Секция 2. Моделирование биосферных процессов

Моделирование биосферных процессов на основе искусственных нейронных сетей

Арзамасцев А. А., Козадаев А. С.

Тамбовский государственный университет им. Г. Р. Державина, ул. Интернациональная 33, Тамбов, 392000, Россия

e-mail: `arz_sci@mail.ru`, `akozadaev@inbox.ru`

В докладе рассматривается моделирование биосферных процессов с помощью аппарата искусственных нейронных сетей: поиск структуры сети, выбор оптимального числа входов.

Для моделирования и прогнозирования биосферных процессов часто используются детерминированные модели, методы математической статистики и анализа временных рядов (ВР) [2, 3]. В последнее время, в связи с развитием искусственного интеллекта и разработкой специализированных симуляторов, появилась возможность решать эти проблемы с помощью технологий искусственных нейронных сетей (ИНС) [1, 4, 5, 7].

Будем считать, что ВР порождается объектом, представляющий собой «черный ящик» с известным числом входов $x_1,x_2,\ldots,x_n$ и одним выходом — y.

Будем считать также заданными временные изменения $x_1(t)$,$x_2(t)\ldots,x_n(t)$ на отрезке $t\in[t_0,t_k]$. Математическая модель такого объекта в операторной форме обычно имеет вид:

$$y=f(x)=f(x_1,x_2,...,x_n)\,. \tag{1}$$

В простейшем случае свойства каналов такого объекта, представляют собой обычные пропорциональные звенья с передаточными функциями $W_1(p)=k_1,W_2(p)=k_2,\ldots,W_n(p)=k_n$. В других случаях передаточные функции каналов могут быть более сложными. Например, наиболее распространенными из них являются: апериодические звенья первого и второго порядков с передаточными функциями: $W(p)=\frac{k}{Tp+1}$ и $W(p)=\frac{k}{(T_1p+1)(T_2p+1)}$ и звено чистого запаздывания с передаточной функцией $W(p)=e^{-p\tau}$. Здесь T,T_1,T_2,k,τ — некоторые коэффициенты, зависящие от природы объекта. Поэтому, в более общем случае уравнение (1), может быть записано в виде:

$$y = f(X, t) = f(x_1, x_2, ..., x_n, t)\,. \tag{2}$$

Поскольку обычно именно такая система и является порождающей для ВР биосферного процесса, необходимо учитывать ее свойства при его анализе и прогнозировании. Главные из таких свойств, усложняющие прогнозирование значений ряда по его предыстории: инерционность, запаздывание и наличие стохастической составляющей, связанной как с недостатком информации, так и с погрешностями измерений.

При прогнозировании временных рядов, порожденных указанным выше объектом, возможно возникновение следующих ситуаций: а) наблюдаемым является только ряд $\{y\}$; временные ряды $\{x_1\}, \{x_2\}, ..., \{x_n\}$, характеризующие входы объекта не являются наблюдаемыми; в этом случае, необходимо допустить, что вся информация от векторов , преобразованная объектом, содержится в векторе $\{y\}$; назовем задачу оптимального выбора структуры и настроек ИНС для моделирования и прогнозов данного временного ряда задачей 1; б) вектор $\{y\}$ и все векторы $\{x_1\}, \{x_2\}, ..., \{x_n\}$ являются наблюдаемыми; назовем задачу оптимального выбора структуры и настроек ИНС для моделирования и прогнозов данного временного ряда задачей 2; в) наблюдаемыми является вектор $\{y\}$ и часть векторов $\{x_1\}, \{x_2\}, ..., \{x_n\}$; в этом случае, необходимо допустить, что вся информация от ненаблюдаемых входных векторов также содержится в векторе $\{y\}$; назовем задачу оптимального выбора структуры и настроек ИНС для моделирования и прогнозов данного временного ряда задачей 3.

ИНС-модель временного ряда. В соответствии с теоремами А.Н. Колмогорова о представимости функций нескольких переменных с помощью суперпозиций и сумм функций одного переменного, можно утверждать что, каждая непрерывная функция n переменных, заданная на единичном кубе n-мерного пространства, представима в виде:

$$f(x_1, x_2, ..., x_n) = \sum_{q=1}^{2n+1} h_q \left[\sum_{p=1}^{n} \varphi_q^p(x_p) \right], \tag{3}$$

где функции $h_q(u)$ непрерывны, а функции $\varphi_q^p(x_p)$ кроме того, еще и стандартны, т.е. не зависят от выбора функции f. Например, функция 2-х переменных x_1 и x_2 может быть представлена в виде:

$$f(x_1, x_2) = \sum_{q=1}^{5} h_q[f_1(x_1) + f_2(x_2)]\,. \tag{4}$$

С точки зрения схемотехники ИНС уравнение (4) может быть представлено с помощью ИНС-модели, содержащей два входных нейрона, два нейрона первого скрытого слоя (f_1, f_2), пять нейронов второго скрытого слоя $(h_1, h_2, .., h_5)$ и один выходной нейрон.

Число степеней свободы такой ИНС-модели, равно числу синаптических связей сети, т.е. в данном случае составляет 17 (19). Для функции n переменных, число синаптических связей, соответствующее формуле (3) и общему принципу схемотехники ИНС составит:

$$P_1 = n + n(2n+1) + 2n + 1 = 2n^2 + 4n + 1 \tag{5}$$

Очевидно, что число строк в обучающей выборке должно быть не меньше числа степеней свободы. Учитывая, что число строк в обучающей выборке можно выразить как $P_2 = k-n-h+1$, получим, что должно удовлетворяться следующее неравенство:

$$P_1 \leqslant P_2,\ 2n^2 + 4n + 1 \leqslant k - n - h + 1\,. \tag{6}$$

Учитывая также, что должно выполняться $n \geqslant 1$, окончательно получим область допустимых значений для числа входов ИНС
для задачи 1:

$$1 \leqslant n \leqslant \frac{-5 + \sqrt{25 + 8(k-h)}}{4}\,. \tag{7}$$

Обучающая выборка, для произвольного, заданного числа входов ИНС для задачи 1 может быть задана матрицами:

$$X = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ y_2 & y_3 & \cdots & y_{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ y_m & y_{m+1} & \cdots & y_{m+n-1} \end{pmatrix},\ Y = \begin{pmatrix} y_{n+h} \\ y_{n+h+1} \\ \cdots \\ y_{n+h+m-1} \end{pmatrix} \tag{8}$$

При этом структура сети: один слой из n входных нейронов, скрытый слой из n нейронов, реализующих преобразование в соответствии с функциями нейронов $f_1, f_2, ..., f_n$, скрытый слой из $2n+1$ нейронов, реализующих преобразование в соответствии с функциями нейронов h и выходной слой, состоящий из одного нейрона. Для определения оптимального числа входов ИНС для задачи 1 необходимо определить минимум функционала:

$$F(W) = \parallel \Delta \parallel = \parallel Y^{tabl} - Y^{net}(W) \parallel = \sqrt{\sum_{i=1}^{p}\sum_{j=1}^{k}\left(Y_{ij}^{tabl} - Y_{ij}^{net}(W)\right)^2} \to min, \tag{9}$$

где вектор W определяется структурой сети S, которая, в свою очередь связана с числом ее входов n, так что имеется однозначное соответствие n и $W(S)$.

Поэтому, вектор W^*, зависящий от структуры сети и минимизирующий (9), соответствует оптимальной ИНС-модели и может быть определен как:

$$W^* = arg \min_{W \in \Omega} F(W) \qquad (10)$$

В докладе рассмотрена технология формирования оптимальной структуры ИНС предназначенной для решения задач 1-3. Обсуждались примеры моделирования и прогнозирования ВР, описывающих изменение температуры воздуха в г. Тамбове [5] и численности смешанной популяции креветки в Индийском океане [6].

Литература

1. Anil K. Jain, Jianchang Mao, Mohiuddin K.M. *Artificial Neural Networks: A Tutorial*, Computer **29** (1996), no. **3**, 31–44.
2. Бокс Дж., Дженкинс Г. *Анализ временных рядов, прогноз и управление*, М. Мир, 1974.
3. Бриллинждер Д. *Временные ряды*, М.: Мир, 1980.
4. *Искусственный интеллект. Модели и методы.* В кн. 2 / под ред. Д.А. Поспелова — М.: Радио и связь, 1990.
5. Козадаев А. С., Арзамасцев А. А. *Прогнозирование временных рядов с помощью аппарата искусственных нейронных сетей. Краткосрочный прогноз температуры воздуха*, Вестник Тамбовского университета, Сер. Естественные и технические науки **11**, Вып. **3** (2006), 299–304.
6. Арзамасцев А.А., Козадаев А.С. *Прогнозирование численности популяции креветки в открытой системе с помощью искусственных нейронных сетей*, Вестник Тамбовского университета, Сер. Естественные и технические науки **10** Вып. **2** (2005), 187–192.
7. Осовский С. *Нейронные сети для обработки информации*, Пер. с польского И. Д. Рудинского, М.:Финансы и статистика, 2002.