О.П. Солдатова, А.Ю. Скобеев

ПРОГНОЗИРОВАНИЕЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ С ПОМОЩЬЮ РАДИАЛЬНО-БАЗИСНЫХ НЕЙРОННЫХ СЕТЕЙ

(ФГБОУ ВПО «Самарский государственный аэрокосмический университет им. академика С.П. Королева (национальный исследовательский университет)»)

Прогнозирование – это ключевой момент при принятии решений. Конечная эффективность любого решения зависит от последовательности событий, возникающих уже после принятия решения. Возможность предсказать неуправляемые аспекты этих событий перед принятием решения позволяет сделать наилучший выбор, который, в противном случае, мог бы быть не таким удачным.

Существуют различные подходы к прогнозированию, и выбор конкретного метода зависит от имеющейся информации и свойств исследуемого показателя. Прогнозирование на основе временных рядов осуществляется в тех случаях, когда имеются сведения о значениях показателя и предполагается, что в будущем закон его развития сохранится. Таким образом, прогнозирование фактически сводится к задаче экстраполяции.

Нейронная сеть выступает в качестве универсального аппроксиматора обучающих данных, поэтому применение нейронных сетей для прогнозирования является весьма перспективным. На вход сети подаются известные значения ряда, а на выходе получают прогноз на необходимое число шагов. Основной задачей становится выбор архитектуры нейронной сети и обучение на имеющихся данных.

В настоящее время прогнозирование с помощью нейронных сетей является активно развивающейся, но еще недостаточно распространенной областью. Наиболее популярные программы-нейроимитаторы обладают широкими возможностями и являются достаточно универсальными, но не приспособлены для задач прогнозирования. Имеющиеся на рынке специализированные программы, напротив, работают по принципу «черного ящика» и обладают ограниченными возможностями для настройки нейронной сети и визуализации результатов. Поэтому существует потребность в программном обеспечении, ориентированном на задачи прогнозирования, но предоставляющем возможность детальной настройки нейронной сети и проведения исследований.

Нейронные сети на основе радиально-базисных функций, с точки зрения математики, выполняют аппроксимацию путем адаптации нескольких одиночных аппроксимирующих функций к ожидаемым значениям, причем эта адаптация проводится только в локальной области многомерного пространства. При таком подходе отображение всего множества данных представляет собой сумму локальных преобразований, а скрытые нейроны составляют множество базисных функций локального типа [1]. В таких нейронных сетях, скрытые нейроны реализуют функции, радиально изменяющиеся вокруг выбранного центра и принимающие ненулевые значения только в окрестности этого центра $\varphi(x)= \varphi(\|x\text{-}c\|)$, где x– входной вектор,c– вектор координат центра.

Обособленный класс алгоритмов обучения радиальных сетей составляют градиентные алгоритмы обучения с учителем, в которых используется алгоритм обратного распространения ошибки. Их основу составляет целевая функция, которая для одного обучающего примера имеет вид:

$$E = \frac{1}{2}\left[\sum_{i=1}^{K} w_i \varphi_i(x) - d\right]^2 \tag{1}$$

Предположим, что применяется самая общая форма гауссовой радиальной функции $\varphi_i(x)$, соответствующей гипер радиально-базисной сети, в которой:

$$\varphi(x) = \exp\left[-\frac{1}{2}\left[Q_i(x-c_i)\right]^T\left[Q_i(x-c_i)\right]\right] \tag{2}$$

а матрица Q_i имеет произвольную структуру.

Подбор значений параметров можно осуществлять, используя градиентные методы оптимизации независимо от объекта обучения – будь то вес или центр. Независимо от выбираемого метода градиентной оптимизации, необходимо, получить вектор градиента целевой функции относительно всех параметров сети. Для этого необходимо решить систему дифференциальных уравнений, представленную формулами (3) [1]:

$$\begin{cases} \dfrac{\partial E}{\partial Q_{ij}} = 0 \\ \dfrac{\partial E}{\partial w_i} = 0 \\ \dfrac{\partial E}{\partial c_i} = 0 \end{cases} \tag{3}$$

Тогда аналитические выражения для частных производных можно записать в более простом виде:

$$\frac{\partial E}{\partial w_0} = y - d \tag{4}$$

$$\frac{\partial E}{\partial w_i} = \exp\left(-\frac{1}{2}u_i\right)(y-d) \tag{5}$$

$$\frac{\partial E}{\partial c_j^{(i)}} = \hat{v}_j^{(i)} = -\exp\left(-\frac{1}{2}u_i\right) w_i (y-d) \sum_{j=1}^{N} Q_{jr}^{(i)} z_j^{(i)} \tag{6}$$

$$\frac{\partial E}{\partial Q_{jr}^{(i)}} = v_t^{(i)} \hat{z}_t^{(i)} = -\exp\left(-\frac{1}{2}u_i\right) w_i (y-d)\left(x_j - c_j^{(i)}\right) z_r^{(i)} \tag{7}$$

$$z_r^{(i)} = \sum_{j=1}^{N} Q_{jr}^{(i)} (x_j - c_j^{(i)}) \tag{8}$$

$$u_i = \sum_{j=1}^{N} \left[z_j^{(i)}\right]^2 \tag{9}$$

i – индекс нейрона скрытого слоя, $i=1,2,...,K$; j, r- индексы компонентов входного вектора x, j, $r=1,2,...,N$; Q_{rj} – элементы матрицы Q.

Экспериментальные исследования разработанной программной модели гиперрадиально-базисной нейронной сети проведены на примере прогнозирования экономических рядов для акций компаний «Газпром», «Сбербанк» и курса рубля по отношению к доллару США при следующих исходных данных: размер обучающей выборки около 800 значений, размер тестирующей выборки около 400 значений.

Результаты исследованийзависимости средней относительной погрешности результатов от коэффициента обученияпри 7 нейронов в скрытом слое, размере скользящего окна равного 6 и числе эпох обучения равного 100представлены в таблице 1.

Таблица 1. Зависимость средней относительной погрешности результатов от коэффициента обучения

Коэффициент обучения	Средняя относительная погрешность обучения	Средняя относительная погрешность тестирования
0,01	0,0931	0,0828
0,011	0,0858	0,0814
0,012	0,0678	0,0692
0,013	0,0671	0,0694
0,014	0,0702	0,0712
0,015	0,0768	0,0787
0,016	0,0651	0,0695
0,017	0,0602	0,0628
0,018	0,0632	0,0696
0,019	0,0679	0,0742

Результаты исследованийзависимости средней относительной погрешности результатов от размера скользящего окна при 7 нейронов в скрытом слое, коэффициенте обучения равного 0,017 и числе эпох обучения равного 100 представлены в таблице 2.

Таблица 2. Зависимость средней относительной погрешности результатов от размера скользящего окна

Размер скользящего окна	Средняя относительная погрешность обучения	Средняя относительная погрешность тестирования
3	0,0624	0,0239
4	0,0582	0,0293
5	0,0517	0,0324
6	0,0498	0,0455
7	0,0468	0,0501
8	0,0514	0,0690

Результаты исследованийзависимости средней относительной погрешности результатов от количества нейронов скрытого слоя при коэффициенте обучения равного 0,017, размере скользящего окна равного 6 и числе эпох обучения равного 100 представлены в таблице 3.

Таблица 3. Зависимость средней относительной погрешности результатов от числа нейронов скрытого слоя

Число нейронов скрытого слоя	Средняя относительная погрешность обучения	Средняя относительная погрешность тестирования
2	0,0663	0,0755
3	0,0662	0,0754
4	0,0634	0,0697
5	0,0634	0,0684
6	0,1101	0,1347
7	0,0531	0,0576
8	0,0752	0,0970
9	0,0634	0,0683
10	0,0630	0,0674
11	0,1018	0,1390
12	0,1127	0,1288

Результаты исследованийзависимости средней относительной погрешности результатов от количества эпох обучения при 7 нейронов в скрытом слое, коэффициенте обучения равного 0,017, размере скользящего окна равного 6 и представлены в таблице 4.

Таблица 4. Зависимость средней относительной погрешности результатов от числа эпох обучения

Число эпох обучения	Средняя относительная погрешность обучения	Средняя относительная погрешность тестирования
30	0,0480	0,0391
35	0,0478	0,0334
40	0,0478	0,0322
45	0,0483	0,0294
50	0,0485	0,0288
55	0,0486	0,0283
60	0,0492	0,0268
65	0,0497	0,0257
70	0,0501	0,0253
80	0,0506	0,0246
90	0,0508	0,0242
100	0,0517	0,0231

Оптимальные параметры сети: 7 нейронов скрытого слоя, 100 эпох обучения, коэффициент обучения равен 0,017, а размер скользящего окна равен 6.

Литература

1. Осовский, С. Нейронные сети для обработки информации / пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344с.