B}v)_2 = v, & (\xi, \eta) &\in \gamma_1, \\ (\mathbf{A}u)_3 &= u_\eta, \quad (\mathbf{B}v)_3 = -n_1(\gamma_2) \int_{\xi}^{\gamma_1} v_\eta(\zeta, \eta) d\zeta, & (\xi, \eta) &\in \gamma_2, \\ (\mathbf{A}u)_4 &= u_\eta, \quad (\mathbf{B}v)_4 = -n_1(\gamma_3) \int_{\xi}^{\gamma_4} v_\eta(\zeta, \eta) d\zeta, & (\xi, \eta) &\in \gamma_3, \\ (\mathbf{A}u)_5 &= u_\xi, \quad (\mathbf{B}v)_5 = -n_2(\gamma_3) \int_{\eta}^{\gamma_2} v_\xi(\xi, \tau) d\tau, & (\xi, \eta) &\in \gamma_3, \\ (\mathbf{A}u)_6 &= u_\xi, \quad (\mathbf{B}v)_6 = -n_2(\gamma_4) \int_{\eta}^{\gamma_1} v_\xi(\xi, \tau) d\tau, & (\xi, \eta) &\in \gamma_4. \end{aligned}$$

Здесь $n_i(\gamma_j)$ — i -я компонента внешней нормали $\mathbf{n}$ к участку границы γ_j .

Итак, векторный дифференциальный оператор $\mathbf{A}$, векторный интегродифференциальный оператор $\mathbf{B}$ и вектор-функция $\mathbf{f}$ определяются равенствами

$$\mathbf{A}u = \begin{bmatrix} u_{\xi\eta} \\ u \\ u_{\eta} \\ u_{\eta} \\ u_{\xi} \\ u_{\xi} \end{bmatrix}, \quad \mathbf{B}v = \begin{bmatrix} \int_{\xi}^{\gamma_1 \cup \gamma_4} v_{\eta}(\zeta, \eta) d\zeta + \int_{\eta}^{\gamma_1 \cup \gamma_2} v_{\xi}(\xi, \tau) d\tau \\ v \\ -n_1(\gamma_2) \int_{\xi}^{\gamma_1} v_{\eta}(\zeta, \eta) d\zeta \\ -n_1(\gamma_3) \int_{\xi}^{\gamma_4} v_{\eta}(\zeta, \eta) d\zeta \\ -n_2(\gamma_3) \int_{\eta}^{\gamma_2} v_{\xi}(\xi, \tau) d\tau \\ -n_2(\gamma_4) \int_{\eta}^{\gamma_1} v_{\xi}(\xi, \tau) d\tau \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} 0 \\ \chi_1 \\ \varphi_2 \\ \varphi_3 \\ \psi_3 \\ \psi_4 \end{bmatrix}.$$

Вектор-функции $\mathbf{A}u$, $\mathbf{B}v$ и $\mathbf{f}$ определены на декартовом произведении $\Omega \times \gamma_1 \times \gamma_2 \times \gamma_3 \times \gamma_3 \times \gamma_4$.

Скалярное произведение векторов $\mathbf{A}u$ и $\mathbf{B}v$ вычисляется путем перемножения соответствующих компонент, вычисления интегралов по соответствующим множествам и суммирования получившихся значений по формуле

$$\begin{aligned} \langle \mathbf{A}u, \mathbf{B}v \rangle = & \int_{\Omega} (\mathbf{A}u)_1 (\mathbf{B}v)_1 d\xi d\eta + \int_{\gamma_1} (\mathbf{A}u)_2 (\mathbf{B}v)_2 ds + \int_{\gamma_2} (\mathbf{A}u)_3 (\mathbf{B}v)_3 ds + \\ & + \int_{\gamma_3} (\mathbf{A}u)_4 (\mathbf{B}v)_4 ds + \int_{\gamma_3} (\mathbf{A}u)_5 (\mathbf{B}v)_5 ds + \int_{\gamma_4} (\mathbf{A}u)_6 (\mathbf{B}v)_6 ds. \end{aligned}$$

Вычисление скалярного произведения векторов $\mathbf{A}u$ и $\mathbf{B}v$ при помощи формулы многомерного интегрирования по частям [11] дает равенство

$$\begin{aligned} \langle \mathbf{A}u, \mathbf{B}v \rangle = & \int_{\Omega} (u_{\xi} v_{\xi} + u_{\eta} v_{\eta}) d\xi d\eta + \int_{\gamma_2} u_{\eta} \int_{\xi}^{\gamma_1} v_{\eta}(\zeta, \eta) d\zeta n_1(\gamma_2) ds + \\ & + \int_{\gamma_3} \left[u_{\eta} \int_{\xi}^{\gamma_4} v_{\eta}(\zeta, \eta) d\zeta n_1(\gamma_3) + u_{\xi} \int_{\eta}^{\gamma_2} v_{\xi}(\xi, \tau) d\tau n_2(\gamma_3) \right] ds + \int_{\gamma_4} u_{\xi} \int_{\eta}^{\gamma_1} v_{\xi}(\xi, \tau) d\tau n_2(\gamma_4) ds + \\ & + \int_{\gamma_1} uv ds - \int_{\gamma_2} u_{\eta} n_1(\gamma_2) \int_{\xi}^{\gamma_1} v_{\eta}(\zeta, \eta) d\zeta ds - \int_{\gamma_3} u_{\eta} n_1(\gamma_3) \int_{\xi}^{\gamma_4} v_{\eta}(\zeta, \eta) d\zeta ds - \\ & - \int_{\gamma_3} u_{\xi} n_2(\gamma_3) \int_{\eta}^{\gamma_2} v_{\xi}(\xi, \tau) d\tau ds - \int_{\gamma_4} u_{\xi} n_2(\gamma_4) \int_{\eta}^{\gamma_1} v_{\xi}(\xi, \tau) d\tau ds, \end{aligned}$$

в котором все криволинейные интегралы, кроме интеграла по γ_1 , сокращаются, и формула для скалярного произведения приобретает следующий простой вид:

$$\langle \mathbf{A}u, \mathbf{B}v \rangle = \int_{\Omega} (u_{\xi} v_{\xi} + u_{\eta} v_{\eta}) d\xi d\eta + \int_{\gamma_1} uv ds,$$

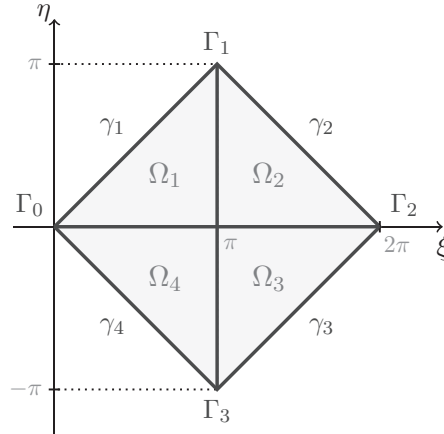


Рис. 2. Подобласти $\Omega_1, \Omega_2, \Omega_3, \Omega_4$ области Ω краевой задачи (1)–(2)

из которого вытекает $\mathbf{B}$ -симметричность оператора $\mathbf{A}$. Далее имеем

$$\langle \mathbf{A}u, \mathbf{B}u \rangle = \int_{\Omega} (u_{\xi}^2 + u_{\eta}^2) d\xi d\eta + \int_{\gamma_1} u^2 ds,$$

откуда, очевидно, вытекает $\mathbf{B}$ -положительность оператора $\mathbf{A}$.

Теорема 1. *Вариационный функционал В. М. Шалова (см. [6]) для рассматриваемой краевой задачи (1)–(2) имеет вид*

$$\begin{aligned} D[u] = & \int_{\Omega} (u_{\xi}^2 + u_{\eta}^2) d\xi d\eta + \int_{\gamma_1} u^2 ds - 2 \int_{\gamma_1} \chi_1 u ds + \\ & + 2 \int_{\gamma_1} \chi_1 u ds + 2 \int_{\gamma_2} \varphi_2 n_1(\gamma_2) \int_{\xi}^{\gamma_1} u_{\eta}(\zeta, \eta) d\zeta ds + 2 \int_{\gamma_3} \varphi_3 n_1(\gamma_3) \sin \xi \int_{\xi}^{\gamma_4} u_{\eta}(\zeta, \eta) d\zeta ds + \\ & + 2 \int_{\gamma_3} \psi_3 n_2(\gamma_3) \int_{\eta}^{\gamma_2} u_{\xi}(\xi, \tau) d\tau ds + 2 \int_{\gamma_4} \psi_4 n_2(\gamma_4) \int_{\eta}^{\gamma_1} u_{\xi}(\xi, \tau) d\tau ds. \quad (3) \end{aligned}$$

При использовании в качестве функционала потерь интегральный функционал обычно оценивается при помощи метода интегрирования Монте-Карло (см. [1]). Функционал (3) содержит четыре повторных интеграла, которые при оценке методом Монте-Карло потребуют значительных вычислительных ресурсов, поэтому желательно построить форму вариационного функционала (3), не содержащую повторных интегралов.

3. Преобразование вариационного функционала. Область краевой задачи Ω можно рассматривать как объединение четырех подобластей $\Omega_1, \Omega_2, \Omega_3, \Omega_4$ (см. рис. 2):

$$\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4.$$

Перейдем в криволинейных интегралах по $\gamma_2, \gamma_3, \gamma_4$ к переменным ξ, η и интегрированию по подобластям $\Omega_1, \Omega_2, \Omega_3, \Omega_4$ с учетом того, что дифференциал длины дуги на $\partial\Omega$ равен $ds = \sqrt{2}d\zeta$ или $ds = \sqrt{2}d\eta$ в зависимости от параметризации кривой.

При вычислении криволинейного интеграла вдоль γ_2 используем параметризацию по переменной η , а именно, $\xi = 2\pi - \eta$, $0 \leq \eta \leq \pi$, $ds = \sqrt{2}d\eta$; тогда

$$\begin{aligned}
\int_{\gamma_2} \varphi_2(\eta) n_1 \int_{\xi}^{\gamma_1} u_{\eta}(\zeta, \eta) d\zeta ds &= \int_0^{\pi} \sqrt{2} \frac{1}{\sqrt{2}} \varphi_2(\eta) \int_{2\pi-\eta}^{\gamma_1} u_{\eta}(\zeta, \eta) d\zeta d\eta = \\
&= \int_0^{\pi} \varphi_2(\eta) \int_{2\pi-\eta}^{\eta} u_{\eta}(\xi, \eta) d\xi d\eta = - \int_0^{\pi} \int_{\eta}^{2\pi-\eta} \varphi_2(\eta) u_{\eta}(\xi, \eta) d\xi d\eta = - \int_{\Omega_1 \cup \Omega_2} \varphi_2(\eta) u_{\eta}(\xi, \eta) d\xi d\eta.
\end{aligned}$$

При вычислении первого криволинейного интеграла вдоль γ_3 также используем параметризацию по переменной η , а именно, $\xi = \eta + 2\pi$, $-\pi \leq \eta \leq 0$, $ds = \sqrt{2}d\eta$; тогда

$$\begin{aligned}
\int_{\gamma_3} \varphi_3 n_1 \sin \xi \int_{\xi}^{\gamma_4} u_{\eta}(\zeta, \eta) d\zeta ds &= \int_{-\pi}^0 \sqrt{2} \varphi_3(\eta) \frac{1}{\sqrt{2}} \int_{\eta+2\pi}^{\gamma_4} u_{\eta}(\zeta, \eta) d\zeta d\eta = \\
&= \int_{-\pi}^0 \varphi_3(\eta) \int_{\eta+2\pi}^{-\eta} u_{\eta}(\xi, \eta) d\xi d\eta = - \int_{-\pi}^0 \int_{-\eta}^{\eta+2\pi} \varphi_3(\eta) u_{\eta}(\xi, \eta) d\xi d\eta = - \int_{\Omega_3 \cup \Omega_4} \varphi_3(\eta) u_{\eta}(\xi, \eta) d\xi d\eta.
\end{aligned}$$

При вычислении второго криволинейного интеграла вдоль γ_3 используем параметризацию по другой переменной ξ , а именно, $\eta = \xi - 2\pi$, $\pi \leq \xi \leq 2\pi$, $ds = \sqrt{2}d\xi$; тогда

$$\begin{aligned}
\int_{\gamma_3} \psi_3 n_2 \int_{\eta}^{\gamma_2} u_{\xi}(\xi, \tau) d\tau ds &= \int_{\pi}^{2\pi} \sqrt{2} \psi_3(\xi) \frac{-1}{\sqrt{2}} \int_{\xi-2\pi}^{\gamma_2} u_{\xi}(\xi, \tau) d\tau d\xi = \\
&= - \int_{\pi}^{2\pi} \psi_3(\xi) \int_{\xi-2\pi}^{2\pi-\xi} u_{\xi}(\xi, \eta) d\eta d\xi = - \int_{\pi}^{2\pi} \int_{\xi-2\pi}^{2\pi-\xi} \psi_3(\xi) u_{\xi}(\xi, \eta) d\eta d\xi = - \int_{\Omega_2 \cup \Omega_3} \psi_3(\xi) u_{\xi}(\xi, \eta) d\xi d\eta.
\end{aligned}$$

При вычислении криволинейного интеграла вдоль γ_4 также используем параметризацию по переменной ξ , а именно, $\eta = -\xi$, $0 \leq \xi \leq \pi$, $ds = \sqrt{2}d\xi$; тогда

$$\begin{aligned}
\int_{\gamma_4} \psi_4 n_2 \int_{\eta}^{\gamma_1} u_{\xi}(\xi, \tau) d\tau ds &= \int_0^{\pi} \sqrt{2} \psi_4(\xi) \frac{-1}{\sqrt{2}} \int_{-\xi}^{\gamma_1} u_{\xi}(\xi, \tau) d\tau d\xi = \\
&= - \int_0^{\pi} \psi_4(\xi) \int_{-\xi}^{\xi} u_{\xi}(\xi, \eta) d\eta d\xi = - \int_0^{\pi} \int_{-\xi}^{\xi} \psi_4(\xi) u_{\xi}(\xi, \eta) d\eta d\xi = - \int_{\Omega_1 \cup \Omega_4} \psi_4(\xi) u_{\xi}(\xi, \eta) d\xi d\eta.
\end{aligned}$$

Итак, вычисление повторных интегралов в (3) дает следующие равенства:

$$\begin{aligned}
\int_{\gamma_2} \varphi_2(\eta) n_1 \int_{\xi}^{\gamma_1} u_{\eta}(\zeta, \eta) d\zeta ds &= - \int_{\Omega_1 \cup \Omega_2} \varphi_2(\eta) u_{\eta}(\xi, \eta) d\xi d\eta, \\
\int_{\gamma_3} \varphi_3 n_1 \sin \xi \int_{\xi}^{\gamma_4} u_{\eta}(\zeta, \eta) d\zeta ds &= - \int_{\Omega_3 \cup \Omega_4} \varphi_3(\eta) u_{\eta}(\xi, \eta) d\xi d\eta, \\
\int_{\gamma_3} \psi_3 n_2 \int_{\eta}^{\gamma_2} u_{\xi}(\xi, \tau) d\tau ds &= - \int_{\Omega_2 \cup \Omega_3} \psi_3(\xi) u_{\xi}(\xi, \eta) d\xi d\eta, \\
\int_{\gamma_4} \psi_4 n_2 \int_{\eta}^{\gamma_1} u_{\xi}(\xi, \tau) d\tau ds &= - \int_{\Omega_1 \cup \Omega_4} \psi_4(\xi) u_{\xi}(\xi, \eta) d\xi d\eta.
\end{aligned}$$

С учетом вычисленных повторных интегралов вариационный функционал (3) для краевой задачи (1)–(2) принимает вид

$$D[u] = \int_{\Omega} (u_{\xi}^2 + u_{\eta}^2) d\xi d\eta + \int_{\gamma_1} u^2 ds - 2 \int_{\gamma_1} \chi_1 u ds - 2 \int_{\Omega_1 \cup \Omega_2} \varphi_2(\eta) u_{\eta}(\xi, \eta) d\xi d\eta - \\ - 2 \int_{\Omega_3 \cup \Omega_4} \varphi_3(\eta) u_{\eta}(\xi, \eta) d\xi d\eta - 2 \int_{\Omega_2 \cup \Omega_3} \psi_3(\xi) u_{\xi}(\xi, \eta) d\xi d\eta - 2 \int_{\Omega_1 \cup \Omega_4} \psi_4(\xi) u_{\xi}(\xi, \eta) d\xi d\eta.$$

Для того, чтобы сгруппировать интегралы в полученном выражении, определим вспомогательные функции Φ , Ψ следующими формулами:

$$\Phi(\eta) = \begin{cases} \varphi_2(\eta), & 0 \leq \eta \leq \pi, \\ \varphi_3(\eta), & -\pi \leq \eta < 0, \end{cases}, \quad \Psi(\xi) = \begin{cases} \psi_3(\xi), & \pi \leq \xi \leq 2\pi, \\ \psi_4(\xi), & 0 \leq \xi < \pi, \end{cases}$$

и получим окончательную форму вариационного функционала для рассматриваемой краевой задачи.

Теорема 2. *Вариационный функционал для краевой задачи (1)–(2) может быть записан в следующем компактном виде*

$$D[u] = \int_{\Omega} (u_{\xi}^2 + u_{\eta}^2 - 2\Phi u_{\eta} - 2\Psi u_{\xi}) d\xi d\eta + \int_{\gamma_1} (u^2 - 2\chi_1 u) ds. \quad (4)$$

4. Обучение нейронной сети для гиперболического уравнения. Решение $u(\xi, \eta)$ краевой задачи (1)–(2) может быть аппроксимировано искусственной нейронной сетью с выходом $f(\xi, \eta; \theta)$, где ξ, η — входные значения нейронной сети, а θ — вектор параметров (весов и смещений) нейронной сети:

$$u(\xi, \eta) \approx f(\xi, \eta; \theta).$$

При обучении нейронной сети в качестве функционала потерь (ошибки) может быть использован функционал невязки (residual functional)

$$\mathcal{L}_R(f) = \int_{\Omega} f_{\xi\eta}^2 d\xi d\eta + \int_{\gamma_1} (f - \chi_1)^2 ds + \int_{\gamma_2} (f_{\eta} - \varphi_2)^2 ds + \\ + \int_{\gamma_3} ((f_{\eta} - \varphi_3)^2 + (f_{\xi} - \psi_3)^2) ds + \int_{\gamma_4} (f_{\xi} - \psi_4)^2 ds. \quad (5)$$

При использовании функционала (5) приходится вычислять пять интегралов, входящих в $\mathcal{L}_R(f)$, для чего необходимо строить случайные выборки из пяти различных областей: Ω , γ_1 , γ_2 , γ_3 , γ_4 , и вычислять частные производные неизвестной функции f до второго порядка включительно.

Также при обучении нейронной сети в качестве функционала потерь (ошибки) может быть использован функционал на основе построенного ранее квазиклассического функционала (4)

$$\mathcal{L}_Q(f) = \int_{\Omega} (f_{\xi}^2 + f_{\eta}^2 - 2\Phi f_{\eta} - 2\Psi f_{\xi}) d\xi d\eta + \int_{\gamma_1} (f^2 - 2\chi_1 f) ds. \quad (6)$$

При использовании квазиклассического функционала (6) приходится вычислять два интеграла, входящих в $\mathcal{L}_Q(f)$, и, соответственно, строить две случайных выборки: из области Ω и из фрагмента границы γ_1 , и вычислять частные производные неизвестной функции f первого порядка.

Поэтому при использовании квазиклассического функционала (6) возникает экономия требуемых вычислительных ресурсов по сравнению с использованием функционала невязки (5) за счет использования меньшего числа случайных выборок и вычисления частных производных более низкого порядка.

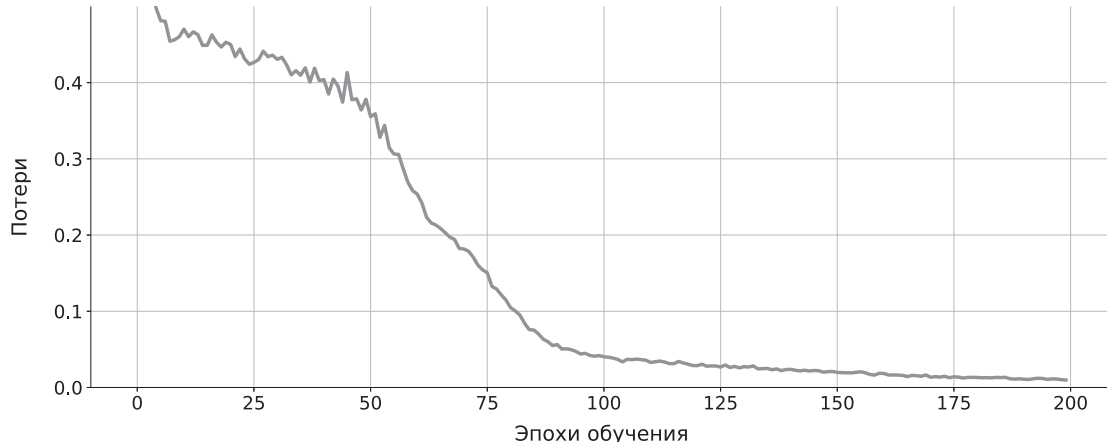


Рис. 3. Кривая обучения нейронной сети при помощи функционала невязки

Проведем вычислительный эксперимент по обучению нейронных сетей для аппроксимации решения гиперболического уравнения (1) с граничными условиями (2) для случая, когда

$$\chi_1(\xi) \equiv 0, \quad \varphi_2(\eta) = -\frac{1}{2} \sin 2\eta, \quad \varphi_3(\eta) = -\frac{1}{2} \sin 2\eta, \quad \psi_3(\xi) = \frac{1}{2} \sin 2\xi, \quad \psi_4(\xi) = \frac{1}{2} \sin 2\xi,$$

и точное решение краевой задачи (1)–(2) равно

$$u(\xi, \eta) = -\frac{1}{4} \cos 2\xi + \frac{1}{4} \cos 2\eta.$$

В качестве функционалов потерь в вычислительном эксперименте будем использовать функционал невязки (5) и квазиклассический функционал (6).

Вычислительный эксперимент проводится для указанной ниже архитектуры нейронной сети при следующих значениях гиперпараметров:

- используется нейронная сеть прямого распространения (FF) с плотными (Dense) слоями;
- число скрытых слоев равно 4 с функцией активации $\tanh$ (гиперболический тангенс);
- количество нейронов в скрытых слоях равно 100;
- в выходном слое один нейрон без функции активации;
- производится 200 эпох обучения с 10 шагами стохастического градиентного спуска (SGD) для каждой эпохи;
- используется оптимизатор Adam с начальной скоростью обучения 0,0001;
- при сэмплировании выбирается по 1000 точек из области Ω и по 500 точек для каждого участка границы $\partial\Omega$.

Программный код обучения нейронной сети для краевой задачи (1)–(2) реализован с использованием фреймворка TensorFlow и выполняется на компьютере MacBook Pro с процессором M2Max.

При обучении нейронной сети на функционале невязки (5) в качестве функционала потерь кривая обучения показана на рис. 3, абсолютная ошибка полученной нейросетевой аппроксимации визуализирована на рис. 4. Процесс обучения продолжался 119,5 с, в результате построена модель нейронной сети со следующими показателями качества (среднеквадратической ошибки MSE , средней абсолютной ошибки MAE , коэффициента детерминации R^2):

$$MSE = 0,00966, \quad MAE = 0,07097, \quad R^2 = 83,92\%.$$

При обучении нейронной сети на квазиклассическом функционале (6) в качестве функционала потерь кривая обучения показана на рис. 5, абсолютная ошибка полученной нейросетевой аппроксимации визуализирована на рис. 6. Процесс обучения продолжался 41,6 с, в результате построена модель нейронной сети со следующими показателями качества:

$$MSE = 0,00001, \quad MAE = 0,00215, \quad R^2 = 99,97\%.$$

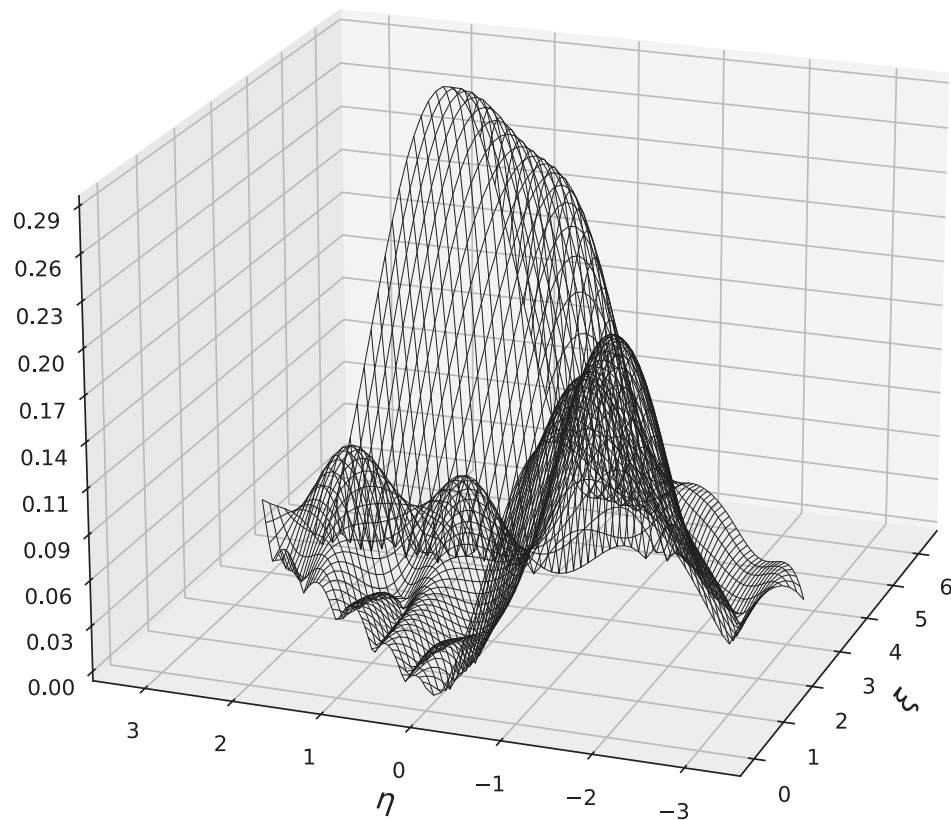


Рис. 4. Поверхность ошибки нейронной сети, обученной с функционалом невязки

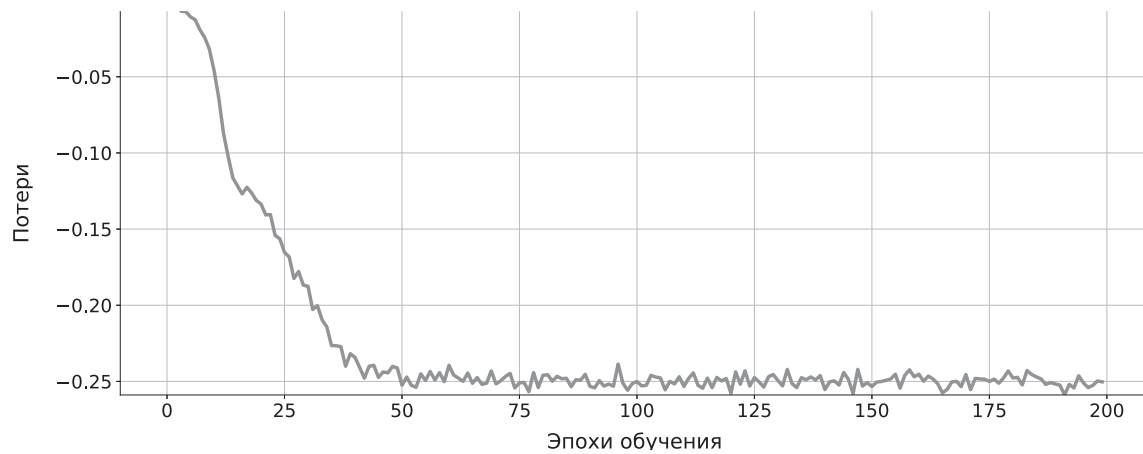


Рис. 5. Кривая обучения нейронной сети при помощи квазиклассического функционала

Таким образом, при использовании квазиклассического функционала (6) процесс обучения нейронной сети ускоряется примерно в три раза, при этом достигаются существенно более высокие значения показателей качества нейросетевой модели, в частности, более низкие значения абсолютной ошибки аппроксимации.

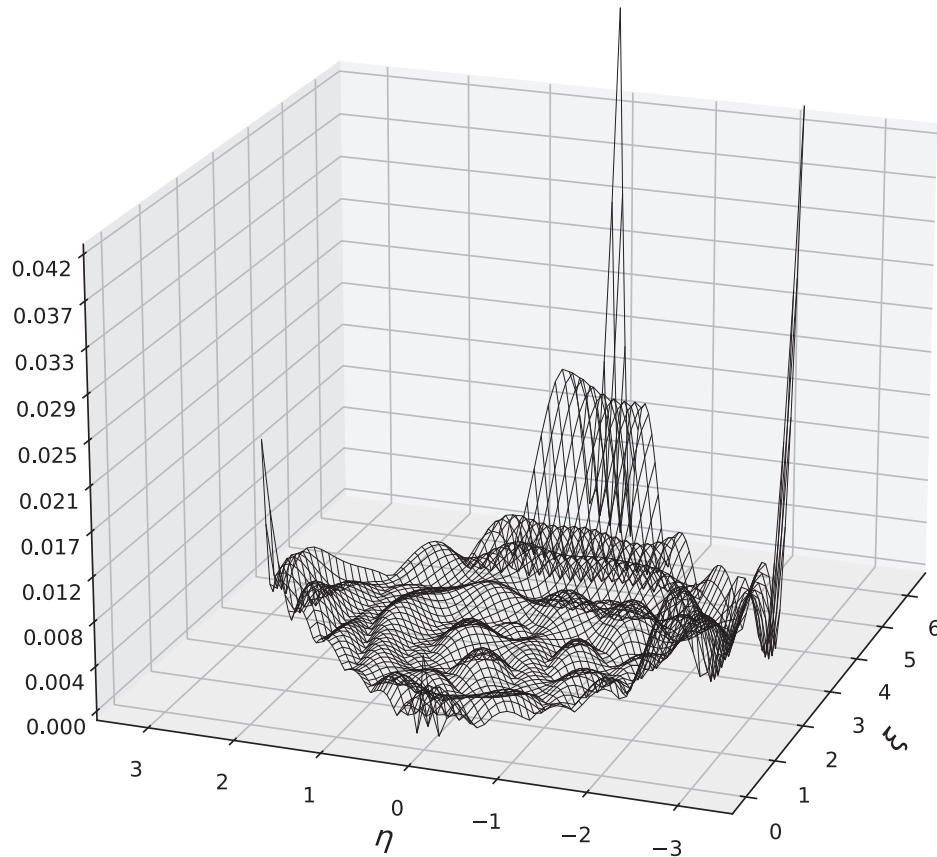


Рис. 6. Поверхность ошибки нейронной сети, обученной с квазиклассическим функционалом

5. Заключение. При помощи метода симметризирующего оператора В. М. Шалова получена компактная вариационная формулировка краевой задачи для гиперболического уравнения, содержащая всего два интеграла по области краевой задачи и фрагменту границы области, с подынтегральными выражениями, зависящими от неизвестной функции и ее первых производных. Полученный вариационный функционал при его использовании в качестве функционала потерь нейронной сети имеет ряд преимуществ перед функционалом потерь на основе невязки, в том числе более высокую скорость обучения нейронной сети и более высокое качество аппроксимации решения краевой задачи.

СПИСОК ЛИТЕРАТУРЫ

1. *Соболь И. М.* Численные методы Монте-Карло. — М.: Наука, 1973.
2. *Филиппов В. М.* Вариационный метод решения краевых задач для волнового уравнения// Диффер. уравн. — 1984. — 20, № 11. — С. 1961–1968.
3. *Филиппов В. М., Савчин В. М., Шорохов С. Г.* Вариационные принципы для непотенциальных операторов// Итоги науки техн. Совр. пробл. мат. Нов. достиж. — 1992. — 40. — С. 3–176.
4. *Шалов В. М.* Некоторое обобщение пространства К. Фридрихса// Докл. АН СССР. — 1963. — 151, № 2. — С. 292–294.
5. *Шалов В. М.* Решение несамосопряженных уравнений вариационным методом// Докл. АН СССР. — 1963. — 151, № 3. — С. 511–512.
6. *Шалов В. М.* Принцип минимума квадратичного функционала для гиперболического уравнения// Диффер. уравн. — 1965. — 1, № 10. — С. 1338–1365.
7. *E W., Yu B.* The deep Ritz method: A deep learning-based numerical algorithm for solving variational problems// Commun. Math. Stat. — 2018. — 6, № 1. — P. 1–12.

8. *Geneva N., Zabaras N.* Modeling the dynamics of PDE systems with physics-constrained deep autoregressive networks// J. Comput. Phys. — 2020. — 403. — 109056.
9. *Raissi M., Perdikaris P., Karniadakis G.* Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations// J. Comput. Phys. — 2019. — 378. — P. 686–707.
10. *Sirignano J., Spiliopoulos K.* DGM: A deep learning algorithm for solving partial differential equations// J. Comput. Phys. — 2018. — 375. — P. 1339–1364.
11. *Young W. H.* On multiple integration by parts and the second theorem of the mean// Proc. London Math. Soc. — 1917. — 2, № 16. — P. 273–293.
12. *Zhu Y., Zabaras N., Koutsourelakis P.-S., Perdikaris P.* Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data// Journal of Computational Physics — 2019. — 394. — P. 56–81.

ДЕКЛАРАЦИЯ АВТОРА

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Финансирование. Работа выполнена в рамках проекта № 002092-0-000 Российского университета дружбы народов имени Патриса Лумумбы.

Финансовые интересы. Автор заявляет об отсутствии подлежащих раскрытию финансовых или нефинансовых интересов, связанных с публикуемым материалом.

Шорохов Сергей Геннадьевич (Shorokhov Sergey Gennadievich)

Российский университет дружбы народов имени Патриса Лумумбы, Москва

(Peoples' Friendship University of Russia named after Patrice Lumumba, Moscow, Russia)

E-mail: shorokhov-sg@rudn.ru