О РЕШЕНИИ ОДНОЙ СИСТЕМЫ РАЗНОСТНО-ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ И ЕЕ ПРИМЕНЕНИИ

Н.В. Кокош (г. Минск, БЕЛАРУСЬ), Л.П. Махнист (г. Брест, БЕЛАРУСЬ)

В работе рассматривается однослойная нейронная сеть, состоящая из n нейронных элементов распределительного слоя и m - выходного слоя.

Задача обучения нейронной сети с фиксированной нелинейной гладкой функцией активации F состоит в нахождении весовых коэффициентов w_{ij} $\left(i=\overline{1,n},\ j=\overline{1,m}\right)$ и порогов нейронных элементов T_j $\left(j=\overline{1,m}\right)$, которые минимизируют квадратичную ошибку сети $E_S=\dfrac{1}{2}\sum\limits_{k=1}^{L}\sum\limits_{j=1}^{m}\left(y_j^k-t_j^k\right)^2$, как функции отклонения выходных значений сети $y_j^k=F\left(\sum\limits_{i=1}^{n}w_{ij}x_i^k-T_j\right)$ от t_j^k – эталонных значений j-ого выходного нейрона сети для k-ого образа, определяемого вектором $\overline{x^k}=\left(x_1^k,\dots,x_n^k\right)$ $\left(k=\overline{1,L}\right)$ [1]. Для нахождения такого решения $\overline{W}=\left(w_{11},w_{21},\dots,w_{n1},T_1,w_{12},w_{22},\dots,w_{n2},T_2,\dots,w_{1m},w_{2m},\dots,w_{nm},T_m\right)^T$ можно применять различные градиентные методы, например, метод сопряженных градиентов и его модификации. Учитывая, что в соответствии с идеей метода сопряженных градиентов на каждой итерации t алгоритма модификация вектора $\overline{W}$ производится по формуле

$$\overline{W}(t+1)=\overline{W}(t)-\alpha(t)\cdot\nabla E_S(t)+\beta(t)\cdot\overline{\Delta W}(t),$$

где $\overline{\Delta W}(t)=\overline{W}(t)-\overline{W}(t-1)$, $\nabla E_S(t)$ – вектор градиента функции E_S, получим:

Теорема. Величины квазиоптимальных параметров $\alpha(t)$ и $\beta(t)$ шага обучения нейронной сети с использованием метода сопряженных градиентов в момент времени t определяется соотношениями:

$$\alpha(t)=\frac{\left\|\nabla E_S\right\|^2\cdot\left(\nabla^2E_S\cdot\overline{\Delta W},\ \overline{\Delta W}\right)-\left(\nabla E_S,\ \overline{\Delta W}\right)\cdot\left(\nabla^2E_S\cdot\nabla E_S,\ \overline{\Delta W}\right)}{\left(\nabla^2E_S\cdot\nabla E_S,\ \nabla E_S\right)\cdot\left(\nabla^2E_S\cdot\overline{\Delta W},\ \overline{\Delta W}\right)-\left(\nabla^2E_S\cdot\nabla E_S,\ \overline{\Delta W}\right)^2},$$

$$\beta(t)=\frac{\left\|\nabla E_S\right\|^2\cdot\left(\nabla^2E_S\cdot\nabla E_S,\ \overline{\Delta W}\right)-\left(\nabla E_S(t),\ \overline{\Delta W}\right)\cdot\left(\nabla^2E_S\cdot\nabla E_S),\ \nabla E_S\right)}{\left(\nabla^2E_S\cdot\nabla E_S,\ \nabla E_S\right)\cdot\left(\nabla^2E_S\cdot\overline{\Delta W},\ \overline{\Delta W}\right)-\left(\nabla^2E_S\cdot\nabla E_S,\ \overline{\Delta W}\right)^2},$$

где ∇^2E_S – матрица Гессе вторых производных функции E_S.

Заметим, что вектор $\overline{W}(0)$ можно получить, например, случайной инициализацией, а $\overline{W}(1)$ может быть вычислен на основании соотношений для шага обучения нейронной сети с использованием метода наискорейшего спуска [1].

Литература. 1. Гладкий И.И., Головко В.А., Махнист Л.П. Обучение нейронных сетей с использованием метода наискорейшего спуска // Вестник Брестского государственного технического университета. – Брест: БГТУ, 2001. – № 5: Физика, математика, химия.– С. 47-55.

О ПРИМЕНЕНИИ МЕТОДА ПРЯМЫХ ДЛЯ РЕШЕНИЯ НЕЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ ЗАДАЧ В ЧАСТНЫХ ПРОИЗВОДНЫХ

А.В. Кот (г.Брест, БЕЛАРУСЬ)

Пусть в прямоугольной области $G\{\alpha_x \le x \le \beta_x; \alpha_y \le y \le \beta_y\}$ необходимо найти решение эллиптического дифференциального уравнения

$$A(x,y)\frac{\partial^2 u}{\partial x^2} + B(x,y)\frac{\partial^2 u}{\partial y^2} + C(x,y)\frac{\partial u}{\partial x} + D(x,y)\frac{\partial u}{\partial y} + E(x,y)u = F(x,y), \quad (1)$$

удовлетворяющего граничным условиям

$$u(x,\alpha_y) = \varphi_0(x),\ u(x,\beta_y) = \varphi_1(x)\ (\alpha_x \le x \le \beta_x),$$
$$u(\alpha_x,y) = \psi_0(y),\ u(\beta_x,y) = \psi_1(y)\ (\alpha_y \le y \le \beta_y), \quad (2)$$

где $\varphi_i(x)$, $\psi_i(y)$ $(i=0,1)$ — заданные функции.

Дискретизация задачи (1) – (2) по x приводит нас к методу прямых. В итоге мы получаем систему n обыкновенных нелинейных дифференциальных уравнений второго порядка

$$A_k(x)U_k'' + B_k(x)\sum_{j=0}^{m} c_{kj}^2 U_{k+j-l}(x) + C_k(x)U_k'(x) +$$
$$+ D_k(x)\sum_{j=0}^{m} c_{kj}^1 U_{k+j-l}(x) + E_k(x)U_k(x) = F_k(x)\ (k=1,2,\ldots,n) \quad (3)$$

с условиями

$$U_k(\alpha_x) = \psi_0(y_k),\ U_k(\beta_x) = \psi_1(y_k)\ (k=1,2,\ldots,n). \quad (4)$$

Через $U_k(x)$ здесь обозначены приближенные значения решения $u(x,y)$ на прямой $y = y_k$.

Система (3) обыкновенных дифференциальных уравнений с граничными условиями (4) аппроксимирует дифференциальную задачу (1) – (2). Эту систему можно рассматривать как совокупность обыкновенных дифференциальных уравнений (3) (для каждого k получается отдельное уравнение) с краевыми условиями (4).

Для решения каждого из этих уравнений применяем метод, который, как доказано в [1], является устойчивым.