

**Задачи к семинарскому занятию на тему
"Многослойные нейронные сети"**

Задача 1

Многослойная нейронная сеть (МНС) имеет 3 входа, 2 выхода и 2 рабочих (скрытых) слоя, содержащих 3 и 2 нейрона соответственно в первом и втором слое.

а) Нарисуйте схему сети и укажите на схеме обозначения синаптических коэффициентов, смещений и выходов нейронов, используя символику расширенных векторов выходов нейронов $\tilde{S}^{(q)}$ и матриц синаптических коэффициентов слоев $\tilde{W}^{(q)}$.

б) Раскройте содержание расширенной матрицы синаптических коэффициентов $\tilde{W}^{(2)}$ второго слоя нейронов. Какова размерность матрицы $\tilde{W}^{(2)}$? Чему соответствует каждая строка ее элементов?

в) Напишите выражение для выхода первого нейрона второго слоя, если активационная характеристика во втором слое – гиперболический тангенс.

Задача 2

Система дистанционного управления движением робототехнического устройства вдоль прямой линии использует оценку текущей скорости устройства, формируемую многослойной нейронной сетью. На вход МНС поступают текущее и три предыдущих измерения координатного положения робота на прямой: $x_t, x_{t-1}, x_{t-2}, x_{t-3}$. Сеть использует более одного предыдущего измерения в связи с большой ошибкой измерения координаты x в целях повышения точности текущей оценки скорости.

МНС содержит один скрытый слой с тремя нейронами и логистической активационной характеристикой нейронов. Выходной нейрон является линейным.

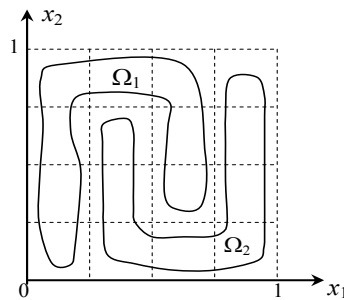
а) Нарисуйте схему нейросетевой обработки данных с указанием обозначений для выходов нейронов и синаптических коэффициентов.

б) Напишите расчетное выражение для оценки скорости, формируемой нейронной сетью.

в) Покажите в форме таблицы структуру обучающей выборки, принимая во внимание заданные в условии задачи размерности переменных. Предполагается, что для получения выборочных примеров были проведены специальные эксперименты с установленным на робототехническом устройстве дорогостоящим датчиком скорости высокой точности.

Задача 3

Решается задача классификации данных. Объекты, характеризующиеся вектором признаков размерности $M = 2$, принадлежат двум классам. Области принадлежности объектов классам 1 и 2 представлены на рисунке.



а) Предложите архитектуру МНС, решающей задачу классификации (с указанием активационных характеристик нейронов).

б) Постройте обучающую выборку, содержащую 10 примеров – по 5 примеров каждого класса.

Задача 4

Требуется построить МНС, которая формирует на выходе оценки экстраполированных значений $x_{t+1}, x_{t+2}, x_{t+3}$ временного ряда по наблюдениям $x_t, x_{t-1}, x_{t-2}, x_{t-3}$.

а) Предложите архитектуру МНС, решающей задачу прогноза временного ряда. Укажите K – число слоев, $N_q, q=1, K$ – распределение нейронов по слоям, выберите активационные характеристики нейронов в слоях.

б) Постройте два примера обучающей выборки, если выборочная реализация временного ряда задана таблицей:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14
x	1.3	1.5	1.7	1.8	1.8	1.6	1.5	1.3	1.2	1.2	1.3	1.4	1.5	1.4

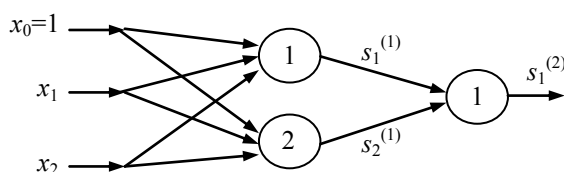
Задача 5

Нейронная сеть имеет один вход, один скрытый слой, содержащий 5 нейронов с гауссианой $f(h) = \exp(-h^2/2)$ в качестве активационной характеристики, и один линейный выходной нейрон.

Используя решение задачи 3 семинара 1 (о функции, реализуемой нейроном с одним входом и гауссовой активационной характеристикой), напишите выражение для выхода нейронной сети. Объясните, что построенная двухслойная нейронная сеть способна аппроксимировать (моделировать) гладкую функцию одной переменной.

Задача 6

Нейронная сеть имеет архитектуру, представленную на рисунке.



Нейроны первого слоя имеют гауссову активационную характеристику, а выходной нейрон – логистическую.

Покажите, что настройкой синаптических коэффициентов и смещений нейронов сети возможно реализовать на ней колоколообразную функцию двух переменных.

Контрольные вопросы по методу обратного распространения ошибки

1. Напишите выражение для критериальной функции $D(\cdot)$, в соответствии с которой осуществляется обучение МНС.
2. По какому правилу осуществляется настройка синаптических коэффициентов МНС при использовании простого градиентного метода?
3. Какую задачу решает метод обратного распространения ошибки? Какие переменные он позволяет рассчитать?
4. Какой подход является альтернативным по отношению к методу обратного распространения ошибки?
5. Покажите на примере (для МНС заданной архитектуры) эффективность метода обратного распространения ошибки по сравнению с альтернативным подходом.
6. Что называется двойственной переменной в сети обратного распространения? По отношению к какой переменной она является двойственной?

7. Допустим, что известны двойственные переменные $\Delta_m^{(p)(q+1)}$, $m = \overline{1, N_{q+1}}$, для всех нейронов слоя $(q+1)$ при использовании примера p обучающей выборки. Каким выражением определяется двойственная переменная $\Delta_i^{(p)(q)}$?

8. Напишите расчетное выражение для $\frac{\partial D_p}{\partial \tilde{w}_{ij}^{(q)}}$, $q = \overline{1, K}$, $i = \overline{1, N_q}$, $j = \overline{0, N_{q-1}}$, если двойственные

переменные вычислены и известны их значения.

9. В чем состоит мнемоническое правило построения схемы обратного распространения ошибки?
10. Сравните количество нейронов в схемах прямого и обратного распространения.

Задача 7

Нейронная сеть имеет три входа ($M = 3$) и два выходных нейрона с активационной характеристикой

$$f_1(h) = \frac{\alpha h}{1 + \alpha |h|}, \quad \alpha > 0. \text{ Скрытые слои отсутствуют } (K=1).$$

а) Для одного обучающего примера p выведите выражения для производных критериальной функции $D(\cdot)$ по синаптическим коэффициентам $w_{ij}^{(1)}$, $i = 1, 2$; $j = \overline{0, 3}$.

б) Постройте расчетную схему для производных $\frac{\partial D_p}{\partial \tilde{w}_{ij}^{(1)}}$, $i = 1, 2$; $j = \overline{0, 3}$.

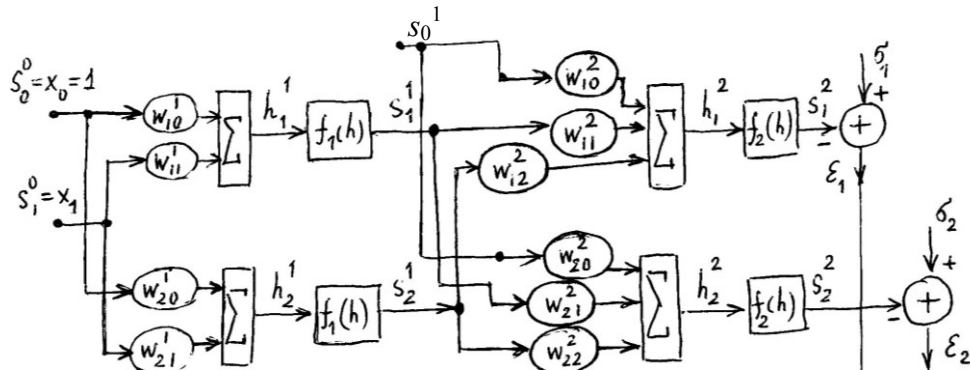
в) В какой точке схемы реализуется двойственная переменная $\Delta_1^{(p)(1)}$?

г) Как учесть в схеме реальный объем выборки $P > 1$ при расчете $\frac{\partial D}{\partial \tilde{w}_{ij}^{(1)}}$?

д) Какое нелинейное преобразование реализуется в схеме обратного распространения ошибки?

Задача 8

Для аппроксимации векторной функции одной переменной применяется нейронная сеть, содержащая два слоя.



Первый слой содержит 2 нейрона с активационной характеристикой $th(h)$, а второй – два нейрона с линейной активационной характеристикой. Обучение сети выполняется методом обратного распространения ошибки в режиме "по примерам".

Напишите выражения для производных критерия по всем настраиваемым синаптическим коэффициентам и постройте соответствующую схему обратного распространения ошибки.

Задача 9

В условиях предыдущей задачи постройте на основе мнемонического правила схему обратного распространения ошибки для вычисления двойственных переменных и убедитесь, что она совпадает со схемой, построенной в задаче 8 на основе аналитических выражений.

Задача 10

Покажите, что в схеме обратного распространения ошибки в условиях задачи 8 нет необходимости в реализации функции $th(h)$. Напишите выражения для нелинейных преобразователей схемы.

Задача 11

Рассчитайте коррекцию синаптического коэффициента $w_{11}^{(2)}$ после предъявления обучающего примера ($s_1^0 = x_1 = 1$; $\sigma_1 = -0.1$; $\sigma_2 = 0.3$) в условиях задачи 8, если известно, что на предыдущем такте обучения были получены следующие значения синаптических коэффициентов:

$$\tilde{W}^{(1)} = \begin{pmatrix} -0.5 & 0.4 \\ 0.2 & -0.3 \end{pmatrix}; \quad \tilde{W}^{(2)} = \begin{pmatrix} -0.1 & -0.2 & 0.01 \\ 0.6 & 0.4 & -0.1 \end{pmatrix}.$$

Параметр скорости обучения α положите равным 0.05.

При расчете $th(h)$ для малых значений h воспользуйтесь приближением $th(h) \approx h$.

Задача 12

Один из нейронов МНС имеет 12 входов с предыдущего слоя. Активационные характеристики предыдущего и текущего слоев являются функциями $th(h)$. Предполагается, что опасность "паралича" нейрона наступает, если $|h|$ превышает значение 1.

Рассчитайте диапазон допустимых значений синаптических коэффициентов рассматриваемого нейрона при его инициализации в следующих условиях:

а) выходы нейронов предыдущего слоя – независимые случайные величины, распределенные равномерно во всем диапазоне своих возможных значений;

б) максимальное значение $|h|$ достигается при $|h| = 2 \sigma_h$.