

СЕКЦИЯ 5

НЕЛИНЕЙНЫЕ ПРОЦЕССЫ И ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ В ЗАДАЧАХ УПРАВЛЕНИЯ КРУПНОМАСШТАБНЫМИ СИСТЕМАМИ

ИСПОЛЬЗОВАНИЕ ПРОГНОЗА ФИЗИЧЕСКИ ИНФОРМИРОВАННОЙ НЕЙРОННОЙ СЕТИ ДЛЯ РЕШЕНИЯ ОБРАТНОЙ КОЭФФИЦИЕНТНОЙ ЗАДАЧИ

Арсеньев-Образцов С.С., Соколов А.А.

РГУ нефти и газа (НИУ) имени И.М. Губкина, Москва, Россия
arseniev@gubkin.ru, mail.sokolov.aa@yandex.ru

Аннотация. Анализируется эффективность использования прогноза физически информированной нейронной сети (PINN) в качестве начального приближения для итерационного метода решения обратной коэффициентной задачи на примере волнового уравнения. Коэффициент ищется в параметрической форме. Дискретизация осуществляется методом конечных разностей (МКР), оптимизация выполняется различными градиентными методами.

Ключевые слова: обратная задача, параметризация решения, оптимизация, волновое уравнение, метод конечных разностей, PINN (Physics-Informed Neural Networks), BFGS (Broyden, Fletcher, Goldfarb, Shanno), TNC (Truncated Newton Constrained).

Введение

Для многих процессов, происходящих в природе, удается построить математические модели, задаваемые с помощью различного рода систем уравнений. При этом, коэффициенты и свободные члены уравнений выражают некоторые важные свойства среды или характеристики изучаемого процесса. Обычно, для конкретных прикладных они неизвестны и зачастую не могут быть измерены непосредственно, например, динамические петрофизические параметры коллекторов углеводородного сырья. Однако их можно определить в результате решения так называемых обратных коэффициентных задач. Данные задачи, как правило, являются некорректными по Адамару и решаются численно с помощью различных методов оптимизации с учетом дополнительной информации.

Существует несколько типов обратных задач, но мы остановимся только на коэффициентных. Общепринятый подход [1] состоит в параметризации искомым полем коэффициентов и реализации некоторой итерационной процедуры, в ходе которой выполняется уточнения вектора параметров путем минимизации заданного целевого функционала, отражающего расхождение между решением прямой задачи в определенных областях относительно дополнительно измеренных значений. Например, сейсмограммам, полученными в результате проведенной сейсмо-съемки месторождения. Многомерная поверхность такого критерия качества может иметь большое количество локальных и, возможно, глобальных минимумов, не имеющих физического смысла. Повышение точности приближения полем коэффициентов осуществляется за счёт увеличения количества оптимизируемых параметров, что приводит к росту размерности пространства. При исследовании процессов, описываемых уравнениями в частных производных, проверка точек в пространстве параметров на экстремальность обычно осложняется высокой стоимостью вычисления значений целевой функции и её градиента.

Помимо используемого метода оптимизации на успешность применения итерационной процедуры критическое влияние оказывает начальное приближение. При его удачном выборе многие описанные трудности в существенной мере удается преодолеть. При определении начального приближения, как правило привлекается дополнительная информация или специальные знания, полученные из соответствующей предметной области. Начальные приближения можно выбрать случайно или предварительно решить вспомогательную задачу с существенно меньшим числом параметров. Иногда удаётся использовать упрощенную постановку прямой задачи, т.е. прокси-модель процесса без учёта некоторых факторов.

В глубоком обучении для борьбы с «затуханием» градиентов используется случайная инициализация весов. Также применяется метод переноса обучения, когда веса предварительно

обученной нейросети используют в качестве начального приближения для обучения модели иной архитектуры или с другими наборами данных.

В данной работе начальное приближение получается с помощью физически информированной нейронной сети [2, 3] на примере обратной коэффициентной задачи для одномерного и двумерного волнового уравнений с неизвестным полем скорости. Акустическое уравнение широко используется в задачах сейсморазведки, например, при решении задачи полного обращения волновых полей [4]. С помощью проведенных вычислительных экспериментов при использовании методов градиентной оптимизации целевой функции была выполнена экспериментальная оценка эффективности предложенного подхода

1. Постановка задачи

1.1. Математическая модель распространения акустических волн

Наиболее простая модель, задается в виде неоднородного волнового уравнения с переменным коэффициентом:

$$u_{tt} = c^2 \Delta u + f. \quad (1)$$

В прямой задаче уравнение (1) необходимо решить относительно функции $u = u(t, \vec{x})$, означающей смещение или давление в зависимости от интерпретации. Уравнение задано в пространственно-временной области $\Omega = \Omega_T \times \Omega_S$. Коэффициент $c = c(\vec{x})$ задает пространственное распределение скорости распространения волны сжатия в среде. Точечный источник волн моделируется с помощью функции $f = f(t, \vec{x}) = p(t)\delta(\vec{x} - \vec{x}_0)$, где в качестве $p(t)$ можно использовать, например, импульс Рикера или Пузырева, а дельта-функцию $\delta(\vec{x} - \vec{x}_0)$ приблизить функцией Гаусса с пиком в точке \vec{x}_0 . На границе области $\partial\Omega = \Omega_T \times \partial\Omega_S$ для простоты зададим граничные условия (ГУ) Дирихле, а в качестве начального условия (НУ) будем использовать $\{u(0, \vec{x}) = u_t(0, \vec{x}) = 0\}$.

1.2. Дискретизация и параметризация коэффициента (поля скорости волны)

На области решения Ω введем регулярную сетку точек $\omega = \omega_T \times \omega_S = \{(t_k, \vec{x}_m) | t_k \in \omega_T, \vec{x}_m \in \omega_S\}$. Под решением прямой задачи будем понимать сеточную функцию v определённую на ω и удовлетворяющую дискретной (конечноразностной) форме исходного уравнения.

В обратной задаче необходимо определить поле скорости волны $c = c(\vec{x})$ с учетом дополнительных измерений, выполненных в заданных точках сетки. Решение будем искать приближенно проекционным методом в параметрическом виде:

$$c = c(\vec{x}) \approx \sum_{i=1}^{N_p} \alpha_i \beta_i(\vec{x}). \quad (2)$$

Здесь $\alpha_i \in \mathbb{R}$ – коэффициенты разложения, $\{\beta_i\}_{i=1}^{N_p}$ – система базисных функций, в качестве которых будут использованы функции, заданные на конечном носителе. При решении прямой задачи должны быть заданы значения функции-источника \vec{f} , начальные и граничные условия (НУ и ГУ) в соответствующих узлах сетки ω . При решении обратной задачи для простоты будем считать, что на каждом k -м временном слое известны истинные смещения \vec{u}_k во всех узлах пространственной сетки. Альтернатива состоит в том, что значения решения известны на некотором подмножестве узлов соответствующего временного слоя. Например, при решении задачи сейсморазведки – это измерения сейсмоприемников, расположенных обычно на дневной поверхности соответствующей области.

2. Метод решения

2.1. Итерационная процедура оптимизации

Классический подход к поиску решения обратной задачи (2) состоит в определении коэффициентов разложения $\vec{\alpha} = (\alpha_1, \dots, \alpha_{N_p})^T$ на основе решения задачи оптимизации функционала J (3) при выполнении условий (6), (7). Соотношение (4) представляет собой норму невязки между вычисленным $\vec{v}_k(\vec{\alpha})$ и истинным волновым полем \vec{u}_k в узлах пространственной сетки k -го временного слоя. Поскольку, обратные задачи практически всегда некорректны по Адамару, обычно применяется некоторая регуляризация $R_\theta(\vec{\alpha})$, заданная, например в форме Тихонова (5).

$$J = J(\vec{v}(\vec{\alpha}), \vec{\alpha}; \theta) = \sum_{k=2}^{N_t} J_k(\vec{v}_k, \vec{\alpha}) + R_\theta(\vec{\alpha}) \rightarrow \min_{\vec{\alpha}}; \quad (3)$$

$$J_k(\vec{v}_k, \vec{\alpha}) = \frac{1}{2} \|\vec{v}_k(\vec{\alpha}) - \vec{u}_k\|_2^2; \quad (4)$$

$$R_\theta(\vec{\alpha}) = \frac{\theta}{2} \|\vec{\alpha}\|_2^2; \quad (5)$$

$$F_k(\vec{v}_k, \vec{v}_{k-1}, \vec{v}_{k-2}, \vec{\alpha}) = 0, \quad k = \overline{2, N_t}; \quad (6)$$

$$\alpha_i^{\min} \leq \alpha_i \leq \alpha_i^{\max}, \quad i = \overline{1, N_p}. \quad (7)$$

Под ограничениями (6) понимается выполнение трехслойной конечно-разностной схемы для волнового уравнения (1) на каждом временном слое. Условия (7) позволяют зафиксировать скорость распространения волны в характерном для исследуемого физического процесса диапазоне, что в свою очередь способствует вычислительной устойчивости и имеет определенный регуляризирующий эффект.

Поставленную обратную задачу можно приближенно решить с помощью различных итерационных методов условной оптимизации. На каждом шаге ν итерационной процедуры необходимо вычислять значения волнового поля $\vec{v}(\vec{\alpha}^{(\nu)})$ в узлах сетки ω . Существенные вычислительные затраты, требуемые для выполнения данной процедуры, делают популярными в подобных задачах быстро сходящиеся градиентные методы оптимизации [4]. Традиционный способ получения искомого градиента $\nabla_{\vec{\alpha}} J$ в точке $\vec{\alpha}^{(\nu)}$ основан на переходе к безусловной оптимизации и функции Лагранжа (8).

$$L = L(\vec{v}(\vec{\alpha}), \vec{\lambda}, \vec{\alpha}; \theta) = \sum_{k=2}^{N_t} (J_k + \vec{\lambda}_k^T F_k) + R_\theta(\vec{\alpha}). \quad (8)$$

Необходимые условия экстремума (8) можно определить исходя из дискретного принципа максимума Понтрягина, приравняв первую вариацию Лагранжиана к нулю и приведя её к виду (9).

$$\delta L = \sum_{k=2}^{N_t} (L_{\vec{v}_k}, \delta \vec{v}_k) + \sum_{k=2}^{N_t} (L_{\vec{\lambda}_k}, \delta \vec{\lambda}_k) + (L_{\vec{\alpha}}, \delta \vec{\alpha}) = 0; \quad (9)$$

$$L_{\vec{\lambda}_k} = F_k = F_k(\vec{v}_k, \vec{v}_{k-1}, \vec{v}_{k-2}, \vec{\alpha}) = 0, \quad k = \overline{2, N_t}; \quad (10)$$

$$L_{\vec{v}_k} = \nabla_{\vec{v}_k} J_k^T + \nabla_{\vec{v}_k} F_k^T \vec{\lambda}_k + \nabla_{\vec{v}_k} F_{k+1}^T \vec{\lambda}_{k+1} + \nabla_{\vec{v}_k} F_{k+2}^T \vec{\lambda}_{k+2} = 0, \quad k = \overline{N_t - 1, 2}; \quad (11)$$

$$L_{\vec{v}_k} = \nabla_{\vec{v}_k} J_k^T + \nabla_{\vec{v}_k} F_k^T \vec{\lambda}_k + \nabla_{\vec{v}_k} F_{k+1}^T \vec{\lambda}_{k+1} = 0, \quad k = N_t; \quad (12)$$

$$\vec{\lambda}_{N_t+1} = 0; \quad (13)$$

$$L_{\vec{\alpha}} = \nabla_{\vec{\alpha}} J_k + (\nabla_{\vec{v}_k} F_k^T \vec{\lambda}_k)^T. \quad (14)$$

Искомый градиент функции L по коэффициентам $\vec{\alpha}$ представлен соотношением (14). Множители Лагранжа или сопряженные переменные $\vec{\lambda}_k$ определяются из условия трансверсальности (13), а также сопряженной задачи (11)-(12). Уравнения (11)-(12) решаются относительно $\vec{\lambda}_k$ в обратном времени от N_t до 2, а роль начальных условий выполняет (13). Выражение (13) справедливо, когда зафиксировано итоговое время решения задачи при свободном конечном состоянии системы (значение волнового поля \vec{v}_{N_t} на последнем шаге по времени N_t). Соотношение (10) представляет собой конечно-разностную схему для прямой задачи, которую удобно записать в матрично-векторном виде (15).

$$F_k = A_1 \vec{v}_k + A_2 \vec{v}_{k-1} + A_3 \vec{v}_{k-2} + \vec{b}_{k-1} = 0, \quad k = \overline{2, N_t}. \quad (15)$$

Матрицы A_1, A_2, A_3 определяются выбором разностной схемы, имеют диагональную структуру и сильно разрежены. Элементы по крайней мере одной из матриц зависят от скорости волны и соответственно коэффициентов разложения $\vec{\alpha}$. Матрицы постоянны для каждого шага по времени, так как рассматривается случай, когда скорость (2), не зависит от времени. Векторы правой части \vec{b}_{k-1} ненулевые, они содержат информацию о граничных условиях и функции-источнике f . Следует отметить, что в допустимой области градиент (14) не всегда можно приравнять к нулю, так как оптимальное значение может достигаться на границе, определяемой ограничениями (7). Условия (7) можно учесть на уровне градиентного метода оптимизации. Например, при «выходе» параметра α_i за границы допустимой области присваивать ему значение границы.

Таким образом, выполнив необходимые операции матрично-векторного дифференцирования можно вычислить искомый градиент $L_{\vec{\alpha}}$. Основной объем необходимых операций приходится на решение прямой и сопряженной задачи и несущественно возрастает при увеличении количества оптимизируемых параметров $\vec{\alpha}$.

При вычислении градиента с помощью численного дифференцирования на каждую компоненту придется решить по дополнительной прямой задаче, что зачастую неприменимо на практике. К тому же, возникает проблема выбора шага приращения из-за некорректности численного дифференцирования. Наиболее перспективной альтернативой может быть применение техники автоматического дифференцирования. Тем не менее в данной работе применялся традиционный подход. Численное дифференцирование использовалось исключительно для проверки работы программного обеспечения.

2.2. Начальное приближение

В работе [2] был представлен метод построения конечно-элементного приближения для обратной коэффициентной задачи с помощью физически информированной нейронной сети. Архитектура обучаемой части преобразования выполнена по аналогии с EfficientNetV2 [5]. Ключевая особенность предложенного алгоритма заключается в предсказании непосредственно разложения $\vec{\alpha}_{net}$ в виде (2) по базису из функций, заданных на конечном носителе.

Подобный подход позволяет провести декомпозицию рассматриваемой области, представленной сеткой ω , на части $\omega_{d_i} = \omega_T \times \omega_{s_i}, \omega_{s_i} \subset \omega_s$, а обработку подобластей ω_{d_i} выполнять независимо друг от друга. Если базис для всей области допускает построение непрерывного решения, то усреднив значения $\vec{\alpha}_{d_i}$ при базисных функциях на границах соседних подобластей можно устранить возникающие разрывы функции. Необходимо отметить, что усреднение параметров в рассмотренных примерах [2] приводило к компенсации ошибок предсказаний и соответственно уменьшению нормы невязки решения обратной задачи.

Нейронная сеть позволяет получить эффективную аппроксимацию решения, но с некоторыми ограничениями по точности сосредоточенными преимущественно вблизи границы области и особенно в угловых точках, в которых не проводится усреднение коэффициентов и используется лишь одно предсказание. К тому же, поскольку модели нейронных сетей используют преимущественно вычислительно эффективные преобразования, что в совокупности с высокопроизводительной параллельной реализацией современных библиотек глубокого обучения, позволяет получать прогноз за пренебрежимо малое время. Таким образом, процедура предсказания вектора коэффициентов $\vec{\alpha}_{net}$ может быть использована для получения начального приближения $\vec{\alpha}^{(0)} = \vec{\alpha}_{net}$ для последующей итерационной процедуры оптимизации (решения обратной задачи).

3. Вычислительный эксперимент

3.1. Задание области решения и дискретизации

С целью анализа вычислительной эффективности предложенной процедуры была проведена серия экспериментов решения обратных коэффициентных задач в одномерном и двумерном случае с областью решения Ω_{1D} и Ω_{2D} соответственно (16)-(17). Прямая задача решалась с помощью явной конечно-разностной схемы на сетках $\Omega_{1D} \sim \omega_{1D}$ и $\Omega_{2D} \sim \omega_{2D}$ (18)-(20). Уравнение (1) было обезразмерено.

$$\Omega_T = \{0 \leq t \leq 1\}, \Omega_x = \{-1 \leq x \leq 1\}, \Omega_y = \{-1 \leq y \leq 1\}; \quad (16)$$

$$\Omega_{1D} = \Omega_T \times \Omega_x, \Omega_{2D} = \Omega_T \times \Omega_x \times \Omega_y; \quad (17)$$

$$\omega_T = \{t_k = kh_t \mid k = 0..2000, h_t = \frac{1}{2000}\}; \quad (18)$$

$$\omega_x = \{x_i = ih_x \mid i = 0..100, h_x = \frac{2}{100}\}, \omega_y = \{y_j = jh_y \mid j = 0..100, h_y = \frac{2}{100}\}; \quad (19)$$

$$\omega_{1D} = \omega_T \times \omega_x, \omega_{2D} = \omega_T \times \omega_x \times \omega_y. \quad (20)$$

3.2. Подготовка и параметризация

В первой серии экспериментов начальное приближение $\vec{\alpha}^{(0)}$ было получено с помощью нейронной сети [2]. Сеть в обоих случаях обучалась в течение 100 эпох на синтетическом наборе данных, сформированном из 1024 решений прямой задачи с различными положениями источника и

полями скорости. В одномерном случае каждое исходное решение прямой задачи было разделено на 4 равные части по пространству, таким образом был сформирован набор из 4096 образцов. В двумерном случае пространственная область разбивалась на 16 равных частей, что в результате позволило сформировать набор данных из 16384 образцов. В обоих случаях в образцах использовался лишь каждый десятый временной слой. В процессе обучения нейросети минимизировалось среднее от квадратов невязки истинного коэффициента и кусочно-постоянного приближения в узлах сетки соответствующей подобласти. На графическом ускорителе Nvidia P-100 обучение потребовало ~16.5 минут в одномерном случае и ~1 час 10 минут в двумерном случае.

В одномерной задаче исходное поле скорости $c(x)$ было задано линейной функцией, а для приближения в каждой подобласти использовалось по 4 кусочно-постоянных базисных функции. Всего после объединения подобластей получаем вектор $\vec{\alpha}_{net}^{1D}$ из 16 параметров. В двумерной задаче в качестве истинной скорости c брался прямоугольный элемент с 4 базисными функциями, что эквивалентно функции вида $c = c(x, y) = c_0 + c_1x + c_2y + c_3xy$. Пример одной базисной функции $\beta_{01}(x, y)$ представлен на рисунке 1. В каждой подобласти также использовались прямоугольные конечные элементы, определяемые 4 параметрами для каждой из 16 подобластей, что в итоге дает 64 параметра. После объединения подобластей, чтобы получить непрерывную скорость волны, коэффициенты при смежных базисных функциях усредняются, что позволяет сократить количество оптимизируемых параметров до 25 компонент вектора $\vec{\alpha}_{net}^{2D}$.

Во второй серии экспериментов в качестве начального приближения использовались векторы $\vec{\alpha}_{const}^{1D}$, $\vec{\alpha}_{const}^{2D}$ длины 16 и 25 соответственно. Их компоненты были равны некоторой константе, которая для каждой задачи генерировалась случайно из допустимой области. Во всех случаях значения истинного коэффициента скорости не выходили за пределы допустимой области.

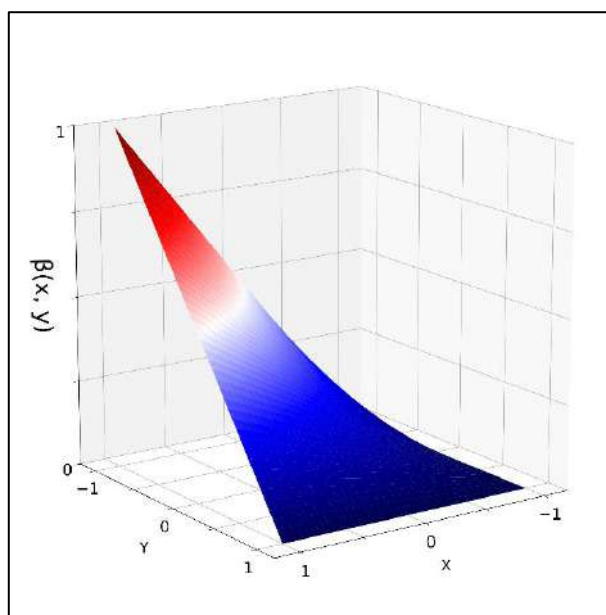


Рис. 1. Пример базисной функции, используемой в двумерной задаче

3.3. Оптимизация и результаты

После получения начального приближения выполнялась итерационная процедура оптимизация двумя градиентными методами: квазиньютоновским методом L-BFGS-B и усеченным методом Ньютона TNC [6]. Реализации обоих методов были взяты из библиотеки SciPy (v1.15.1) [7] с ограничениями вида (7), которые определяли допустимую область каждого из оптимизируемых параметров до отрезка [0, 10].

На каждой итерации подсчитывалось взвешенное количество вызовов целевой функции и вычислений градиента с коэффициентами 1 и 2 соответственно. Данный показатель выражает общее количество решений прямой и сопряженной задач в ходе оптимизации и позволяет грубо оценить вычислительные затраты используемого метода. Проверка программной реализации вычисления градиента через сопряженную задачу выполнялась путем сравнения с конечно-разностным вычислением градиента на тестовых обратных задачах. Качество решения задачи минимизации оценивалось по значению целевой функции, которая является неотрицательной.

Эксперимент был направлен на оценку «выигрыша» от использования начального приближения нейросетью относительно случайной инициализации. Под «выигрышем» понимается взвешенное количество вызовов функции решения прямой и сопряженной задач до достижения порогового значения целевой функции равного 10^{-7} . Порог был выбран эмпирически так, чтобы при начальном приближении нейросетью в среднем выполнять небольшое количество итераций, а полученное решение обратной задачи отражало бы исходную зависимость с достаточной точностью. Помимо прочего, был установлен лимит на количество итераций в 500 и 50 для одномерных и двумерных задач соответственно. Стоит отметить, что за одну итерацию методы могут совершать несколько вычислений значения функции и градиента. Классический критерий остановки метода оптимизации, заключающийся в малости изменений целевой функции или градиента, в SciPy регулируется параметром «tol», который был выбран равным 10^{-16} . Регуляризация не применялась.

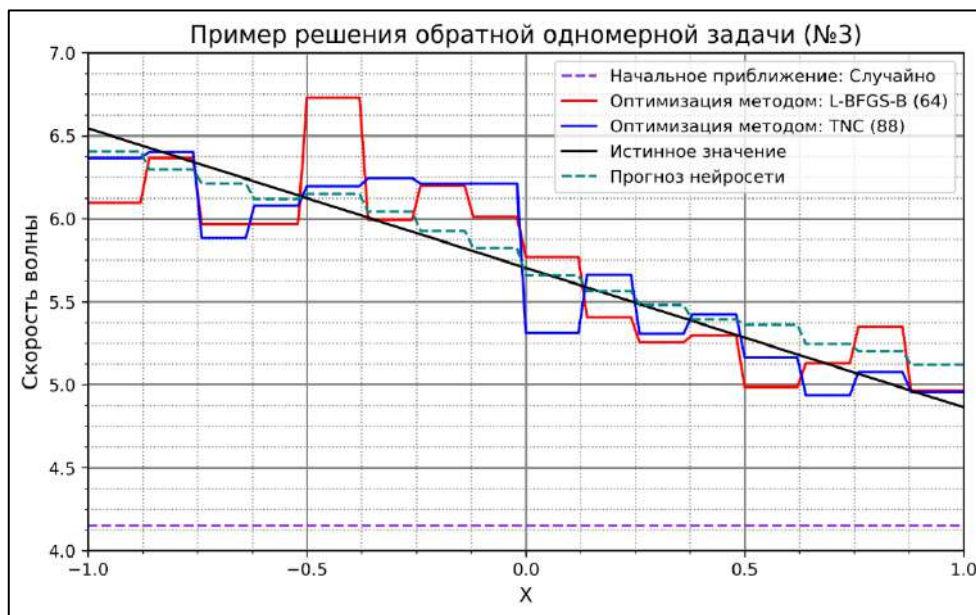


Рис. 2. Пример решения обратной задачи различными методами оптимизации из одинакового начального приближения. В скобках рядом с методом указано взвешенное количество вызовов функции и градиента

В таблице 1 представлены результаты эксперимента для одномерной задачи. Средние значения целевой функции при инициализации нейросетью оказались меньше порога и лишь в немногих случаях алгоритм был прерван из-за достижения малых значений параметра «tol». Метод TNC оказался немного эффективней: за меньшее количество обращений удается достичь заданного порога с меньшим средним значением. При инициализации случайным значением можно выделить два случая. Первый проиллюстрирован рисунком 2 и его можно назвать «успешным». Пороговое значение достигнуто, выполнено несколько десятков итераций, а график поля скорости в целом отражает реальную зависимость. При втором варианте проявляются, как недостатки использованных методов оптимизации, так и некорректность обратных задач, особенно без использования регуляризации. В этом случае порогового значения целевой функции достичь не удастся, в основном из-за остановки по классическому критерию сходимости, что как правило сопровождается «упором» в границу допустимой области одного или нескольких оптимизируемых параметров.

Таблица 1. Результаты эксперимента для 128 одномерных обратных задач

Метод оптимизации	Способ инициализации	Среднее взвешенное количество вызовов	Среднее значение целевой функции после оптимизации
L-BFGS-B	Нейросеть	61,2	$8,74 \cdot 10^{-8}$
L-BFGS-B	Случайно	245,0	$7,82 \cdot 10^{-6}$
TNC	Нейросеть	50,5	$7,14 \cdot 10^{-8}$
TNC	Случайно	424,5	$7,54 \cdot 10^{-6}$

Таблица 2. Результаты эксперимента для 32 двумерных обратных задач

Метод оптимизации	Способ инициализации	Среднее взвешенное количество вызовов	Среднее значение целевой функции после оптимизации
L-BFGS-B	Нейросеть	2,7	$3,41 \cdot 10^{-8}$
L-BFGS-B	Случайно	159,6	$4,14 \cdot 10^{-7}$
TNC	Нейросеть	5,9	$3,13 \cdot 10^{-8}$
TNC	Случайно	273,0	$4,22 \cdot 10^{-7}$

В двумерном случае усредненные показатели экспериментов представлены таблице 2. Нетрудно заметить, что во всех случаях удаётся достичь лучших результатов за существенно меньшее число итераций при большем количестве оптимизируемых параметров. Причина в том, что используемый в двумерном случае базис позволяет точно приблизить решение обратной задачи.

Результирующие показатели по каждой задаче представлены на рисунках 3 и 4. При случайной инициализации (красный график) на рисунке 3, аналогично одномерному случаю, можно наблюдать две группы экспериментов. В первой менее чем за ~50 итераций удается достичь заданного порога, а во второй оптимизация выполняется существенно дольше и прерывается по причине выполнения одного из критериев остановки. Нетрудно заметить, что подобное поведение наблюдается в задачах с одним и тем же номером независимо от метода оптимизации. При инициализации нейронной сетью количество выполненных итераций на графиках (a) и (b) в целом не превышает 50.

На рисунке 4 приведены значения целевой функции в начальный момент времени (выколота точка соответствующего цвета) и после оптимизации (закрашенная точка) в логарифмическом масштабе. При инициализации нейронной сетью заданный порог (черная пунктирная линия) достигается во всех рассмотренных примерах.

В случае инициализации методом предсказания нейронной сети при дальнейшей оптимизации после достижения заданного порога 10^{-7} значение целевой функции продолжает быстро убывать, что непременно сказывается на точности получаемого поля скоростей. На рисунке 5 представлены абсолютная и относительная погрешность поля скоростей в начальный момент времени при использовании начального приближения, полученного нейросетью в одной из задач. Результат оптимизации по прошествии 47 итераций метода BFGS представлен на рисунке 6. Процедура оптимизации была остановлена по достижении целевой функцией порога 10^{-10} . Наибольшие погрешности возникают вблизи границы и по мере оптимизации постепенно устраняются во всей области решения. Относительная погрешность во всех узлах сетки не превосходит 3%

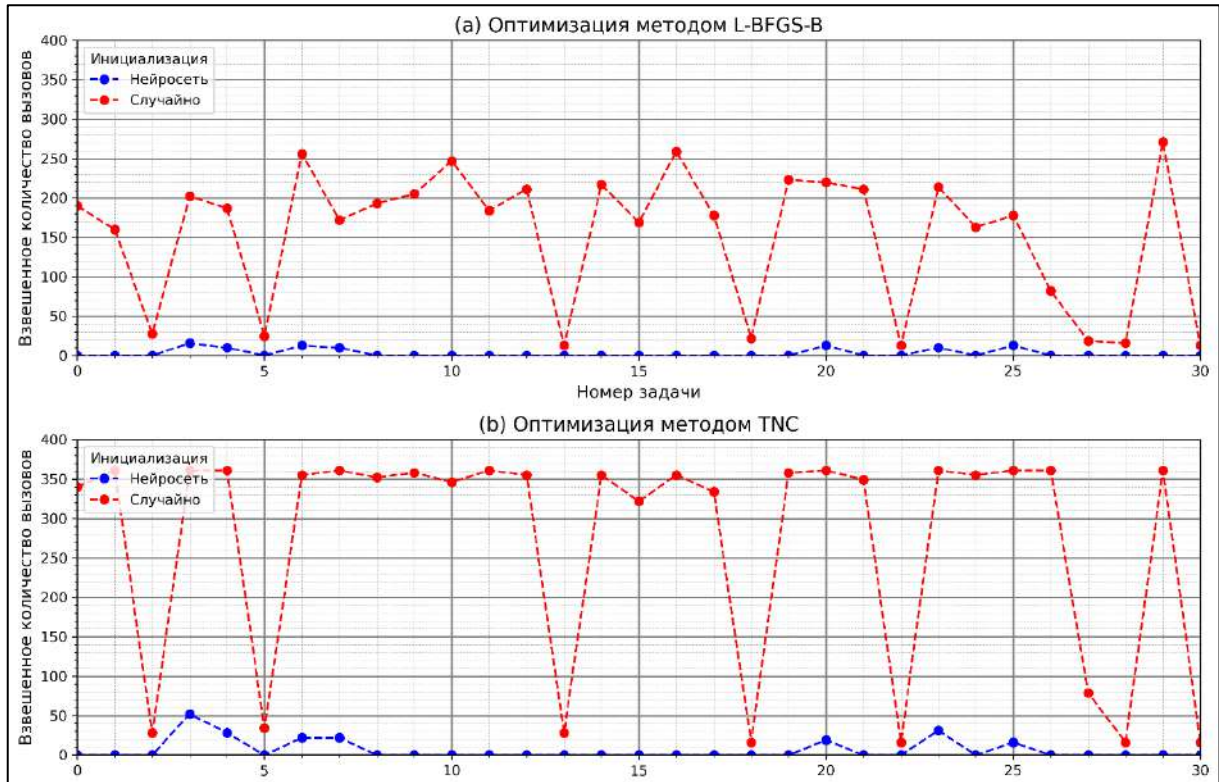


Рис. 3. Взвешенное количество вызовов функции и градиента при решении двумерных обратных задач выполненное до достижения условий останова. (a) Оптимизация выполнена методом L-BFGS-B. (b) Оптимизация выполнена методом TNC

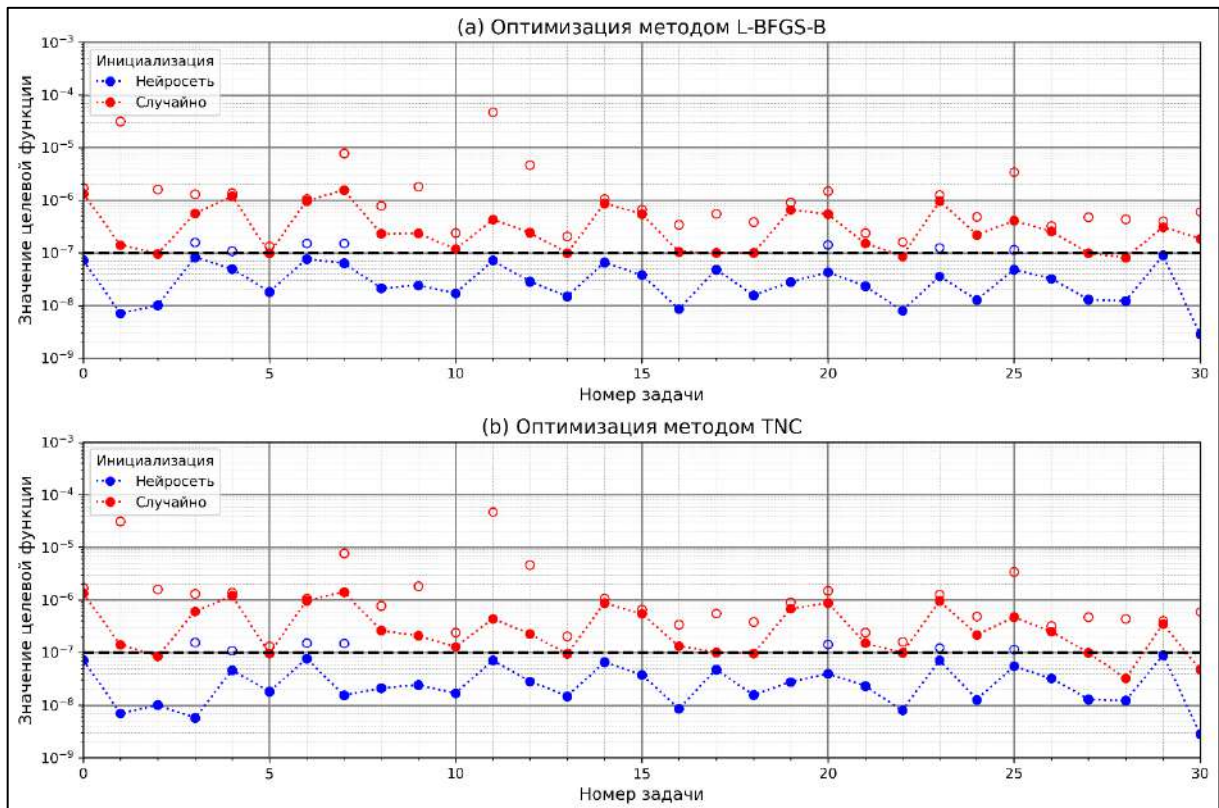


Рис. 4. Значение целевой функции градиента при решении двумерных обратных задач. Выколотые точки соответствуют нулевой итерации, закрасенные итоговой. Черной пунктирной линией обозначено пороговое значение. (a) Оптимизация выполнена методом L-BFGS-B. (b) Оптимизация выполнена методом TNC

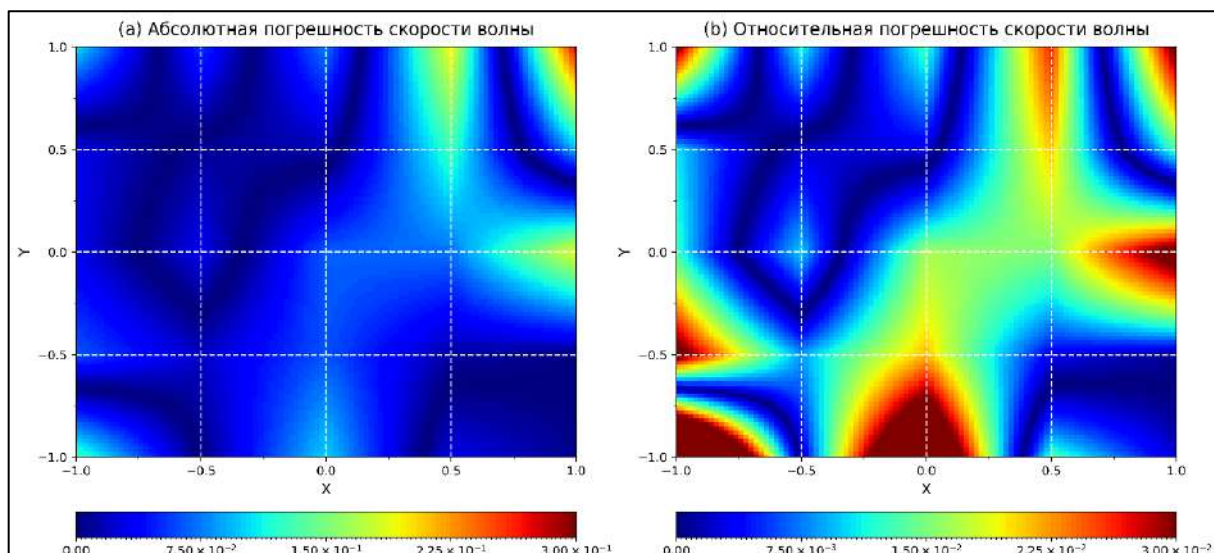


Рис. 5. Пример погрешностей начального приближения скорости волны, полученного с помощью нейронной сети. Подобласти отделены белой пунктирной линией. (a) Абсолютная погрешность. (b) Относительная погрешность

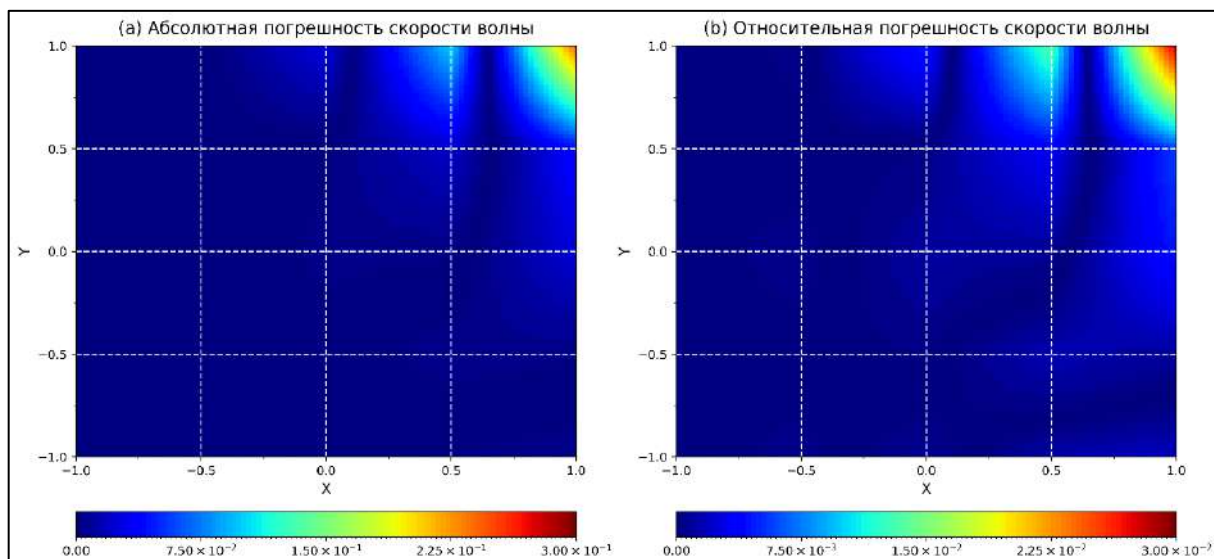


Рис. 6. Результат оптимизации по прошествии 47 итераций метода BFGS. Подобласти отделены белой пунктирной линией. Значение целевой функцией достигло порога 10^{-10} . (a) Абсолютная погрешность. (b) Относительная погрешность

4. Заключение

В работе была продемонстрирована работоспособность предложенного метода получения и использования нейросетевого начального приближения при итерационном решении обратной коэффициентной задачи распространения акустических волн. В приведенных результатах экспериментов начальное приближение, полученное с помощью нейронной сети, позволяет существенно уменьшить количество необходимых итераций метода решения обратной задачи для достижения приемлемого результата. Объединение классического итерационного метода решения обратной задачи и алгоритма, основанного на использовании физически информированной нейронной сети, позволяет компенсировать недостатки каждого из них по отдельности. Быстрый прогноз нейросетью дает возможность устранить продолжительную и наиболее непредсказуемую часть работы метода оптимизации. Недостаточная обоснованность и точность прогноза нейронной сети нивелируется благодаря использованию классического метода решения обратной задачи на заключительном этапе. В целом ускорение, полученное синергией классического подхода и нейронных сетей на проведенной серии экспериментов, достигает двух порядков, что показано в таблице 1.

Безусловно, стоит отметить ограничения и недостатки используемого подхода, заключающиеся в необходимости обучения нейронной сети и, естественно, зависимости результатов обучения от используемых данных, а также применения дискретизации только на равномерных сетках. Однако, подобные ограничения во многих задачах не имеют решающего значения и могут быть устранены на подготовительном этапе.

Литература

1. Самарский А.А. Численные методы решения обратных задач математической физики / А.А. Самарский, П.Н. Вабищевич – М.: Издательство ЛКИ, 2009. – 480 с.
2. Арсеньев-Образцов С.С., Соколов А.А. Построение конечно-элементного приближения решения обратной задачи с помощью физически информированных нейронных сетей // Автоматизация и информатизация ТЭК. – 2025. – № 3(620). – С. 47–55.
3. Karniadakis G.E., et al. Physics-informed machine learning // Nature Reviews Physics. – 2021. – Vol. 3. – P. 422–440.
4. Virieux J., Asnaashari A., Brossier R., et al. An introduction to full waveform inversion. // Encyclopedia of exploration geophysics. SEG. – 2017. – P. R1-1–R1-40.
5. Tan M., Le Q.V. EfficientNetV2: Smaller Models and Faster Training // Proc. of the 38-th International Conference on Machine Learning. – Virtual Event, 2021. – P. 10096–10106.
6. Nash S.G. A survey of truncated-Newton methods // J. of Computational and Applied Mathematics. – 2000. – Vol. 124, Issues 1–2. – P. 45–59.
7. Virtanen P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. // Nat Methods – 2020. – P. 261–272.