

## ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ КВАРТАЛЬНОЙ ДИНАМИКИ ВВП

Коробкова С.В. Замаев Б.А.

Научный Центр Нейрокомпьютеров  
Министерство Экономического Развития РФ

В данной задаче требуется спрогнозировать квартальный индекс физического объема ВВП по отношению к соответствующему кварталу предыдущего года. Показатель индекса физического объема ВВП – характеризует изменения физического объема произведенной продукции (товаров и услуг).

Индексы физического объема ВВП являются во всем мире наиболее важным показателем роста и колебаний в экономической конъюнктуре. Индексы физического объема ВВП, сопоставленные с индексами физического объема других показателей, обеспечивают данные для анализа многих аспектов экономики, динамики производительности труда, динамики эффективности использования отдельных элементов национального богатства и т.д. Индексы физического объема отдельных элементов ВВП необходимы для выявления изменений в структуре экономики, анализа изменений в уровне жизни населения, а также для выработки экономической политики и принятия управленческих решений.

В данной работе описано решение задачи прогнозирования квартального индекса физического объема ВВП (горизонт прогнозирования три квартала) и изучения информативности признаков на нейронной сети.

### Исходными данными для решения данной задачи являются следующие:

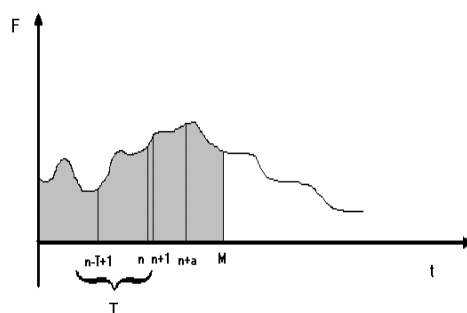
1. Валовой внутренний продукт по кварталам за период с 1995 по первый квартал 2002 года.
2. Индекс физического объема ВВП по кварталам за период с 1995 по первый квартал 2002 года.
3. Дефлятор ВВП по кварталам за период с 1995 по первый квартал 2002 года.
4. Численность занятого в экономике населения по кварталам за период с 1995 по первый квартал 2002 года.
5. Общая численность безработных по кварталам за период с 1995 по первый квартал 2002 года.
6. Индекс потребительских цен по кварталам за период с 1995 по второй квартал 2002 года.
7. Индекс цен производителей промышленной продукции по кварталам за период с 1995 по второй квартал 2002 года.
8. Объем платных услуг по кварталам за период с 1995 по второй квартал 2002 года.
9. Темп роста в % к соотв. периоду прошлого года по кварталам за период с 1995 по второй квартал 2002 года.
10. Среднемесячная номинальная зарплата по кварталам за период с 1996 по первый квартал 2002 года.
11. Средний размер назначенной месячной пенсии по кварталам за период с 1998 по первый квартал 2002 года.
12. Экспорт товаров и услуг по кварталам за период с 1995 по второй квартал 2002 года.
13. Импорт товаров и услуг по кварталам за период с 1995 по второй квартал 2002 года.

### МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧИ

Даны значения многомерной временной дискретной функции  $F$  на интервале  $[1, M]$ . Так как все признаки, зависят от времени представим  $F(x_1, \dots, x_{13})$  как  $F(t)$  – функцию от времени.

Требуется в момент времени  $n$  вычислить упрежденные значения функции  $F$  на временном интервале  $[n+1, n+\alpha]$ .

На рис. 1 серым цветом обозначены известные значения функции  $F$ , то есть значения функции на интервале  $[1, M]$ . Здесь  $n$  - текущий момент времени;  $[n+1, n+\alpha]$  - временной интервал, для которого требуется вычислить упрежденные значения функции  $F$ ; и  $[n-T+1, n]$  - временной интервал, значения функции и значения ее переменных на котором в данный момент времени подаются на вход нейронной сети.

Рис. 1. Вид одномерной функции  $F(t)$ 

В каждый момент времени  $n \in \{T, \dots, M - \alpha\}$  формируется элемент обучающей выборки, в которой входным сигналом нейронной сети является вектор значений функции  $F$  и ее переменных на интервале  $[n - T + 1, n]$ , а желаемым выходным сигналом является вектор значений функции  $F$  на интервале  $[n + 1, n + \alpha]$ . Таким образом, количество обучающих векторов равно  $M - T - \alpha$ .

#### АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ НА НЕЙРОННОЙ СЕТИ.

Анализ цикличности ряда осуществляется со следующими целями:

- определение оптимального размера скользящего окна (размерности вектора входных данных  $T$ )
- Определить оптимальный размер скользящего среднего для медианной фильтрации.

Размерность вектора входных данных можно выбирать кратным длине одного из основных циклов ряда. Это может улучшить обучаемость нейронной сети, так как при определении оптимального размера вектора входных данных путем экспериментов требуются значительные вычислительные затраты, а единой теории об оптимальности размеров входного окна не существует. Также можно выбирать в качестве обучающей выборки не весь интервал наблюдения а только те фрагменты, которые соответствуют фазе цикла исходного временного ряда.

При решении данной задачи для определения цикличности ряда можно использовать эндогенные показатели, которые имеют поквартальную или помесечную составляющую.

На основе информации об их существовании можно осуществить последующий выбор размера скользящего окна  $T$  и сразу правильно выбрать размер скользящего среднего для последующих приложений к выбору размера нейронной сети.

#### Входной сигнал нейронной сети

Входным сигналом нейронной сети является вектор значений функции  $F$  и ее переменных на интервале  $[n - T + 1, n]$ , равный:

$$\bar{x}(n) = \left\{ \begin{array}{l} (F(n - T + 1), x_1, x_2, \dots, x_k), \\ (F(n - T + 2), x_1, x_2, \dots, x_k), \\ \dots, \\ (F(n - 1), x_1, x_2, \dots, x_k), \\ (F(n), x_1, x_2, \dots, x_k) \end{array} \right\}, \quad (1)$$

при  $n \in \{T, \dots, M - \alpha\}$ .

#### Выходной сигнал нейронной сети

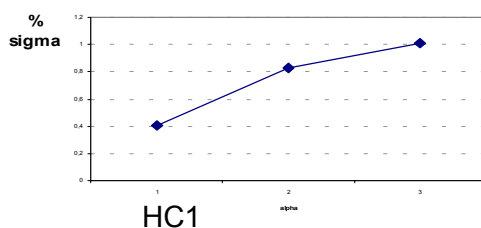
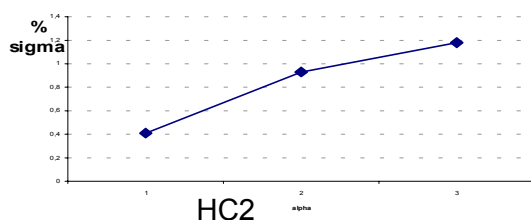
Выходным сигналом НС в момент времени  $n$  является вектор аппроксимированных нейронной сетью упрежденных на интервале  $[n + 1, n + \alpha]$  значений функции  $F$ , обозначаемый через  $\bar{y}(n) = \{y_1(n), \dots, y_\alpha(n)\}$ .

#### Желаемый выходной сигнал нейронной сети

Желаемым выходным сигналом нейронной сети в момент времени  $n$  является вектор значений функции  $F$  на интервале  $[n + 1, n + \alpha]$ .

### Ошибка решения

Ошибка решения определяется как разность между желаемым и действительным выходным сигналом нейронной сети в дискретный момент времени  $n$ , то есть ошибка решения показывает расстояние до правильного значения. В нашем случае, ошибка решения для текущего момента времени  $n$  представляет собой вектор,  $j$ -ый элемент которого  $e_j(n)$  равен:



$$e_j(n) = y_j(n) - F(n + j), \quad (2)$$

где  $n$  - текущий момент времени,  $n \in \{T, \dots, M - \alpha\}$ ;

$j$  - текущая позиция на интервале упреждения,  $j \in \{1, \dots, \alpha\}$ .

### Функция активации

В решаемой задаче диапазон изменения значений функции  $F(x)$  составляет  $[0, 2m]$ , где  $2m$  - максимальное значение функции  $F$  на интервале  $[0, M]$ .

С учетом вышесказанного, далее будет использоваться функция активации типа арктангенса вида:

$$y_{h_k}(n) = m \left( \frac{2}{\pi} \arctg(g_{h_k}(n)) \right) + m, \quad (3)$$

### Описание структуры разомкнутой нейронной сети

В данной задаче в выходном слое должно быть  $\alpha$  нейронов. Структура разомкнутой  $K$ -слойной сети описывается уравнениями (3) и (4).

Число настраиваемых параметров  $K$ -слойной нейронной сети равно:

$$\sum_{k=1}^K (H_{(k-1)} + 1) \cdot H_k, \quad (5)$$

где  $H_{(k-1)}$  - количество нейронов  $(k-1)$ -го слоя; при  $(k-1) = 0$ ,  $H_0 = T$ ;

$H_k$  - количество нейронов  $k$ -го слоя НС.

### РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Для прогнозирования индекса физического объема ВВП были применены два типа нейронной сети:

- трехслойная нейронная сеть прямого распространения с числом нейронов первого слоя 5, второго слоя – 3 (НС1);
- трехслойная нейронная сеть прямого распространения с числом нейронов первого слоя 10, второго слоя – 3 (НС2);

Приведены графики зависимости среднеквадратичной ошибки (sigma) от времени экстраполяции (alpha).

Расчеты показывают, что точность (среднеквадратичная ошибка) прогнозирования индекса физического объема ВВП не превышает 1,2 процентных пункта для времени экстраполяции 1, 2, 3 квартала.

**ИНФОРМАТИВНОСТЬ ПРИЗНАКОВ**

Прогнозирование индекса физического объема ВВП производилось с использованием конкретных физических признаков, которыми являлись признаки, указанные выше. Необходимо определить информативность этих признаков при решении задачи предсказания индекса физического объема ВВП. Предположим, что среднеквадратическая ошибка предсказания при использовании полного набора признаков равна  $\sigma(\sigma)$ , а при использовании полного набора признаков с исключением  $i$ -того признака равна  $\sigma_i$ . Тогда информативность  $i$ -того признака может быть оценена как  $(\sigma_i - \sigma)$ .

Обучение нейронной сети производилось при подаче на вход значений признаков за четыре квартала и формированием указания учителя в виде значения индекса физического объема ВВП, упрежденного на 1, 2 и 3 квартала. Исходная выборка делилась на две части, из которых одна использовалась для обучения, а другая – для тестирования – проверки качества работы обученной системы.

Далее приводятся графики среднеквадратичной ошибки ( $\sigma$ ) экстраполяции от времени экстраполяции ( $\alpha$ ) при удалении из выборки следующих признаков: точности прогноза необходимо: а) расширить набор показателей для увеличения полного пространства признаков; б) по мере накопления отчетной статистической информации увеличивать длину временных рядов исходных показателей.

**Среднеквадратичная ошибка**

Горизонт прогнозирования	полное пр-во признаков	GDP	GDP2	IPC1	IPP1	экон	безр	PSqv	Pu_t	Wqv	Pqv	Eqv	Iqv
1 квартал	НС1	0,4	0,8	0,4	0,5	0,6	0,5	0,4	0,5	0,4	0,4	0,6	0,4
2 квартала	НС1	0,9	1,3	1,1	1,1	0,8	0,6	1,2	0,8	2,3	1,4	0,4	1,1
3 квартала	НС1	1,2	1,1	1,4	1,2	1,2	1	1,5	1,2	1,4	1,2	0,8	1,4
1 квартал	НС2	0,4	0,4	0,5	0,7	0,6	0,4	0,5	0,4	0,5	0,4	0,8	0,5
2 квартала	НС2	0,8	0,6	0,9	0,9	1,1	1,1	1,4	1,2	1,6	1,2	1,3	1,1
3 квартала	НС2	1	1,1	1,4	1,2	1,2	1,1	1,1	1,1	1	1,2	1,4	1,3

**Информативность признаков**

Горизонт прогнозирования	GDP	GDP2	IPC1	IPP1	экон	безр	PSqv	Pu_t	Wqv	Pqv	Eqv	Iqv
1 квартал НС1	0,4	0	0,1	0,2	0,1	0	0,1	0	0	0	0,2	0
2 квартала НС1	0,4	0,2	0,2	-0,1	-0,3	0,3	-0,1	1,4	0,5	-0,5	0,1	0,2
3 квартала НС1	-0,1	0,2	0	0	-0,2	0,3	0	0,2	0	-0,4	0,2	0,2
1 квартал НС2	0	0,1	0,3	0,2	0	0,1	0	0,1	0	0,4	0,1	0,2
2 квартала НС2	-0,2	0,1	0,1	0,3	0,3	0,6	0,4	0,8	0,4	0,5	0,3	0,2
3 квартала НС2	0,1	0,4	0,2	0,2	0,1	0,1	0,1	0	0,2	0,4	0,3	-0,1

### **Выводы**

На основе проведенных прогнозных расчетов квартального индекса физического объема ВВП (горизонт прогнозирования три квартала) сделаны следующие выводы. 1. Использование НС позволяет сделать прогноз квартальной динамики ВВП с достаточной степенью точности. 2. Для повышения статистической информации увеличивать длину временных рядов исходных показателей.

### **ЛИТЕРАТУРА**

1. Neural Network Algorithms for Function Extrapolation and their Application in Call-center Forecasting Tasks. *A.I. Galuskin, D.S. Tomashevich, N.S. Tomashevich, M.U. Muromsky, E.A. Shachnev.* "Neurocomputers: Design and Applications", Volume 1, Issue 3,4,5, 2000. Begell House, inc. Publishers.
2. Neural Networks for Conditional Probability Estimation. Dirk Husmeier Atenæum Press Ltd., Gateshead, Tyne&Wear, 275p.



## NEURAL NETWORK APPLICATION FOR THE QUARTER GDP DINAMICS PREDICTION

Korobkova S., Zamaraev B.

Scientific Center of Neurocomputers  
The Department of Economic Development, Russia

In this task we should predict the quarter index of the physical size of GDP relative to the size of it in the corresponding quarter of the previous year. The size of the index of the physical size of GDP characterizes the change in the quantity of the manufactured production (services and goods).

The indexes of the physical size of GDP are the most important characteristics of growth and fluctuations in the state of the market worldwide.

This paper presents the solution of the task of the quarter index of the physical size of GDP prediction (the prediction horizon is three quarters) and the task of self-descriptiveness study using neural networks.

**MATHEMATICAL TASK DEFINITION**

There are thirteen characteristics of the multidimensional function  $F$  in the range  $[1, M]$ . All these characteristics depend on time so we can represent the function  $F(x_1, \dots, x_{13})$  as time function  $F(t)$ .

We should calculate the preceded values of the function  $F$  in the time range  $[n+1, n+\alpha]$ .

The figure 1 shows the given values of the function (grey). In this figure  $n$  is the current moment of time  $[n+1, n+\alpha]$  is the range of time where we should calculate the preceded function values and  $[n-T+1, n]$  the time range the values of function of which are the current inputs of the neural network.

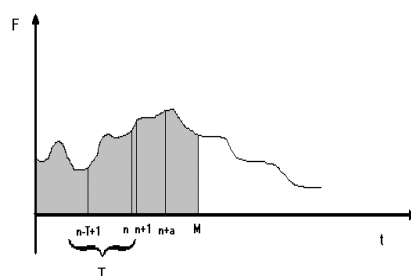


Fig.1. The single-dimensional function  $F(t)$

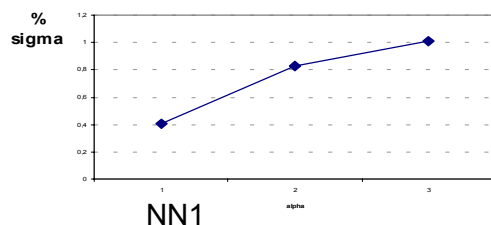
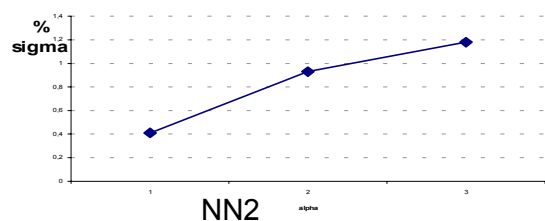
**THE EXPERIMENTS**

We used two types of the neural network for the prediction of the index of the physical size of GDP:

- three layers feedforward neural network with the first layer neuron number 5 and the second layer neuron number – 3 (NN1);
- three layers feedforward neural network with the first layer neuron number 10 and the second layer neuron number – 3 (NN2);

Here we show the graphs of the mean square error (sigma) that depends of the extrapolation time (alpha).

	2001			2002			
	II quarter	III quarter	IV quarter	I quarter			
GDP1 (report data)	105,2	105,8	104,3	103,7			
	NN1	NN2	NN1	NN2	NN1	NN2	
Prediction values							
ceiling		0,4	0,4	0,9	0,8	1,2	1
floor		-0,4	-0,4	-0,9	0,8	-1,2	-1



### CONCLUSIONS

The calculations show that the neural network algorithms are applicable for the task of the quarter index of the physical size of GDP prediction and can provide the desired accuracy of calculations.

### REFERENCES

1. Neural Network Algorithms for Function Extrapolation and their Application in Call-center Forecasting Tasks. *A.I. Galuskin, D.S. Tomashevich, N.S. Tomashevich, M.U. Muromsky, E.A. Shachnev.* "Neurocomputers: Design and Applications", Volume 1, Issue 3,4,5, 2000. Begell House, inc. Publishers.
2. Neural Networks for Conditional Probability Estimation. Dirk Husmeier Atenæum Press Ltd., Gateshead, Tyne&Wear, 275p.