Формулы (2) и (3) дают возможность определить ожидаемые выходы нейронов предыдущего слоя $t_i = x_i$. По формуле (1) производится изменение весовых коэффициентов и порогов предыдущего слоя.

Получены простые формулы (1)–(3) для рассматриваемой архитектуры сети.

Список литературы

1. Maniakov, N. Traing algorithm for forecasting multilayer neural network / N. Maniakov, L. Makhnist, V. Rubanov // Pattern Recognition and Information Processing : Proc. of The Seventh Intern. Conf. (PRIP'2003), Minsk, Republic of Belarus, 21–23 May 2003 : in 2 vol. – Minsk, 2003. – Vol. 1. – P. 26–30.
2. Makhnist, L. Some Methods of Adaptive Multilayer Neural Network Training / L. Makhnist, N. Maniakov // International Journal of Computing. – 2004. – Vol. 3. – P. 99–106.
3. Maxnist, L. Convergence analysis of neural networks training based on steepest descent method / L. Maxnist, A. Doudkin, V. Golovko // Pattern Recognition and Information Processing (PRIP'2007) : Proc. of the Ninth Intern. Conf., Minsk, Republic of Belarus, 22–24 May 2007 : in 2 vol. – Minsk, 2007. – Vol. 1. – P. 285–289.
4. Махнист, Л. П. Использование алгоритмов обучения однослойной сети для многослойных нейронных сетей прямого распространения / Л. П. Махнист, В. А. Головко, И. И. Гладкий // Вестн. Брест. гос. техн. ун-та. Сер.: Физика, математика, информатика. – 2020. – № 5. – С. 32–37.

УДК 519.6+517.983

ВЫБОР МОМЕНТА ОСТАНОВА В ИТЕРАЦИОННОМ МЕТОДЕ ЯВНОГО ТИПА РЕШЕНИЯ НЕКОРРЕКТНЫХ УРАВНЕНИЙ ПЕРВОГО РОДА

А. В. Жорох,

БрГУ имени А. С. Пушкина, Брест

Научный руководитель:

О. В. Матысик, кандидат физико-математических наук, доцент

В действительном гильбертовом пространстве H решается линейное уравнение первого рода $Ax = y$, где $A: H \to H$ — положительный ограниченный самосопряженный оператор ($0 \in SpA$, и, следовательно, рассматриваемая задача некорректна). Решать данную задачу будем при помощи *итерационного метода явного типа*

$$x_{n+1} = x_n + \alpha A(y - Ax_n), \quad x_0 = 0. \tag{1}$$

В случае приближенной правой части y_δ $\left(\|y - y_\delta\| \le \delta\right)$ соответствующие процедуре (1) итерации примут вид

$$x_{n+1,\delta} = x_{n,\delta} + \alpha A(y_\delta - Ax_{n,\delta}), \;\; x_{0,\delta} = 0. \qquad (2)$$

Под сходимостью метода (2) понимается утверждение о том, что приближения (2) сколь угодно близко подходят к точному решению уравнения $Ax = y$ при подходящем выборе n и достаточно малых δ.

Рассмотрим сходимость метода (1) при точной правой части y операторного уравнения $Ax = y$. Нетрудно показать по индукции, что $x_n = A^{-1}\left[E - \left(E - \alpha A^2\right)^n\right]y.$ Тогда $x - x_n = A^{-1}\left(E - \alpha A^2\right)^n y.$ Используя [1–2] интегральное представление самосопряженного оператора $A = \int\limits_0^M \lambda\, dE_\lambda$ ($M = \|A\|, E_\lambda$ – соответствующая оператору A спектральная функция), получим $x - x_n = \int\limits_0^M \lambda^{-1}\left(1 - \alpha\lambda^2\right)^n dE_\lambda y$. Для сходимости метода (1) потребуем, чтобы $\left|1 - \alpha\lambda^2\right| < 1, \;\; \lambda \in (0, M]$. Отсюда

$$0 < \alpha < \frac{2}{M^2}. \qquad (3)$$

Разобьем выписанный интеграл на два интеграла

$$x - x_n = \int\limits_0^\varepsilon \lambda^{-1}\left(1 - \alpha\lambda^2\right)^n dE_\lambda y + \int\limits_\varepsilon^M \lambda^{-1}\left(1 - \alpha\lambda^2\right)^n dE_\lambda y.$$

При условии (3) $\left|1 - \alpha\lambda^2\right| \le q < 1$ для $\lambda \in [\varepsilon, M]$, поэтому

$$\left\|\int\limits_\varepsilon^M \lambda^{-1}\left(1 - \alpha\lambda^2\right)^n dE_\lambda y\right\| \le q^n \left\|\int\limits_\varepsilon^M \lambda^{-1} dE_\lambda y\right\| = q^n \left\|\int\limits_\varepsilon^M dE_\lambda x\right\| \to 0, \;\; n \to \infty.$$

Для первого интеграла

$$\left\|\int\limits_0^\varepsilon \lambda^{-1}\left(1 - \alpha\lambda^2\right)^n dE_\lambda y\right\| \le \left\|\int\limits_0^\varepsilon \lambda^{-1} dE_\lambda y\right\| = \left\|\int\limits_0^\varepsilon dE_\lambda x\right\| = \|E_\varepsilon x\| \to 0,$$

так как E_ε сильно стремится к нулю при $\varepsilon \to 0$ [1–2].

Таким образом, $\|x - x_n\| \to 0, \;\; n \to \infty$ и, значит, сходимость метода (1) к точному решению x для случая $x_0 = 0$ при условии (3) доказана, т.е. доказана

Теорема 1. *При условии* (3) *итерационный процесс* (1) *сходится в исходной норме гильбертова пространства*.

Скорость убывания к нулю $\|x - x_n\|$ неизвестна и может быть сколь угодно малой. Для ее оценки предположим, что решение x истокопредставимо, т.е. $x = A^s z,\ s > 0$. Значит, $y = A^{s+1} z$ и $x - x_n = \int\limits_0^M \lambda^{-1}\left(1 - \alpha\lambda^2\right)^n dE_\lambda y = \int\limits_0^M \lambda^s \left(1 - \alpha\lambda^2\right)^n dE_\lambda z.$

Чтобы получить оценку для $\|x - x_n\|$, оценим максимум модуля подынтегральной функции $f(\lambda) = \lambda^s\left(1 - \alpha\lambda^2\right)^n$. Для этого $f'(\lambda)$ приравняем к нулю $f'(\lambda) = \lambda^{s-1}\left(1 - \alpha\lambda^2\right)^{n-1}\left[s - (2n+s)\alpha\lambda^2\right] = 0$. Отсюда видно, что производная обращается в нуль при равенстве нулю любого из трех сомножителей. Но при обращении в нуль первых двух $f(\lambda)$ тоже обращается в нуль. Поэтому остается $s - (2n+s)\alpha\lambda^2 = 0$, откуда $\lambda^* = \left[\frac{s}{(2n+s)\alpha}\right]^{\frac{1}{2}}$ – стационарная точка. Она является точкой локального максимума, так как $f''(\lambda^*) = -\left[\frac{s}{(2n+s)\alpha}\right]^s \left(\frac{2n}{2n+s}\right)^{n-1} 2(2n+s)\alpha^2 < 0$.

Оценим $f(\lambda^*) = \left|f(\lambda^*)\right|$:

$$f(\lambda^*) = \lambda^s\left(1 - \alpha\lambda^2\right)^n \big|_{\lambda=\lambda^*} = s^{\frac{s}{2}} \alpha^{-\frac{s}{2}} (2n+s)^{-\frac{s}{2}} \left(\frac{2n}{2n+s}\right)^n =$$

$$= s^{\frac{s}{2}} \alpha^{-\frac{s}{2}} (2n)^{-\frac{s}{2}} \left(\frac{2n+s}{2n}\right)^{-n-\frac{s}{2}} = s^{\frac{s}{2}} \alpha^{-\frac{s}{2}} (2n)^{-\frac{s}{2}} \left[\left(1 + \frac{s}{2n}\right)^{\frac{2n}{s}}\right]^{\frac{s}{2n}\left(-n-\frac{s}{2}\right)} <$$

$$s^{\frac{s}{2}} (2n\alpha e)^{-\frac{s}{2}}.$$

Нетрудно показать, что последнее значение не меньше максимального значения модуля $|f|$ для всех α, удовлетворяющих условию

$$0 < \alpha \le \frac{5}{4M^2}. \qquad (4)$$

Таким образом, при α, удовлетворяющему условию (4), справедлива следующая оценка $\max\limits_{\lambda\in[0,M]} |f(\lambda)| < s^{s/2}(2n\alpha e)^{-s/2}$, следовательно, $\|x - x_n\| < s^{s/2}(2n\alpha e)^{-s/2}\|z\|$.

Оценим $\|x_n - x_{n,\delta}\|$, где $x_n - x_{n,\delta} = \int\limits_0^M \lambda^{-1}\left[1 - \left(1 - \alpha\lambda^2\right)^n\right] dE_\lambda (y - y_\delta)$. Для этого оценим подынтегральную функцию (она положительна) $g_n(\lambda) = \lambda^{-1}\left[1 - \left(1 - \alpha\lambda^2\right)^n\right]$. По индукции нетрудно доказать, что для $n \ge 2$ при условии (4) справедлива оценка $g_n(\lambda) \le 2n^{\frac{1}{2}}\alpha^{\frac{1}{2}}$. Общая оценка погрешности метода (2) для истокопредставимого решения при условии (4) имеет вид

$$\left\|x - x_{n,\delta}\right\| \le s^{\frac{s}{2}}\left(2n\alpha e\right)^{-\frac{s}{2}}\|z\| + 2n^{\frac{1}{2}}\alpha^{\frac{1}{2}}\delta,\ n \ge 2. \qquad (5)$$

Итак, доказана

Теорема 2. *Итерационный процесс* (2) *при условии* (4) *сходится, если выбрать число итераций* n *из условия* $n^{\frac{1}{2}}\delta \to 0$, $n \to \infty$, $\delta \to 0$. *Для процесса* (2) *при условии* (4) *и* $x = A^s z,\ s > 0$, *справедлива оценка погрешности* (5).

Оптимизируем оценку (5). Для этого при заданном δ найдем такое $n_{\text{опт}}$, при котором оценка погрешности будет минимальной. Приравняв нулю производную по n от правой части неравенства (5), получим априорный момент останова:
$n_{\text{опт}} = s^{\frac{s+2}{s+1}} 2^{-\frac{s+2}{s+1}} \alpha^{-1} e^{-\frac{s}{s+1}} \|z\|^{\frac{2}{s+1}} \delta^{-\frac{2}{s+1}}$.

Список литературы

1. Матысик, О.В. Явные и неявные итерационные процедуры решения некорректно поставленных задач / О.В. Матысик. – Брест: Брест. гос. ун-т, 2014. – 213 с.

2. Матысик, О.В. Итерационная регуляризация некорректных задач / О.В. Матысик. – Saarbrücken : LAP LAMBERT Acad. Publ., 2015. – 188 с.