УДК 519.6

ЧИСЛЕННЫЕ МЕТОДЫ НУЛЕВОГО ПОРЯДКА ПОИСКА БЕЗУСЛОВНОГО ЭКСТРЕМУМА

А. А. АРХИПЧУК
Научный руководитель Л. И. СОТСКАЯ, канд. физ.-мат. наук, доц.
Белорусско-Российский университет
Могилев, Беларусь

В практической деятельности часто возникает необходимость оптимизировать какой-либо процесс, реализация данной задачи может быть исполнена благодаря математическим методам. Реальные прикладные задачи оптимизации очень сложны. В связи с этим необходимы новые математические модели и методы развития аппарата оптимизации. Следует склоняться к таким методам, которыми проще управлять в процессе решения задач. Некоторые из этих методов рассмотрены нами в данной работе.

Первый метод, который рассмотрен нами, – это метод золотого сечения [1]. Данный алгоритм относится к численным методам поиска экстремума некоторой гладкой функции на заданном отрезке. Сам метод основан на разбиении отрезка по принципу золотого сечения.

Алгоритм метода состоит в следующем.

1. Задается начальный интервал $L_0 = [a_0, b_0]$ и точность eps, с которой требуется найти экстремум, а также задается значение золотого сечения $\left(gold\_val = \frac{1+\sqrt{5}}{2} \cong 1{,}618 \right)$.

2. Находятся две промежуточные точки на заданном отрезке:

$$x_1 = b_0 - \frac{b_0 - a_0}{gold\_val};\ x_2 = a_0 + \frac{b_0 - a_0}{gold\_val}. \quad (1)$$

3. В полученных точках вычисляются значения функции $f(x_1)$, $f(x_2)$ и сравниваются.

4. Пока длина текущего отрезка больше eps, выполняются следующие шаги в зависимости от результатов сравнения $f(x_1)$, $f(x_2)$:

а) если $f(x_1) > f(x_2)$, то $a_0 = x_1$, $x_1 = x_2$, $f(x_1) = f(x_2)$, далее вычисляем новую точку x_2, используя формулу (1), и находим значение функции в данной точке;

б) если $f(x_1) < f(x_2)$, то $b_0 = x_2$, $x_2 = x_1$, $f(x_2) = f(x_1)$, далее вычисляем новую точку x_1, используя формулу (1), и находим значение функции в данной точке.

5. Шаги 2–4 повторяются до тех пор, пока длина отрезка не станет меньше заданной точности, после чего выводится значение минимума:

$$x = \frac{a_0 + b_0}{2},\ y = f(x).$$

Пример – Для функции $f(x) = 2x^2 - 12x$ найти экстремум.

Зададим начальный интервал $\left[a_0, b_0\right]$ и eps = 0,0001; используя алгоритм, представленный выше, найдем экстремум.

На рис. 1 показан график функции и поиск экстремума по шагам.

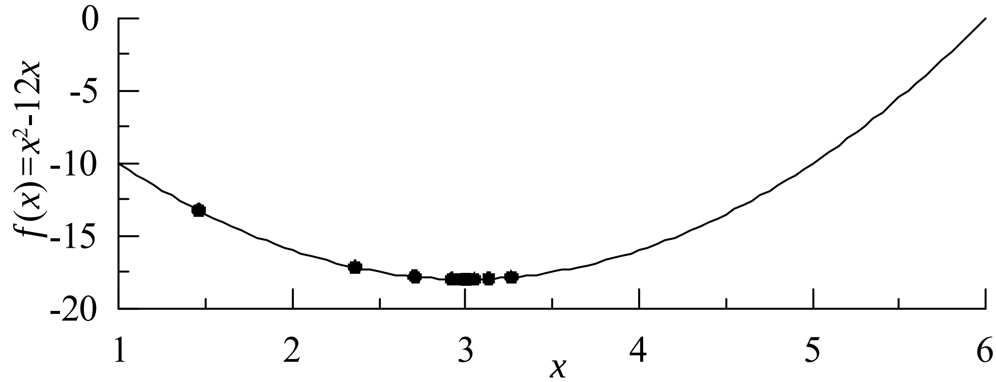

Рис. 1. Зависимость значений функции от аргумента (линия – график функции $f(x) = 2x^2 - 12x$, точки – значения функции в точках приближения, вычисленных по принципу золотого сечения)

Метод Фибоначчи – метод одномерной минимизации, работающий по принципу последовательного сокращения интервала неопределенности [1]. Данный метод рассмотрен нами на примере применения в нейронных сетях. Обычно в нейронных сетях используются методы первого порядка, такие как градиентный спуск, но и метод Фибоначчи нашел свое применение в данном разделе. Метод Фибоначчи основывается на ряде Фибоначчи, и каждый член данного ряда будет использован как некоторое количество эпох. Так как используется понятие «нейронная сеть», то поясним некоторые термины для дальнейшего понимания задачи.

Термин «эпоха» обозначает ситуацию, когда весь набор данных прошел через нейронную сеть в прямом и обратном направлении только один раз [2]. Может возникнуть вопрос: почему одной эпохи недостаточно? Связано это с тем, что используется неполный набор данных. Поэтому обновления весов после одного прохождения недостаточно. Чтобы оптимизировать обучение и подстроить кривую под данные, необходимо использовать градиентный спуск. Точность вычислений можно определить как долю правильных ответов, полученных моделью на тестовом наборе данных.

Метод Фибоначчи использован нами для поиска оптимального количества эпох, т. к. одна эпоха приводит к «недообучению», а избыток эпох – к «переобучению» и некорректной работе нашей нейронной сети.

Как было изложено выше, каждый член ряда Фибоначчи – это некоторое значение эпох, на основе которого мы обучаем нашу нейронную сеть. Возникает вопрос: как мы поймем, какое минимальное количество эпох будет необходимо? Ответ заключается в точности. Как только достигается лучшая точность, не превышающая 1, так будет выявлено оптимальное количество эпох.

Разберем пример классификации рукописных цифр из библиотеки mnist. Сначала подготавливаем данные, т. е., используя сторонние библиотеки, загружаем в наш код данные, после делим их на тренировочные, тестовые выборки и каждый элемент данных выборок переводим в векторы значений такого вида [0, 0, 0, 0, 0, 0, ..., 255, 255, 255, 255, 255, 0, 0, 0, 0, 0, ..., 0, 0, 0].

Далее составляем модель нашей нейронки и компилируем ее:

```
model = keras.Sequential([
keras.layers.Dense(64, activation='relu'),
keras.layers.Dense(10, activation='softmax')])
model.compile(optimizer='adam', loss='sparse_categorical_crossentropy',
metrics=['accuracy'])
```

Используя цикл, подставляем элемент ряда Фибоначчи как некоторое количество эпох в модель, после чего запускается обучение, и, пройдя по всему ряду, находим оптимальное количество эпох с лучшей точностью.

На рис. 2 показан график, из которого видно, что лучшая точность достигается при прохождении 55 эпох. И теперь полученное значение можно использовать как в данной нейронной сети, сократив время ожидания ее обучения, так и в другой подобной нейронной сети.

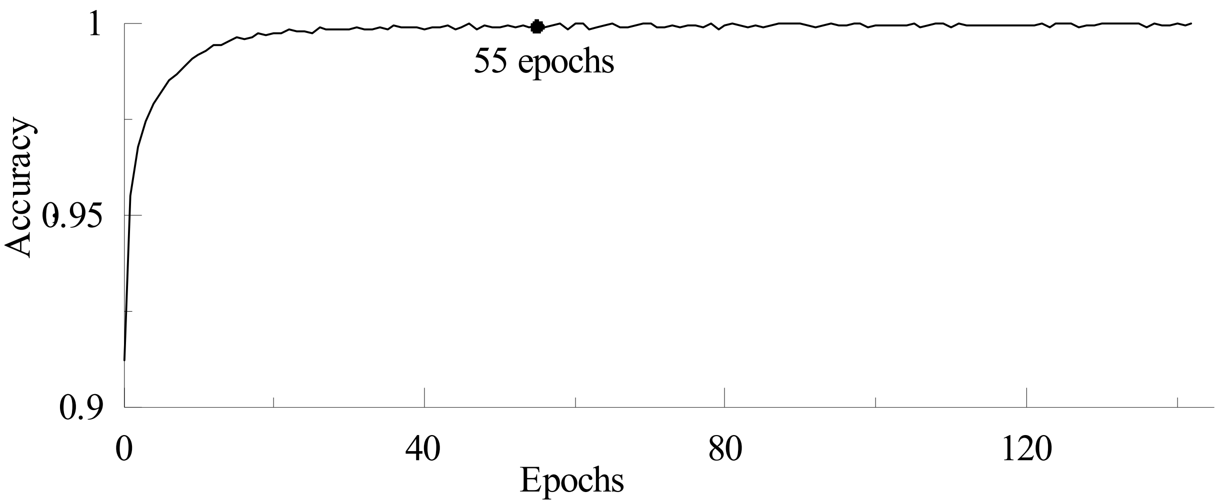

Рис. 2. Зависимость правильной классификации рукописных цифр от количества эпох, вычисленная с использованием метода Фибоначчи

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. **Пантелеев, А. В.** Методы оптимизации в примерах и задачах / А. В. Пантелеев, Т. А. Летова. – Москва: Высшая школа, 2008. – 545 с.

2. **Рашид, Тарик.** Создаем нейронную сеть / Тарик Рашид. – Санкт-Петербург: Альфа-книга, 2017. – 274 с.