

УДК 004.942:519.6

Регуляризация процедуры определения энергии фотонов методами искусственного интеллекта

А. Ю. Баев, В. Т. Лазурик

Харьковский национальный университет имени В.Н. Каразина, Украина

Рассмотрена возможность использования искусственного интеллекта для решения некорректной задачи определения дискретного спектра фотонов на основе значений функции распределения дозы измеренных методом поглотителей. Предложен статистический подход сравнения различных численных методов обработки экспериментальных данных. Проведен сопоставительный анализ метода наименьших квадратов и нейронной сети для решения частного случая задачи восстановления спектра фотонного излучения.

Ключевые слова: *нейронные сети, метод наименьших квадратов, обратная задача спектроскопии.*

Розглядається можливість використання штучного інтелекту для розв'язання некоректної задачі визначення дискретного спектру фотонів на основі значень функції розподілення дози виміряних методом поглиначів. Запропоновано статистичний підхід порівняння різноманітних чисельних методів обробки експериментальних даних. Виконано порівняльний аналіз метода найменших квадратів та нейронної мережі для вирішення окремого випадку задачі відновлення спектру фотонного випромінювання.

Ключові слова: *нейронні мережі, метод найменших квадратів, обернена задача спектроскопії.*

It's shown how to use artificial intelligence to solve ill-posed problem of discrete photon's spectrum recognition. The values of dose distribution function which were measured by absorber's method were used. The statistical approach was offered for the purpose to compare various numerical methods of the experimental data evaluation problem. The comparative analysis between the least-squares method and the neural network was made in the particular case of the photon radiation's spectrum reconstruction problem.

Key words: *neural networks, least-squares methods, inverse and ill-posed problem of spectroscopy.*

1. Введение

Решение ряда прикладных задач спектроскопии сводятся к решению интегрального уравнения Фредгольма I рода [1]. Одним из эффективных методов решения уравнения является метод квадратур [2], в котором интегральное уравнение заменяется конечномерной суммой вида:

$$\sum_j A_{ij} Y_j = f_i \quad (1)$$

Для решения системы линейных алгебраических уравнений (1) применяются различные численные методы [3], однако получаемые этими методами решения являются некорректным по Адамару, поэтому для решения уравнения (1) применяют специальный класс алгоритмов регуляризации [4]. В то же время развитие теории искусственного интеллекта дает предпосылки к использованию методов машинного обучения [5] для решения уравнения (1).

В работе рассматривается простейший случай задачи восстановления дискретного спектра γ -излучения по величинам дозы, измеренной за

поглотителями различной толщины [7]. В качестве метода искусственного интеллекта использована нейронная сеть (НС) обученная по алгоритму обратного распространения ошибки [5, 6]. Для выявления преимуществ НС и сопоставительного анализа разработан статистический подход сравнения методов, решающих задачу обработки экспериментальных данных (ЭД). В качестве эталонного метода использован метод наименьших квадратов (МНК). Процесс обучения с учителем предполагает использование истинных (известных) значений [6], поэтому для построения системы обработки ЭД с помощью искусственного интеллекта разработаны модели физического процесса и процесса получения экспериментальных данных.

2. Статистический подход сравнения методов обработки данных

Для сравнения методов обработки разработан алгоритм (см. рис. 1), последовательно выполняющий процедуры: 1) моделирование процесса получения ЭД, 2) обработка ЭД каждым из сравниваемых методов и 3) сравнение методов по выбранным критериям (см. рис. 1).



Рис. 1. Общая схема эксперимента

2.1 Модель процесса получения ЭД

Модель процесса получения ЭД включает два этапа: получение истинных значений и добавление неопределенности, обусловленной несовершенством измерительных приборов и внешними воздействиями на полезный сигнал.

Получение истинных значений предполагает наличие знаний об идеальном физическом процессе. В общем случае истинная зависимость, в соответствии с которой наблюдаются значения измеряемой величины на выходе идеального измерительного прибора, имеет вид:

$$y = f(x, c_0, \dots, c_s) = f(x, \bar{c}), \quad (2)$$

где x – независимая, а y – зависимая переменные, \bar{c} – набор физических параметров. Компоненты вектора \bar{c} задают вид зависимости и имеют ряд ограничений, определяемых физическим смыслом. Результат процесса получения истинных данных есть набор пар (x_i, y_i) , где $i = 1, n$, n – количество измерений в одном эксперименте.

Особенности конструкции измерительных приборов и влияние внешних воздействий определяют отклонения измеренных значений $(\tilde{x}_i, \tilde{y}_i)$, вследствие чего ЭД не совпадают с истинными данными. В разработанной модели указанные воздействия имитируются как нормально распределенная погрешность измерений:

$$\Phi_X(u) = u + Q_X k_X \sqrt{-2 \ln(r_1)} \cos(r_2), \quad (3)$$

где k_X – погрешность определения значения u , r_1, r_2 – независимые равномерно распределенные в $[0; 1]$ случайные величины, значение Q_X определяет тип погрешности: абсолютная для $Q_X = const$ и относительная для $Q_X = u$.

Выбор x выполняется в интервале $[0, 1]$ так, что:

I. Выбирается сетка $x_i : (x_{i+1} - x_i) = const$, после чего к каждому узлу добавляется нормально распределенная погрешность измерений:

$$\tilde{x}_i = \Phi_x(x_i), \quad (4)$$

II. В качестве \tilde{x}_i выбираются случайные равномерно распределенные случайные числа, т.е.

$$\tilde{x}_i = r_3. \quad (5)$$

где r_3 – независимая равномерно распределенная в $[0; 1]$ случайная величина.

Отклонение значений y выполнено в соответствии с:

$$\tilde{y}_i = \Phi_Y(y_i), \quad (6)$$

Отметим, что многие современные методы используют сведения об обрабатываемых данных в виде параметров [1, 4], основанных на теоретических знаниях предметной области: области определения, коэффициенты корреляции и т.д. Процесс получения такой информации интерпретируется как обучение и требует наборов данных, предназначенных для подгонки значений соответствующих параметров. В этой связи разработанная модель процесса получения ЭД предусматривает возможность генерации выборки для обучения и тестирования. В зависимости от задачи, решаемой сравниваемыми методами, указанные выборки могут пересекаться или нет.

Результатом работы алгоритма, реализующего описанную выше модель, является набор $(\tilde{x}_i, \tilde{y}_i)$, который дополняется набором истинных параметров $\bar{t} = (t_1, \dots, t_l)$, на основе которых рассчитываются ошибки методов обработки:

$$(\tilde{x}_1, \tilde{y}_1, \tilde{x}_2, \tilde{y}_2, \dots, \tilde{x}_n, \tilde{y}_n, t_1, \dots, t_l). \quad (7)$$

Для формирования обучающей выборки совокупность векторов (7) записывается в матрицу вида

$$\begin{pmatrix} \tilde{x}_1, \tilde{y}_1, \tilde{x}_2, \tilde{y}_2, \dots, \tilde{x}_n, \tilde{y}_n, t_1, \dots, t_l \\ \vdots \\ \tilde{x}_1, \tilde{y}_1, \tilde{x}_2, \tilde{y}_2, \dots, \tilde{x}_n, \tilde{y}_n, t_1, \dots, t_l \end{pmatrix}_{m \times l}, \quad (8)$$

где m – объем обучающей выборки, l – размерность вектора-результата обработки.

Необходимость создания и хранения матрицы (8) в рамках одного статистического эксперимента диктуется необходимостью многократного прохода по обучающим данным во время обучения. Напротив, во время тестирования данные не собираются в матрицу, а генерируются перед и уничтожаются после обработки, что позволяет экономить вычислительные ресурсы. Стоит отметить, что обучающая выборка может быть разбита на подмножества для непосредственно обучения и валидации процесса обучения в зависимости от используемого алгоритма обучения, а выборка для тестирования

используется только для апробирования методов обработки данных. Таким образом, совокупность настроек полученных в результате обучения можно интерпретировать как параметр регуляризации, т.е. внесение знаний о конкретной предметной области в процесс обработки данных, а обучение – метод определения этого параметра.

В работе для получения матрицы (8) используется набор опорных векторов $\{\bar{c}_e\}$, а для тестирования – набор $\{\bar{c}_t\} : \forall \bar{c} \in \{\bar{c}_t\}, \bar{c} \notin \{\bar{c}_e\}$. Таким образом $\{\bar{c}\} = \{\bar{c}_t\} \cup \{\bar{c}_e\}$. При определении множества $\{\bar{c}\}$ допускаются следующие предположения:

- неопределенность δ^* истинного закона считается много меньше, чем неопределенность процедуры измерения,
- значения компонент различных \bar{c} дискретизированы таким образом, что различия в наборах истинных закономерностей много больше δ^* .

2.2 Обработка ЭД и сравнение методов

В рамках разрабатываемого статистического подхода акцент при сравнении ставится на анализ результатов работы каждого из сравниваемых методов. В этой связи сам метод обработки рассматривается как черный ящик, вход которого в общем случае есть вектор данных (7) и набор параметров настраивающих работу метода, а выход – вектор результатов обработки $\bar{r} = (r_1, \dots, r_l)$. Элементы вектора \bar{r} , а также критерии оценки ошибки результатов зависят от задачи, решаемой сравниваемыми методами.

В работе рассчитываются численные значения компонент вектора \bar{c} , который в совокупности с функцией (2) описывает обрабатываемый набор ЭД. Т.е. $\bar{r} = \bar{c}$. После обработки тестовой выборки каждым из сравниваемых методов рассчитываются статистические характеристики полученных результатов, т.е. рассчитываются математическое ожидание и среднеквадратичное отклонение для каждого компонента вектора \bar{c} .

3. Решение задачи определения спектра фотонного излучения

В работе рассмотрен частный случай задачи определения дискретного спектра фотонов на основе значений функции распределения дозы измеренных методом поглотителей [7]. Для описания истинных зависимостей физических величин использована зависимость вида:

$$f(x, \bar{c}) = Y_1 e^{-\mu_1 x} + Y_2 e^{-\mu_2 x} \quad (9)$$

где Y_1, Y_2 – интенсивности компонент спектра, μ_1, μ_2 – коэффициенты ослабления γ -излучения и

$$\bar{c} = (Y_1, \mu_1, Y_2, \mu_2). \quad (10)$$

Согласно [7] физический смысл коэффициентов интенсивности предполагает следующие ограничения:

$$Y_1 \geq 0, Y_2 \geq 0. \quad (11)$$

Аппроксимирующая функция для МНК выбрана как линейная комбинация $\varphi(x) = Y_1 \varphi_1(x) + Y_2 \varphi_2(x)$, где $\varphi_i(x) = e^{-\mu_i x}$ [7]. Таким образом, работа МНК сводится к нахождению коэффициентов Y_1, Y_2 . Обработка данных МНК может

дать значения без физического смысла (отказ метода), т.е. в рамках моделируемого эксперимента – не выполнение (11). Учет отказов значительно изменяет характер зависимости статистических характеристик от величины неопределенности процесса измерения.

В качестве метода использующего модель искусственного интеллекта выбрана полносвязная НС, обученная по алгоритму обратного распространения ошибки. На вход сети подается вектор $(x_1, f_1, x_2, f_2, \dots, x_n, f_n)$, на выходе – численные значения компонент вектора \bar{c} пропорциональные нормирующему вектору, который формируется на основании множества $\{\bar{c}_e\}$.

4. Результаты и обсуждение

В работе реализовано программное обеспечение для сравнения алгоритмов анализа в соответствии с описанной методикой и проведена серия компьютерных экспериментов, параметры которых указаны в Табл. 1.

Табл.1. Параметры компьютерного эксперимента

Количество экспериментальных точек (n)	5
Области определения	$x \in [0,1]; y \in [0,1]$
Множество $\{\bar{c}_e\}$	0.6 0.4 0.2 0.5 0.8 0.6 0.4 0.7 1.0 0.8 0.6 0.9
Множество $\{\bar{c}_l\}$	$\bar{c}_1 = (0.7 \ 0.5 \ 0.3 \ 0.6)$ $\bar{c}_2 = (0.9 \ 0.7 \ 0.5 \ 0.8)$
Объем обучающей выборки (на каждый \bar{c})	100
Объем тестирующей выборки (на каждый \bar{c})	1000
Неопределенность измерения (абсолютная)	$0 \div 0.1$ с шагом 0.01

На рис. 2, 3, 4 показаны результаты статистического анализа сравниваемых методов обработки. На графиках математического ожидания обозначены истинные значения (прямые линии различного стиля) и рассчитанные значения среднего (точки различной формы). Анализ выполнен для коэффициентов Y_1, Y_2 двух векторов тестовой выборки.

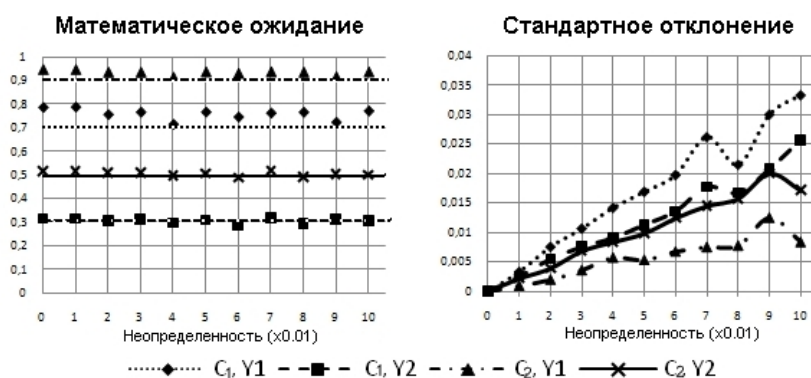


Рис. 2 Результат обработки ЭД с помощью НС

На рис. 2-3 показаны результаты статистического анализа для всех решений каждого метода без учета отказов. На графике среднего значения найденных

методом НС коэффициентов (рис. 2) видно смещение в обеих зависимостях (\bar{c}_1, \bar{c}_2 , для коэффициента Y_1), которого нет для значений определенных МНК (рис. 3). Природа упомянутого смещения, вероятно, обусловлена количеством нейронов в НС и видом их активационной функции (в работе – гиперболический тангенс). Зависимость величины смещения математического ожидания от структуры НС представляет интерес для дальнейшего исследования. Сравнение графиков среднеквадратичного отклонения на рис. 2 и 3 показывает, что НС определяет необходимые коэффициенты с отклонением на два порядка меньшим, чем у решений, полученных с помощью МНК.

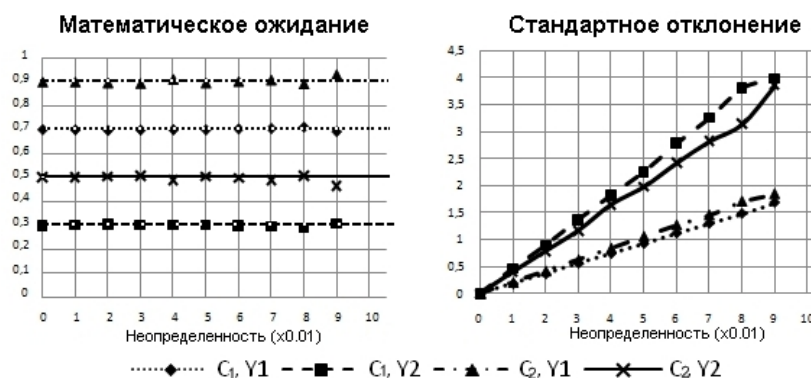


Рис. 3 Результат обработки ЭД с помощью МНК

На рис. 4 представлена зависимость вероятности отказа МНК от величины неопределенности измерения входных данных, т.е. количество результатов обработки для которых не выполняется условие (11). Отметим, что при обучении нейронной сети используется только данные, удовлетворяющие условию (11), и поэтому не следует ожидать решений без физического смысла.

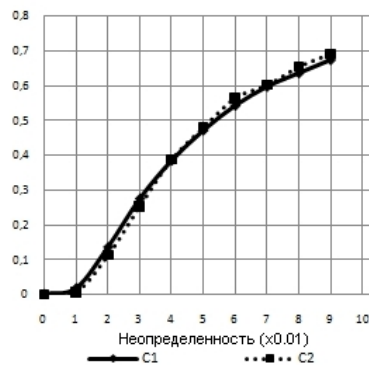


Рис. 4 Вероятность отказа

График на рис. 5 показывает результаты статистического анализа только для имеющих физический смысл решений полученных МНК. На графике математического ожидания видны большие смещения среднего для МНК при неопределенности измерения более 0.02 (обусловлены растущим количеством отказов). Сравнение графиков на рис. 2 и 5 показывает высокое качество результатов, полученных методом, использующим НС, в сравнении с МНК.

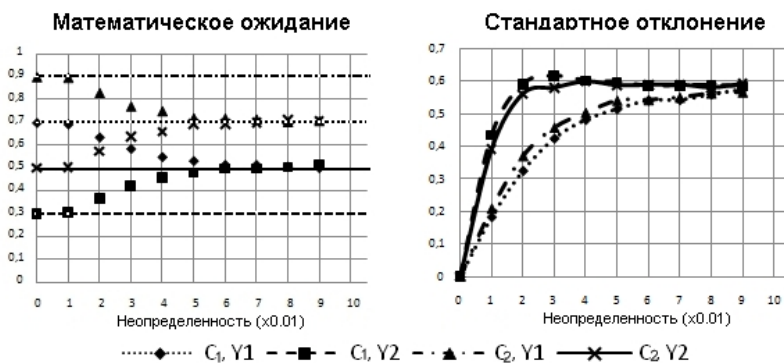


Рис. 5 Результат обработки ЭД с помощью МНК

5. Выводы

Рассмотрена регуляризация решений некорректной задачи восстановления спектра фотонного излучения методами, использующими искусственный интеллект. Развита статистический подход сравнения вычислительных методов обработки эмпирических данных, включающий модели процессов получения и обработки данных и критерии оценивания качества обработки. На основе вычислительного эксперимента для случая задачи восстановления дискретного спектра показано преимущество метода использующего нейронную сеть в сравнении с методом наименьших квадратов. Преимущество искусственного интеллекта обусловлено использованием априорных знаний о предметной области, которые задают ограничения возможных результатов обработки. Взгляд на нейронные сети как на методы регуляризации представляет интерес сравнения искусственного интеллекта с другими современными методами регуляризации и фильтрации.

ЛИТЕРАТУРА

1. Сизиков В.С. Математические методы обработки результатов измерений: Учебник для вузов. – СПб: Политехника, 2001. – 240с.
2. Белоцерковский С.М., Лифанов И.К. Численные методы в сингулярных интегральных уравнениях, - М.: Наука, 1985. – 256с.
3. Одокиенко С.Н. Особенности методов решения интегральных уравнений Вольтера I рода // Математичне та комп'ютерне моделювання. – 2008. – Сер. «Технічні науки», вип. 1 – С.130-139
4. Кирьянов Д.В., Кирьянова Е.Н. Вычислительная физика – М.: Полибук Мультимедиа, 2006. – 352 с.
5. Bishop Christopher M. Neural networks for pattern recognition. – New York: Oxford University Press Inc., 1996. – 482 p.
6. Haykin S. Simon Neural networks and learning machines. – New Jersey: Prentice Hall, 2009. – 936 p
7. Adadurov A., Lazurik V., Rogov Yu., Tokarevskii V., Shechenko S. Spectrometry of Intense Fluxes of Gamma Radiation by Means of the Method of Capsule-Absorbers // IEEE Nuclear Science Symposium and Medical Imaging Conference, Anaheim, California, Abstract Book, – 1996. – p. 17.

Надійшла 30.03.2010.

© А. Ю. Баев, В. Т. Лазурик, 2010