

# Нейронные сети в решении уравнений квантовой физики

В.А. Руднев

Санкт-Петербургский государственный университет  
Кафедра вычислительной физики



## Квантовая механика и проклятие размерности

Число частиц	Размерность задачи	Размерность редуцированной задачи
1	3	1
2	6	1 – 3
3	9	3 – 6
4	12	6 – 9
...	...	...
n	3n	(3n-3) – (3n-6)

Кинетическая энергия:

$$T = -\frac{\hbar^2}{2m} \sum_{k=1}^n \Delta_{r_k}$$

# Квантовая механика и проклятие размерности

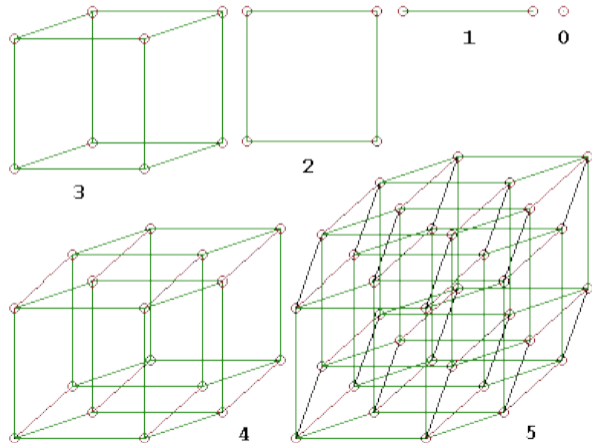
$$(-\Delta + V - E)\Psi = 0$$

$$\Psi \in L_2(\mathcal{R}^n)$$

Размерность сетки скалируется  
как

$$N \sim N_1^n$$

Вычислительные ресурсы  
растут экспоненциально по  
размерности проблемы (числу  
частиц).



# Нейронные сети

Аффинное преобразование

$$W_l(x) = W^{(l)}x + W_0^{(l)}$$

Функция активации

$$z = \sigma(y)$$

Перцептрон

$$\mathcal{P}(x; W) = W_2 \circ \sigma \circ W_1 x$$

Многослойный перцептрон

$$\mathcal{N}_{MLP}(x; W) = W_n \circ \sigma \circ W_{n-1} \circ \dots \circ \sigma \circ W_1 x$$

# Нейронные сети

$P(x)$  сколь угодно точно аппроксимирует непрерывную функцию на компактном множестве.

**Целевой функционал.**

Задачи регрессии:

$$L(W) = \frac{1}{N} \sum_{i=1}^N (\mathcal{N}_{MLP}(x_i; W) - y_i)^2 .$$

Задачи классификации:

$$L(W) = -\frac{1}{N} \sum_{i=1}^N [y_i \log \mathcal{N}_{MLP}(x_i; W) + (1 - y_i) \log(1 - \mathcal{N}_{MLP}(x_i; W))] .$$

Обучение:

$$W = \underset{W}{\operatorname{argmin}} L(W)$$

# Пример

## Вычисление производных

Производные по параметрам  $D_W L(W)$  и производные по входным данным  $D_x \mathcal{N}_{MLP}(x; W)$  вычисляются по формуле производных композиции функций:

$$D(F \circ G) = (DF \circ G) DG .$$

Время вычисления сравнимо со временем вычисления самой функции.  
Производная слоя:

$$D_x(\sigma \circ W) = D_x(\sigma(Wx)) = \sigma'(Wx)W$$

## Можно решать дифференциальные уравнения!

$\mathcal{D}$  - некоторый дифференциальный оператор. Дифференциальное уравнение

$$\mathcal{D}f(x) = b$$

Нейросетевая подстановка:

$$f(x) = \mathcal{N}(x, W)$$

Целевой функционал:

$$L(W) = \frac{1}{N} \sum_{i=1}^N (\mathcal{D} \mathcal{N}(x, W) |_{x=x_i} - b)^2$$



# Уравнение Шредингера, решение Лагариса

## Artificial neural network methods in quantum mechanics

I.E. Lagaris<sup>1</sup>, A. Likas, D.I. Fotiadis

*Department of Computer Science, University of Ioannina, P.O. Box 1186, GR 45110 Ioannina, Greece*

Received 17 March 1997; revised 22 April 1997

---

### Abstract

In a previous article we have shown how one can employ Artificial Neural Networks (ANNs) in order to solve non-homogeneous ordinary and partial differential equations. In the present work we consider the solution of eigenvalue problems for differential and integrodifferential operators, using ANNs. We start by considering the Schrödinger equation for the Morse potential that has an analytically known solution, to test the accuracy of the method. We then proceed with the Schrödinger and the Dirac equations for a muonic atom, as well as with a nonlocal Schrödinger integrodifferential equation that models the  $n + \alpha$  system in the framework of the resonating group method. In two dimensions we consider the well-studied Henon–Heiles Hamiltonian and in three dimensions the model problem of three coupled anharmonic oscillators. The method in all of the treated cases proved to be highly accurate, robust and efficient. Hence it is a promising tool for tackling problems of higher complexity and dimensionality. © 1997 Elsevier Science B.V.

*PACS:* 02.60.Lj; 02.60.Nm; 02.70.Jn; 03.65.Ge

*Keywords:* Neural networks; Eigenvalue problems; Schrödinger; Dirac; Collocation; Optimization

## Решение Лагариса

Задача

$$(H - EI)\Psi(x) = 0$$
$$\Psi(x)|_{x \in \partial\Omega} = 0$$

Представление решения:

$$\Psi(x) = \Psi_W(x) = B(x)\mathcal{P}(x; W),$$

$W$  – веса нейронной сети.

Оценка энергии

$$E_W = \frac{\langle \Psi_W(x) | H | \Psi_W(x) \rangle}{\langle \Psi_W(x) | \Psi_W(x) \rangle}$$

Обучение:

$$W = \underset{W}{\operatorname{argmin}} L(W)$$

$$L(W) = \|(H - E_W I)\Psi_W(x)\|^2$$

# Решение Лагариса

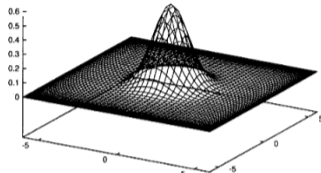


Fig. 4.

Fig. 4. Ground state of the Henon-Heiles problem ( $\epsilon = 0.99866$ ).

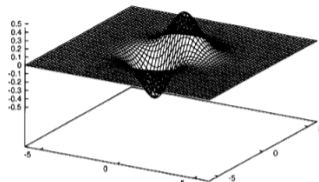


Fig. 5.

Fig. 5. First excited state of the Henon-Heiles problem ( $\epsilon = 1.990107$ ).

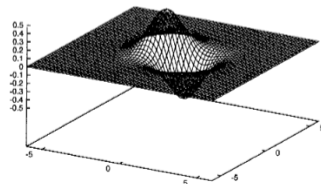


Fig. 6.

Fig. 6. Second excited state (degenerate) of the Henon-Heiles problem ( $\epsilon = 1.990107$ ).

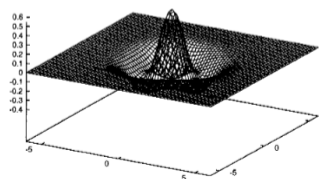


Fig. 7.

Fig. 7. Third excited state of the Henon-Heiles problem ( $\epsilon = 2.957225$ ).

## Общие проблемы традиционного подхода

- Невыпуклая функция потерь, сложности оптимизации
- Проблема локальных минимумов
- Проблема переобучения
- Проблема выборки в пространстве большой размерности

## Смена парадигмы: глубокое обучение

Глубокая сеть: сеть, с числом скрытых слоёв больше двух. Многослойная сеть может обладать преимуществами:

- Сложность воспроизводимой функции растёт геометрически по числу слоёв, а число параметров - линейно.  
<https://arxiv.org/abs/1606.05336>
- Количество параметров, необходимых для аппроксимации, растёт линейно по размерности задачи  
<https://arxiv.org/abs/1905.12882>
- Есть (нестрогие) основания полагать, что все локальные минимумы эквивалентны  
<https://arxiv.org/abs/1702.08580>

# Глубокое обучение - решение

- ~~Невыпуклая функция потерь, сложности оптимизации~~
- ~~Проблема локальных минимумов~~
- ~~Проблема переобучения~~
- Проблема выборки в пространстве большой размерности

## Пакетная оценка целевого функционала

Целевой функционал:

$$L(\mathbf{W}) = \mathbb{E}_{x_i \sim \rho_X} R(x_i; \mathbf{W})^2$$

Оценка градиента

$$\nabla_{\mathbf{W}} L(\mathbf{W}) = \nabla_{\mathbf{W}} \mathbb{E}_{x_i \sim \rho_X} R(x_i; \mathbf{W})^2$$

"Стохастический градиент"

$$\nabla_{\mathbf{W}} L_i(\mathbf{W}) = \nabla_{\mathbf{W}} R(x_i; \mathbf{W})^2$$

Пакетная оценка

$$\nabla_{\mathbf{W}} L_m(\mathbf{W}) = \nabla_{\mathbf{W}} \mathbb{E}_{x_i \in X_m} R(x_i; \mathbf{W})^2$$

$X_m = \{x_1, x_2, \dots, x_n\}$ ,  $x_i \sim \rho_X$ ,  $n$  - размер пакета

## Пример: многомерный квантовый гармонический осциллятор

$n$	0	1	2	3	4	5	6	7	8	9
$E_n$	1.5000	2.5222	2.5228	2.5252	3.5002	3.5024	3.5067	3.5092	3.5094	3.5097
$\mathcal{E}_E$	0.0000	0.0089	0.0091	0.0101	0.0001	0.0007	0.0019	0.0026	0.0027	0.0028

Таблица 8: Оценки энергий  $N = 10$  состояний для трёхмерного гармонического осциллятора.

$n$	0	1	2	3	4	5	6	7	8	9
$E_n$	2.0000	3.0066	3.0091	3.0127	3.0136	3.9959	3.9990	4.0051	4.0070	4.0075
$\mathcal{E}_E$	0.0000	0.0022	0.0030	0.0042	0.0045	-0.0010	-0.0002	0.0013	0.0018	0.0019

Таблица 9: То же, что и в Таблице 8, но для четырёхмерной задачи.

$n$	0	1	2	3	4	5	6	7	8	9
$E_n$	2.4994	3.4942	3.4963	3.5000	3.5015	3.5136	4.4941	4.4941	4.4984	4.5123
$\mathcal{E}_E$	-0.0003	-0.0017	-0.0010	-0.0000	0.0004	0.0039	-0.0013	-0.0013	-0.0004	0.0027

Таблица 10: То же, что и в Таблице 8, но для пятимерной задачи.



## Сложности

- С ростом размерности распределение точек коллокаций становится всё более критическим (объём шара/объём куба  $\rightarrow 0$ ).
- Сложные целевые функционалы требуют оптимизации гиперпараметров.
- Строгая оценка физических наблюдаемых всё равно требует вычисления многомерных интегралов, (квази)Monte-Carlo.
- Подбор оптимальной архитектуры нейронной сети - искусство.

## Ближкие новые идеи

- Сети Колмогорова-Арнольда (KAN).
- Операторные сети (DeepONet).
- Прямое вычисление инвариантных подпространств.



# Backup slides

# Проблемы и решения

## Проблемы подхода

1. Локальные минимумы затрудняют приближение конкретного состояния.
2. Обучение нейронной сети требует больших вычислительных затрат.
3. Многомерное интегрирование существенно усложняет вычисление энергии.

## Решения

- Вычислять несколько –  $N$  – состояний одной сетью.

## Цена решения

- Сложный целевой функционал:

$$L(W) = \alpha \sum_{k=1}^N E_k + \beta \sum_{k=1}^N (\|\psi_k\| - 1)^2 + \gamma \sum_{k=1}^N \sum_{j \neq k} \frac{\langle \psi_k | \psi_j \rangle^2}{\langle \psi_k | \psi_k \rangle \langle \psi_j | \psi_j \rangle} + \\ + \delta \sum_{k=1}^N r_k^{(b.c.)} + \epsilon \sum_{k=1}^N R_k^{(eq.)} + \lambda \|W\|^2$$

## Новый подход: основные теоремы

Одномерная задача самосогласованной эволюции.

$$\begin{cases} \partial_t \Psi(t, x) = -H\Psi(t, x) \\ \Psi(0, x) = \frac{\Psi(T, x)}{\Psi(T, x_0)} \\ b(\Psi(t, x))|_{x \in \partial\Omega} = 0 \end{cases} . \quad (1)$$

Здесь  $H$  - полуограниченный снизу самосопряженный оператор с дискретным спектром,  $b$  определяет соответствующие граничные условия,  $x_0 \in \Omega \setminus \partial\Omega$ .

### Theorem

*Задача (1) имеет нетривиальное решение тогда и только тогда, когда  $\Psi(0, x) = c\phi(x)$ , где  $c \neq 0$ ,  $\phi(x)$  - собственная функция  $H$ ,  $(H - E)\phi(x) = 0$ .*

## Новый подход: одно состояние

Proof.

1. Пусть  $\Psi(0, x) = c\phi(x)$ ;  $\Psi(t, x) = ce^{-Et}\phi(x)$  удовлетворяет уравнению эволюции. Начальное условие удовлетворяется если  $c = \frac{1}{\phi(x_0)}$ , и  $c$  определено с точностью до нормировки.

2. Пусть  $\Psi(t, x)$  удовлетворяет (1).  $\Psi(t, x) = \sum_{l=1}^{\infty} a_l e^{-E_l t} \phi_l(x)$ . Пусть собственные функции нормированы:  $\langle \phi_k | \phi_l \rangle = \delta_{kl}$ . Из “начального условия” следует

$$\sum_{l=1}^{\infty} a_l e^{-E_l T} \phi_l(x_0) = e^{-E_k T}, \quad k = 1, \dots, \infty.$$

Эта система разрешима только если для любых  $l \neq k$   $a_l = 0$ , что означает  $\Psi(0, x) = \frac{\phi_k(x)}{\phi_k(x_0)} \equiv c\phi_k(x)$ . □

## Новый подход: существование решения для нескольких состояний

Зададим начальное состояние как линейную комбинацию состояний отвечающих различным значениям времени.

Рассмотрим  $\{t_i\}_{i=1}^n$ ,  $t_i \neq t_{i'}$ ,  $t_i \in (0, T]$ .

$$\begin{cases} \partial_t \Psi(t, x) = -H\Psi(t, x) \\ \Psi(0, x) = \sum_{i=1}^n \alpha_i \Psi(t_i, x) \\ b(\Psi(t, x))|_{x \in \partial\Omega} = 0 \end{cases} \quad (2)$$

### Theorem

$\exists \alpha_i$ : задача (2) имеет решение  $\Psi(t, x)$  и для  $\forall t \in [0, t]$   $\Psi(t, x) \in I_n(H)$ ,  $I_n(H)$  –  $n$ -мерное инвариантное подпространство оператора  $H$ .



## Новый подход: существование решения для нескольких состояний

$$\begin{cases} \partial_t \Psi(t, x) = -H\Psi(t, x) \\ \Psi(0, x) = \sum_{i=1}^n \alpha_i \Psi(t_i, x) \\ b(\Psi(t, x))|_{x \in \partial\Omega} = 0 \end{cases} .$$

Proof.

Построим решение явно.  $\Psi(t, x) = \sum_{i=1}^n \beta_i e^{-E_i t} \phi_i(x)$  приводит к системе уравнений на коэффициенты  $\alpha_i$ :

$$\sum_{i=1}^n e^{-E_j t_i} \alpha_i = 1, \quad j = 1, \dots, n$$

Эта система разрешима (необязательно однозначно). □

## Новый подход: построение собственного подпространства

### Theorem

Пусть  $\Psi(t, x)$  - решение задачи (2). Тогда для любого фиксированного  $t \in [0, T]$  оно – как функция  $x$  – принадлежит инвариантному подпространству  $I_{n'}(H)$  оператора  $H$ ,  $\dim I_{n'}(H) = n' \leq n$ .

### Proof.

Общий вид решения задачи  $\Psi(t, x) = \sum_{i=1}^{\infty} \beta_i e^{-E_i t} \phi_i(x)$  приводит к системе уравнений на коэффициенты  $\alpha_i$ :

$$\sum_{i=1}^n e^{-E_j t_i} \alpha_i = 1, \quad j = 1, \dots, \infty$$

Эта система разрешима только если значение  $j$  не превышает  $n$ . □

# Базовый алгоритм

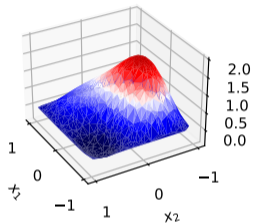
Целевой функционал

$$L(W) = \frac{1}{N} \sum_{k=1}^N |(\partial_t + H) \Psi_W(t, x)|_{t=t_k, x=x_k}|^2 + \frac{1}{N} \sum_{k=1}^N |\Psi_W(0, x_k) - \bar{\Psi}(x)|^2$$

- 1: Установить начальное состояние  $\bar{\Psi}(x) = \Psi_0(x)$
- 2: **loop**
- 3:     произвести шаг обучения сети  $\Psi_W(t_i, x)$
- 4:     **for**  $i = 1, i \leq n, i++$  **do**
- 5:          $\psi_i(x) := \Psi_W(t_i, x)$
- 6:     **end for**
- 7:     положить  $\alpha_i = \operatorname{argmin}_{\alpha_i} \|\Psi_0(x) - \sum_{i=1}^n \alpha_i \psi_i(x)\|^2$
- 8:      $\bar{\Psi}(x) := \sum_{i=1}^n \alpha_i \Psi_i(x)$
- 9: **end loop**

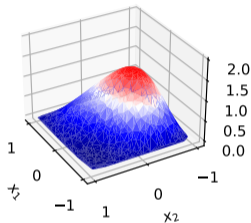
# Some numerical illustrations

$t=0.067$



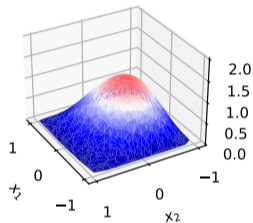
$E=4.939$

$t=0.115$

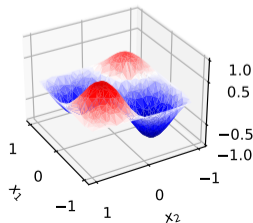
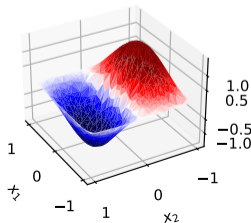
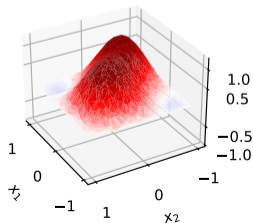


$E=12.350$

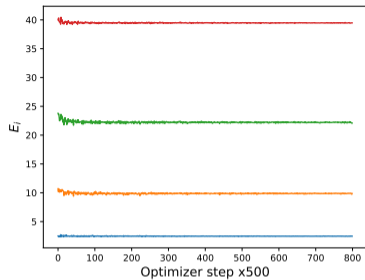
$t=0.200$



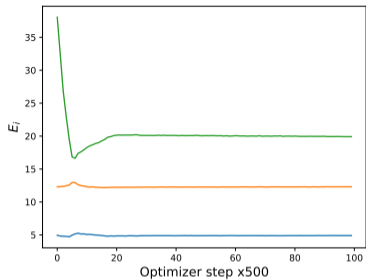
$E=19.717$



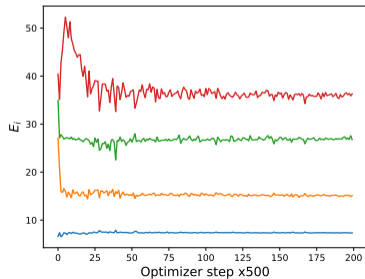
1D



2D



3D



4D

