

Поведение модели на расширенной области определения, подбор гиперпараметров

Рассматривалась следующая задача:

$$\begin{aligned}
 & iq_t + ia_1q_x + a_2q_{xx} + ia_3q_{xxx} + a_4q_{xxxx} + ia_5q_{xxxxx} + a_6q_{xxxxxx} + q(b_1|q|^2 + b_2|q|^4 + b_3|q|^6) = 0, \\
 & x \in [x_L, x_R], t \in [0, t_{max}], \\
 & q(x, 0) = q_0(x), \\
 & q(x_L, t) = 0, \\
 & q(x_R, t) = 0, \\
 & q_x(x_L, t) = 0, \\
 & q_x(x_R, t) = 0, \\
 & q_{xx}(x_L, t) = 0, \\
 & q_{xx}(x_R, t) = 0
 \end{aligned} \tag{BVP6}$$

имеющая решение в виде:

$$q(x, t) = \frac{A_1}{ae^{x-C_0t-x_0} + \frac{\chi}{4a}e^{-(x-C_0t-x_0)}} e^{i(kx-\omega t+\theta_0)}, \tag{1}$$

где

$$\begin{aligned}
 A_1^{(1,2)} &= \pm \sqrt{2} \sqrt{\frac{\chi(a_2 - 6a_4k^2 + 12a_4k + 10a_4 + 75a_6k^4 + 150a_6k^2 + 91a_6)}{b_1}}, \\
 C_0 &= a_1 + 2a_2k + 8a_4k^3 + 96a_6k^5, \\
 \omega &= a_1k + a_2k^2 - a_2 + 3a_4k^4 - 6a_4k^2 - a_4 + 35a_6k^6 - 75a_6k^4 - 15a_6k^2 - a_6.
 \end{aligned} \tag{2}$$

$k = 1$, а χ, a, x_0, θ_0 - произвольные константы.

Параметры имеют вид:

$$\begin{aligned}
 a_5 &= -6a_6k, \\
 a_3 &= -4a_4k - 40a_6k^3, \\
 b_3 &= \frac{720a_6\chi^3}{A_1^6}, \\
 b_2 &= -\frac{24a_4\chi^2 + 360a_6\chi^2k^2 + 840a_6\chi^2}{A_1^4},
 \end{aligned} \tag{3}$$

где a_1, a_2, a_4, a_6, b_1 - произвольные.

В качестве a_i , где $i \in \{1, 2, 3, 4, 5, 6\}$, были взяты: 0.1, 0.1, -4.4, 0.1, -0.6000000000000001, 0.1, а в качестве b_j , где $j \in \{1, 2, 3\}$: 2.0, -0.1103806509211915, 0.0019498436834692017.

Модель была обучена на $x \in [-25, 25]$ с $t_{max} = 0.1$, а тестировалась на $x \in [-26, 26]$ с $t_{max} = 0.15$ и на $x \in [-50, 50]$ с $t_{max} = 1$. В качестве параметров для отладки были взяты: $\chi = 0.1$, $a = 0.1$, $x_0 = 0$, $\theta_0 = 0$.

Использовались 30 тыс. точек коллокаций, 3000 точек для начального условия и 1000 точек для левой и правой границ по x . Для генерации точек коллокаций

использовалось псевдослучайное распределение.

Нейронная сеть включала в себя следующие слои: один входной слой с 2 нейронами, 2 скрытых слоя, каждый из которых содержал 200 нейронов, и один выходной слой с 2 нейронами. В качестве функции активации был выбран синус.

Сначала обучение проводилось на протяжении 30000 итераций. Использовался оптимизатор Adam с параметрами: $initial_learning_rate = 1.1e - 02$, $loss_weights = [1, 1, 1000, 1000, 1, 1000, 1000, 1, 10000, 10000]$. Затем нейронная сеть обучалась в течение еще 10000 итераций с использованием оптимизатора L-BFGS с параметром $loss_weights = [1, 1, 1000, 1000, 1, 1000, 1000, 1, 10000, 10000]$.

Входные и выходные данные не подвергались дополнительной обработке.

Были получены следующие результаты:

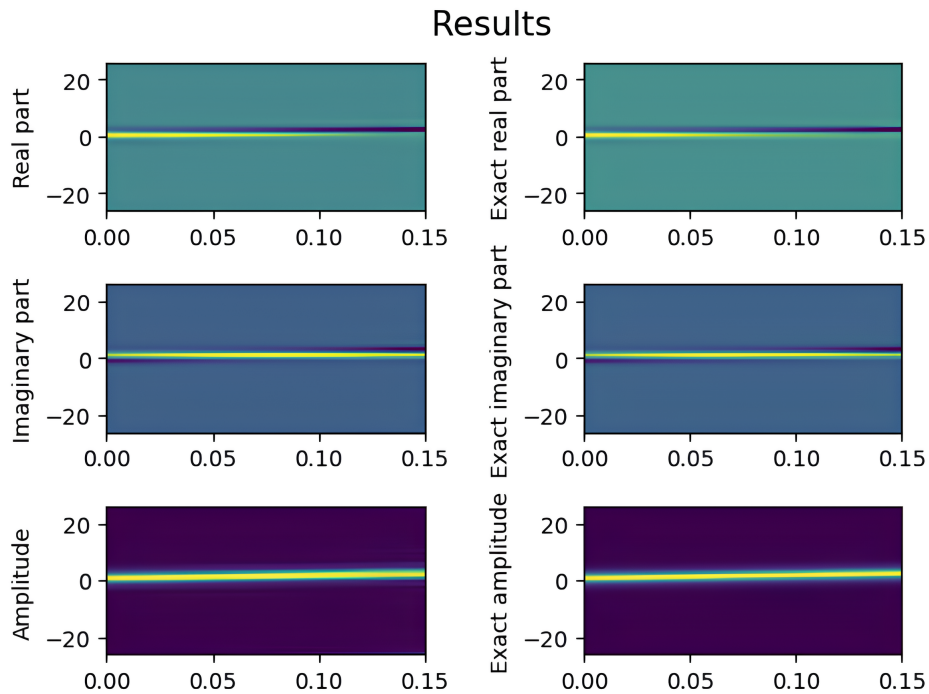


Рис. 1: Результаты решения задачи BVP6 для $x \in [-26, 26]$ с $t_{max} = 0.15$

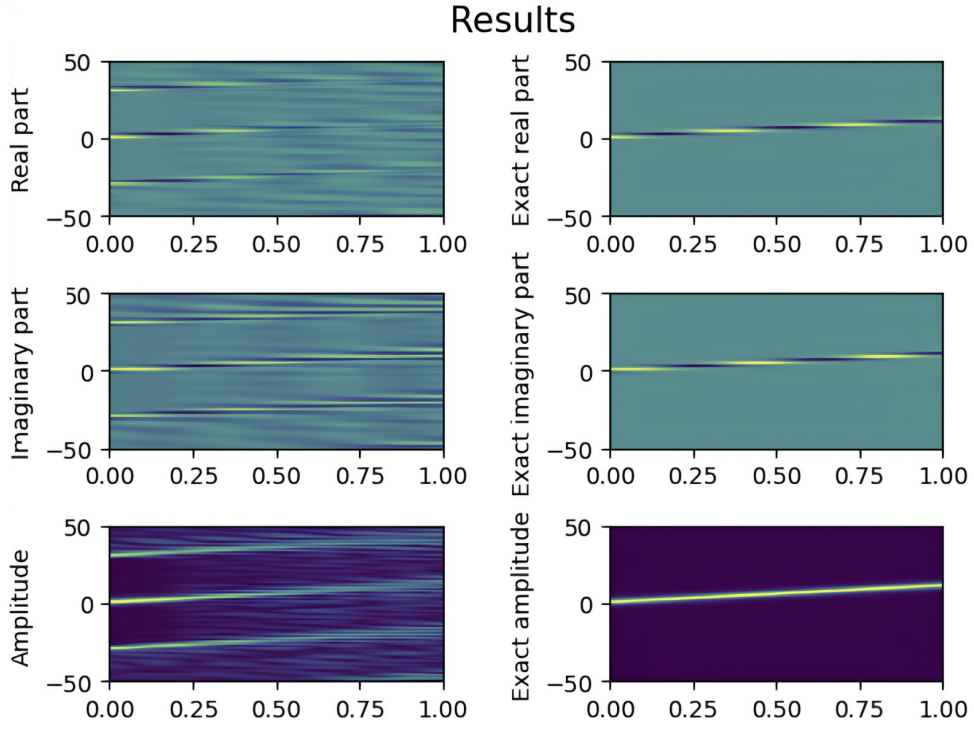


Рис. 2: Результаты решения задачи BVP6 для $x \in [-50, 50]$ с $t_{max} = 1$

$Lw1_{max}$	$Lw1_{mean}$	$Lw2_{max}$	$Lw2_{mean}$	Rel_h
5.2309	1.7480	4.6995	1.5828	0.0790
152.1487	112.4568	165.0656	73.1291	1.3055

Таблица 1: Основные метрики

Данные таблицы 1 свидетельствуют о том, что модель плохо ведёт себя на тех областях, где не училась.

Далее был осуществлен возврат к предыдущей области определения и составлена таблица 2, отражающая изменение Rel_h .

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	float(32/64)	Rel_h
1	1	1000	1000	1	1000	1000	1	1000	1000	32	0.0904
1	1	1000	1000	1	1000	1000	1	1000	1000	64	0.0822
1	1	1000	1000	1	1000	1000	1	10000	10000	32	0.0773
1	1	1000	1000	1	1000	1000	1	10000	10000	64	0.0468
1	1	1000	1000	1	1000	1000	1	100000	100000	32	0.0731
1	1	1000	1000	1	1000	1000	1	100000	100000	64	0.0781

Таблица 2: Точность модели

Наибольшая точность достигается при конфигурации весов:
 $loss_weights = [1, 1, 1000, 1000, 1, 1000, 1000, 1, 10000, 10000]$ и float64.