ОБУЧЕНИЕ НАСТРОЙЩИКА С НЕЙРОННОЙ СЕТИ НА ОСНОВЕ ЭКСТРЕМАЛЬНОГО МЕТОДА ОБУЧЕНИЯ

Мамиров Уктам Фарходович
Ташкентский государственный технический университет, доктор технических наук, доцент
Холикулова Муслима Примберди кизи
Наманганский инженерно-технологический институт, магистрант

Обучение нейронной сети предлагается вести в два этапа: первый этап – это этап первоначальной инициализации нейронной сети. Его целью является обучение нейронной сети на текущие коэффициенты ПИ-регулятора, с которыми функционирует объект управления в момент установки настройщика. Второй этап – обучение нейронной сети в оперативном режиме.

Для функционирования метода обратного распространения ошибки необходимо произвести первоначальную инициализацию нейронной сети, представляющую собой установку первоначальных значений весовых коэффициентов и смещений. Так как нейронная сеть имеет два выхода, отвечающие за K_n и K_u соответственно, то после первоначальной инициализации, при подаче на входы сети текущих значений, на выходе сети должны появиться значения коэффициентов регулятора, используемых в данный момент системой управления. Невыполнение этого условия приведет к резкому ухудшению качества управления на этапе инициализации нейросетевого настройщика. Для решения этой задачи необходимо произвести инициализацию нейронной сети с помощью экстремального метода обучения [1-3]. Для использования данного метода обучения, в соответствии с выбранной структурой нейронной сети (в особенности количества нейронов в скрытом слое), необходимо получить 13 наборов входных значений. Накопив данный массив информации, произвести обучение экстремальным методом.

Метод экстремального обучения (ELM) позволяет без итерационной процедуры обучить нейронную сеть. Обучение происходит за счет подачи 13 заранее подготовленных наборов входных значений, для каждого из которых задана пара векторов $(P\ Y_j\ ,\ _j\ )$, j =1,..., N, где P_j – j-й входной вектор, а Y_j – j-й выходной вектор, N–количество наборов входных значений.

Входные векторы образуют матрицу входных векторов P нейронной сети за 13 тактов ее работы, имеющую размерность N*m (13*4), где m–количество входных нейронов:

$$\mathrm{P} = \begin{bmatrix} \mathrm{P}_1 \\ \Lambda \\ \mathrm{P}_\mathrm{N} \end{bmatrix} = \begin{bmatrix} P_{11} & \Lambda & P_{1m} \\ \Lambda & \Lambda & \Lambda \\ P_{N1} & \Lambda & P_{Nm} \end{bmatrix} \ N \cdot m .$$

Выходные вектора образуют матрицу выходных значений Y нейронной сети, имеющую размерность N*k (13*2), *k*– количество выходных нейронов. Эта матрица состоит из 13 одинаковых векторов по два элемента: текущее значение *Кn* и *Кu*.

$$Y = \begin{bmatrix} Y_1 \\ \Lambda \\ Y_N \end{bmatrix} = \begin{bmatrix} Y_{11} & \Lambda & Y_{1k} \\ \Lambda & \Lambda & \Lambda \\ Y_{N1} & \Lambda & Y_{Nm} \end{bmatrix} L \cdot k\,.$$

Матрица входных весов нейронов скрытого слоя w представлена следующим образом:

$$W = \begin{bmatrix} W_1 \\ \Lambda \\ W_L \end{bmatrix} = \begin{bmatrix} W_{11} & \Lambda & W_{1m} \\ \Lambda & \Lambda & \Lambda \\ W_{L1} & \Lambda & W_{Lm} \end{bmatrix} L \cdot m,$$

где W_i – вектор входных весов i – го нейрона скрытого слоя, $i = 1,\dots,L, L$ – количество нейронов скрытого слоя.

В случае применения метода экстремального обучения входные веса W_{ij} и смещение b_i для каждого i – го нейрона скрытого слоя задаются с помощью генератора случайных величин: $W_{ij} = random$ (1...1), b_i =*random*(0...1).

Матрица выходных весов нейронов скрытого слоя, имеющая размерность L*k(13*2):

$$\beta = \begin{bmatrix} \beta_1 \\ \Lambda \\ \beta_L \end{bmatrix} = \begin{bmatrix} \beta_{11} & \Lambda & \beta_{1m} \\ \Lambda & \Lambda & \Lambda \\ \beta_{L1} & \Lambda & \beta_{Lm} \end{bmatrix},$$

где β_i – вектор выходных весов i-го нейрона скрытого слоя, $i = 1,\dots,L$.

Выходной вектор сети Y_{ij} вычисляется как

$$Y_j = \sum_{i=1}^{L} \beta_i F\left(W_i P_j^T + b_i\right), \quad j = 1,2,\dots,N, \tag{1}$$

В матричном виде расчет будет иметь вид:

$$Y = H\beta, \tag{2}$$

где *H* – матрица выходных значений нейронов скрытого слоя, имеющая размерность *N*x*L*, *N* – количество наборов входных значений, *L* – количество нейронов скрытого слоя:

$$\mathrm{H} = \begin{bmatrix} F\left(w_1 P_1 + b_1\right) & \Lambda & F\left(w_L P_1 + b_L\right) \\ \Lambda & \Lambda & \Lambda \\ F\left(w_1 P_N + b_1\right) & \Lambda & F\left(w_L P_N + b_L\right) \end{bmatrix} N \cdot L\,,$$

где функция F является игмоидальной:

$$F\left(w_1 P_1 + b_1\right) = \frac{1}{1 + \exp\left(W_i P_j + b_i\right)}.$$

Обучение сети заключается в вычислении матрицы весов между нейронами скрытого и выходного слоев β_i по формуле:

$$\beta = H^{+}Y, \tag{3}$$

где $H^{+} = \left(H^{T}H\right)^{-1}H^{T}$ – псевдообратная матрица для случая невырожденной матрицы H.

По теории матриц [2] известно, что выражение (3) является наилучшим приближенным решением уравнения (2) по методу наименьших квадратов, что позволяет избегать локальных экстремумов при обучении сети.

Когда произведено обучение нейронной сети, то есть по формуле (3) получена матрица весовых коэффициентов выходного слоя, β при подаче на вход сети вектора входных значений P, на выходе вычисляется первоначальное значение настроек регулятора. После первоначальной инициализации нейросетевой настройщик переходит в оперативный режим работы, и метод экстремального обучения в дальнейшем не используется.

Вместе с тем здесь при реализации алгоритма стабилизации вида (3) используется псевдообратная матрица H^{+}. Известно [4-7], что задача вычисления псевдообратной матрицы в общем случае неустойчива по отношению к погрешностям в задании матрицы. На практике к рассмотрению таких возмущений может побуждать ограниченная точность, с которой наблюдаемое явление описывается количественной информацией. Влияние погрешностей округлений, производимых в ходе численной процедуры, тоже можно проанализировать так, как если бы оно имело причиной возмущения входных данных.

Условия аппроксимации уравнения (3) примем в виде

$$\left\|H - \overline{H}\right\| \le h, \quad \left\|Y - \overline{Y}\right\| \le \delta, \tag{4}$$

где $\overline{H}$ и $\overline{Y}$ – точные значения матричного оператора и правой части уравнения (3) соответственно.

Известно [4], что нормальное псевдорешение $\beta = H^{+}Y$ является нормальным решением нормальной системы уравнений

$$H^{T}H\beta = H^{T}Y \tag{5}$$

или $H^{T}r = 0$, где $r = Y - H\beta$.

При решении уравнения (3) обычно используется метод регуляризации А.Н.Тихонова [6]:

$$(H^{T}H + \alpha I)\beta = H^{T}Y. \tag{6}$$

Как показано в [5] регуляризованную нормальную систему уравнений (6) можно записать в виде:

$$\left(\begin{array}{c|c} \omega I_m & H \\ \hline H^{T} & -\omega I_n \end{array}\right)\begin{pmatrix} y \\ \beta \end{pmatrix} = \begin{pmatrix} Y \\ 0 \end{pmatrix} \Leftrightarrow \tilde{H}_{\omega} z = \tilde{\beta}, \tag{7}$$

где $y = \omega^{-1} r$, $\omega = \sqrt{\alpha}$.

Матрица $\widetilde{H}_{\omega}$ системы (7) при всех $\alpha > 0$ невырождена и её единственным решением является вектор $z_* = \left(y_*^T,\ \beta_*^T\right)^T$, где $\beta = \left(H^T H + \alpha I_n\right)^{-1} H^T \beta$, $y_* = \omega^{-1} r_*$.

Приведенные алгоритмы позволяют получать устойчивое решение задачи определения весов нейронов скрытого слоя и использовать их при решении задач идентификации и управления линейным неопределенным динамическим объектам. На основе известных теоретических положений теории регулярного оценивания можно утверждать, что использование рандомизированного и квазиоптимального вариантов регуляризованной модификации проекционного алгоритма позволяет увеличить скорость сходимости и получить искомое решение с заранее необходимой точностью. При этом эффективность этих алгоритмов, как известно, в значительной степени повышается при уменьшении спектрального числа обусловленности регуляризованной расширенной системы вида (7), что обусловливает повышенную точность используемого алгоритма.

ЛИТЕРАТУРА:

1. Omatu S., Khalid M., Yusof R. Neuro–Control and its Applications. – London: Springer, 1995. – 255 p.

2. Терехов В.А., Ефимов Д.В., Тюкин И.Ю. Нейросетевые системы управления. Москва, Высшая школа. 2002. –183 с.

3. Алиев Р.А., Алиев Р.Р. Теория интеллектуальных систем и ее применение. –Баку: Чашыоглы, 2001 -720 с.

4. Igamberdiev, H., Yusupbekov, A., Mamirov, U., Abdukaxxarov, I. (2022). Stable Algorithms for Solving the Problem of Determining the Weighting Coefficients of Neural Networks with Radial-Basis Activation Functions. In: Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M., Sadikoglu, F.M. (eds) 11th International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions and Artificial Intelligence - ICSCCW-2021. ICSCCW 2021. Lecture Notes in Networks and Systems, vol 362. Springer, Cham. https://doi.org/10.1007/978-3-030-92127-9_87.

5. Yusupbekov, A.N., Sevinov, J.U., Botirov, T.V., Mamirov, U.F.: Algorithms for the Synthesis of a Neural Network Regulator for Control of Dynamic Objects. In: Aliev R.A., Kacprzyk J., Pedrycz W., Jamshidi M., Babanli M., Sadikoglu F.M. (eds) 14th International Conference on Theory and Application of Fuzzy Systems and Soft Computing. Advances in Intelligent Systems and Computing. vol. 1306. 723-730 (2020) Springer, Cham. https://doi.org/10.1007/978-3-030-64058-3_90.

6. Yusupbekov N.R., Igamberdiev H.Z., Zaripov O.O., Mamirov U.F. Stable Iterative Neural Network Training Algorithms Based on the Extreme Method In: Aliev

R.A., Kacprzyk J., Pedrycz W., Jamshidi M., Babanli M., Sadikoglu F.M. (eds) 14th International Conference on Theory and Application of Fuzzy Systems and Soft Computing – ICAFS-2020. ICAFS 2020. Advances in Intelligent Systems and Computing, vol 1306. Springer, Cham. https://doi.org/10.1007/978-3-030-64058-3_30.

7. Yusupbekov, N.R., Igamberdiev, H.Z., Mamirov, U.F.: Adaptive Control System with a Multilayer Neural Network under Parametric Uncertainty Condition. In: Russian Advances in Fuzzy Systems and Soft Computing: selected contributions to the 8-th International Conference on Fuzzy Systems, Soft Computing and Intelligent Technologies (FSSCIT-2020), Vol. 2782, pp. 228-234. CEUR Workshop Proceedings, Aachen, Germany. doi: http://ceur-ws.org/Vol-2782/paper_32.pdf