

Министерство образования и науки Украины

Донбасская государственная машиностроительная академия

# **НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ**

## **КОНСПЕКТ ЛЕКЦИЙ**

(для студентов специальности 8.04030302 и 7.04030302 «Системы и методы принятия решений» всех форм обучения)

**Краматорск 2014**

Министерство образования и науки Украины

**Донбасская государственная машиностроительная академия**

# **НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ**

## **КОНСПЕКТ ЛЕКЦИЙ**

(для студентов специальности 8.04030302 и 7.04030302 «Системы и методы принятия решений» всех форм обучения)

**Краматорск 2014**

Конспект лекций по курсу «Нейросетевые технологии» (для специальности 8.04030302 и 7.04030302 «Системы и методы принятия решений» всех форм обучения) / Сост.: Гитис В.Б. – Краматорск: ДГМА, 2014. – 20 с.

В конспекте лекций приводятся теоретические сведения по курсу «Нейросетевые технологии» для квалификационного уровня «специалист» и «магистр».

Составитель	В.Б. Гитис, к.т.н., доцент каф. ИСПР
-------------	--------------------------------------

Отв. за выпуск	В.Б.Гитис, к.т.н., доцент каф. ИСПР
----------------	-------------------------------------

## СОДЕРЖАНИЕ

1. Нейросетевая компрессия данных	5
2. Прогнозирование временных рядов с помощью нейронных сетей	7
3. Решение задач оптимизации нейронными сетями	10
4. Нейроуправление	14
Список рекомендуемой литературы	19

## 1. НЕЙРОСЕТЕВАЯ КОМПРЕССИЯ ДАННЫХ

Важной проблемой, возникающей при обработке больших массивов наблюдений, является сжатие данных с целью выделения наиболее существенной информации и выявления факторов, неявным образом определяющих природу изучаемого явления. Одним из наиболее эффективных подходов к решению этой проблемы является аппарат факторного анализа.

Основная идея факторного анализа, допускающая наличие априорно неизвестных скрытых факторов, ведет к следующей неформальной задаче: наблюдая большое количество измеряемых параметров (показателей), выявить небольшое число параметров-факторов, определяющих в основном поведение измеряемых параметров. Или зная значения большого количества функций измеряемых параметров, установить соответствующие значения общих для всех функций аргументов-факторов и восстановить вид этих функций.

Задача нелинейного факторного анализа эффективно может быть решена с помощью автоассоциативного трехслойного персептрона, называемого «*Бутылочное горлышко*».

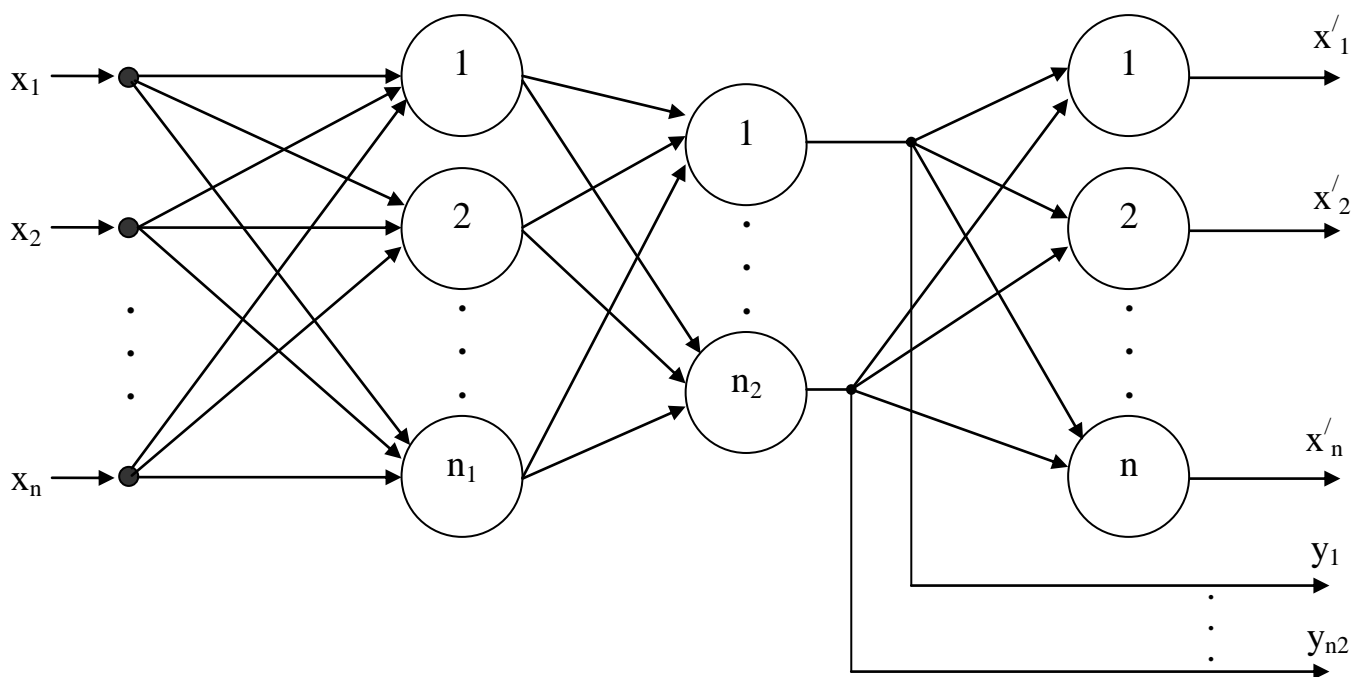


Рисунок 1 – Схема автоассоциативного трехслойного персептрона

В процессе обучения минимизируется целевая функция в виде

$$E = \sum_k E_k = \frac{1}{2} \sum_k \sum_{i=1}^n e_{ik}^2 = \frac{1}{2} \sum_k \sum_{i=1}^n (x_{ik} - x'_{ik})^2,$$

где  $k$  – номер обучающего примера.

В качестве алгоритма обучения может быть использована любая процедура обратного распространения ошибки. В качестве эталонного выхода используется сам входной сигнал  $\overline{X}$ , подлежащий сжатию.

На входной слой поступает  $n$ -мерный вектор входных сигналов  $\overline{X}$ , первый скрытый слой содержит  $n_1$  нейронов, второй скрытый слой –  $n_2 < n$  нейронов и выходной слой –  $n_3 = n$  нейронов. Целью ассоциативного обучения является восстановление на выходе сети сигнала  $\overline{X'}$ , наилучшим образом аппроксимирующего входной сигнал  $\overline{X}$ . Сжатие информации происходит во втором скрытом слое, содержащем меньшее число нейронов чем первый и выходной слою. Именно с выхода второго скрытого слоя снимается «сжатый» сигнал  $\overline{Y}$ , при этом в результате такого подхода к компрессии информации достигается оптимальное решение задачи нелинейного факторного анализа.

## **2. ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ**

При решении задачи прогнозирования многомерных стохастических временных рядов в условиях структурной и параметрической неопределенности в общем случае природа наблюдаемой последовательности неизвестна.

Во временных рядах, в отличие от случайных выборок, данные рассматриваются как последовательность измерений, упорядоченных в неслучайные моменты времени, т.е. временной ряд включает два обязательных элемента – время и конкретное значение показателя (уровень ряда). Как правило, значения показателя фиксируются через равные промежутки времени

Для построения временного ряда должны быть выполнены следующие условия:

- периодизация развития образуется расчленением во времени на однородные этапы, в пределах которых показатель подчиняется общему закону развития;
- значения временного ряда должны быть сопоставимы по всем признакам, по которым осуществляется его формирование, например по кругу охватываемых явлений, единицам измерения, методологии расчёта и т.д.;
- периоды должны соответствовать интенсивности процессов;
- временной ряд должен быть полным, т.е. не допускаются пропуски, если они неизбежны, ряд дополняют условно-расчетными значениями.

Только в случае выполнения перечисленных условий рассматриваемая последовательность является временным рядом и для нее применимы методы анализа и прогнозирования временных рядов.

В задачах анализа временных рядов целью является прогноз будущих значений переменной, зависящей от времени, на основе предыдущих ее значений. При этом не ставится задача выявить неявно отраженные во временном ряду зависимости целевого показателя от существенных для него значений других характеристик предметной области.

Обычно очередное значение временного ряда прогнозируется по некоторому числу его предыдущих значений (прогноз на один шаг вперед во времени). Однако можно выполнять прогноз на любое число шагов (при снижении точности прогнозирования).

После того, как вычислено очередное предполагаемое значение, оно подставляется обратно и с его помощью получается следующий прогноз. Такой способ называется *проекцией временного ряда* или *методом «окон»*. Окно имеет фиксированный размер и способно перемещаться по временной последовательности.

Искусственные нейронные сети являются одним из наиболее адекватных инструментов прогнозирования временных рядов, позволяющих по прошлым наблюдениям восстанавливать нелинейное отображение вида

$$\hat{x}(t) = F(x(t-1), x(t-2), \dots, x(t-k)) + e(t) = x(t) + e(t),$$

где  $\hat{x}(t)$  – оценка (прогноз) значения  $x(t)$ , полученная на выходе нейросети;

$e(t)$  – ошибка прогнозирования;

$k$  – порядок модели (ширина окна).

Нейросеть представляет в данном случае нелинейную авторегрессионную модель (NAR-модель).

Возможность и эффективность использования NAR-модели в задачах прогнозирования определяется теоремой Тэкенса, устанавливающей существование порядка модели, который обеспечивает сколь угодно малое значение ошибки  $e(t)$ , и универсальными аппроксимирующими свойствами нейросетей.

В качестве основы для построения NAR-моделей чаще всего используются многослойные сети с прямой передачей информации.

Количество входов сети равно ширине окна, а количество выходов – величине смещения окна при скольжении (числу временных периодов прогнозирования). Обычно окно смещают на одну позицию.



Сеть позволяет сформировать на выходе сигнал, являющийся функцией времени. Такие сети также называются *нейросетями с временными задержками*. Они получили широкое распространение в задачах прогнозирования временных последовательностей самой различной природы.

При прогнозировании временных рядов большое значение имеет качество обучающей выборки, которое тем выше, чем меньше ее противоречивость и больше повторяемость.

*Противоречивость* – показатель, характеризующий количество одинаковых объектов, принадлежащих к разным классам.

*Повторяемость* – показатель, характеризующий количество одинаковых наборов в рамках одного класса.

Факторы влияющие на противоречивость и повторяемость:

- количество элементов обучающей выборки – чем больше элементов, тем больше противоречивость и повторяемость;
- глубина погружения во временной ряд («окно») – чем больше глубина, тем меньше противоречивость и меньше повторяемость.

*Недостатки:*

1. Большое число настраиваемых весов и низкая скорость обучения, что вызывает особые проблемы при работе в реальном времени;
2. Сложность в подборе оптимальной ширины окна. При наличии четко выраженной цикличности в поведении временного ряда может выявиться существенная зависимость результатов прогнозирования от корреляции между шириной окна и циклом колебаний временного ряда;
3. Обеспечение актуальности прогнозов. Последовательность значений временного ряда является отражением скрытых зависимостей, значение которых со временем может изменяться (увеличиваться или наоборот, практически полностью исчезать). Поэтому в сеть постоянно нужно вводить новые данные, находя правильный баланс со старыми данными.

### 3. РЕШЕНИЕ ЗАДАЧ ОПТИМИЗАЦИИ НЕЙРОННЫМИ СЕТЯМИ

Высокая степень распараллеливания обработки информации позволяет успешно их применять для решения задач комбинаторной оптимизации. Наиболее эффективно нейросетями решаются задачи транспортно-ориентированной оптимизации (задача коммивояжера и ее модификации) и задачи распределения ресурсов (задача о назначениях, задача целераспределения и другие). Решение таких задач традиционными методами математического программирования, большинство из которых ориентировано на вычислительные системы с последовательной архитектурой, сопряжено с большими затратами времени, часто не приемлемыми для многих приложений.

Одной из наиболее сложных проблем оптимизации является *задача коммивояжера*. Она состоит в определении кратчайшего пути, соединяющего некоторое множество городов, причем каждый город должен быть посещен лишь один раз.

В случае  $n$  городов существует  $n!/2n$  различных маршрутов, среди которых необходимо найти кратчайший. Это пример NP – полной задачи или неполиномиально сложной задачи, поиск точного решения которой за время, не превосходящее полинома от размера данных невозможен. Решение таких задач методом перебора является неоправданно трудоемким.

Хопфилд и Тэнк предложили подход к ее приближенному решению на основе сетей Хопфилда.

Для описания возможных маршрутов введен специальный тип матрицы, в которой города образуют строки (A, B, C, D... – названия городов), а столбцы отображают последовательность городов в маршруте. Тогда произвольный маршрут может быть представлен в виде таблицы, в которой единица в строке, отвечающей данному городу, определяет его номер в маршруте.

*Пример:* маршрут  $D \Rightarrow B \Rightarrow E \Rightarrow A \Rightarrow C$ :

Таблица 1 – Пример маршрута

Город	Последовательность				
	1	2	3	4	5
A	0	0	0	1	0
B	0	1	0	0	0
C	0	0	0	0	1
D	1	0	0	0	0
E	0	0	1	0	0

Сопоставим клетке таблицы на пересечении строки  $X$  и столбца  $i$  нейрон  $S_{xi}$  с состоянием  $\{0,1\}$ . Возбужденное состояние данного нейрона сигнализирует о том, что город  $X$  в маршруте следует посещать в  $i$ -тую очередь.

Тогда к сети Хопфилда предъявляются следующие требования:

1. Сеть должна состоять из  $N = n \times n$  нейронов, которые рассматриваются как квадрат из  $n$  строк и  $n$  столбцов, где  $n$  – число городов;
2. Ответ сети должен содержать только один активный нейрон в каждой строке и каждом столбце;
3. Активный нейрон в первом столбце задаёт первый город маршрута, во втором столбце – второй город маршрута, и так далее.

Для решения этой задачи достаточно следующее:

1. Для выполнения условия 2 веса сети должны быть построены таким образом, чтобы каждый нейрон препятствовал активации других нейронов в своей строке и в своём столбце;
2. Для минимизации длины пути необходимо, чтобы нейрон в  $i$ -м столбце тем активнее препятствовал активации нейронов в  $i+1$ -м и  $i-1$ -м столбцах, чем больше расстояние между ними;
3. Для работоспособности сети Хопфилда необходимо, чтобы веса сети не были все отрицательными.

Целевая функция задачи поиска оптимального маршрута будет включать 4 слагаемых:

$$E(S) = E_1 + E_2 + E_3 + E_4 \rightarrow \min,$$

где  $E_1$  – означает, что каждый город должен быть посещен не более чем один раз (в каждой строке матрицы имеется не более одной единицы);

$E_2$  – означает, что под каждым номером должно посещаться не более одного города (в каждом столбце – не более одной единицы);

$E_3$  – означает, что общее число посещений равно числу городов  $n$  (в матрице всего имеется ровно  $n$  единиц).

$$\begin{cases} E_1 = \frac{\alpha}{2} \sum_x \sum_i \sum_{j \neq i} S_{xi} S_{xj} \\ E_2 = \frac{\beta}{2} \sum_x \sum_i \sum_{y \neq x} S_{xi} S_{yj} \\ E_3 = \frac{\gamma}{2} \left( \sum_x \sum_i S_{xi} - n \right)^2 \end{cases}$$

Каждое из трех слагаемых обращается в нуль на допустимых маршрутах, и принимает значения больше нуля на недопустимых. Таким образом, первые три слагаемых отвечают за допустимость маршрута.

Слагаемое  $E_4$  минимизирует длину маршрута:

$$E_4 = \frac{\eta}{2} \sum_x \sum_i \sum_{y \neq x} d_{xy} S_{xi} (S_{yi+1} + S_{yi-1}),$$

где  $d_{xy}$  – расстояние между городами  $X$  и  $Y$ .

При этом отрезок пути  $XY$  включается в сумму только тогда, когда город  $Y$  является относительно города  $X$  либо предыдущим, либо последующим.

Множители  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\eta$  задают относительные веса слагаемых.

Значения весов и порогов сети задаются следующим образом:

$$\begin{cases} W_{xij} = -\alpha\delta_{xy}(1 - \delta_{ij}) - \beta\delta_{ij}(1 - \delta_{xy}) - \gamma - \mu d_{xy}(\delta_{ji+1} - \delta_{ji-1}) \\ T_{xi} = -n, \end{cases}$$

где  $\delta_{xy}$  – символ Кронекера, принимающий значение 1, если  $x = y$  и значение 0 в противном случае:

$$\delta_{xy} = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases}$$

Первый член равен  $-\alpha$  для всех связей в той же строке ( $x = y$ ), кроме связи нейрона с самим собой (при  $i = j$ ). Второй член равен  $-\beta$  для всех связей в том же столбце ( $i = j$ ), кроме связи с самим собой ( $x = y$ ). Четвертый член пропорционален расстоянию между городами  $x$  и  $y$ , если эти города соседние в маршруте ( $i = j-1$  или  $i = j+1$ ).

Если такую сеть привести в случайное начальное состояние, то результирующее стабильное состояние даст субоптимальный путь, длина которого будет близка к оптимальной. При этом сам путь может значительно отличаться от оптимального. Для практического применения сеть следует запустить несколько раз, и выбрать наилучший путь.

Таким образом, решение данной задачи демонстрирует альтернативный подход к задачам оптимизации: если возможно перевести условия некоторой задачи в параметры связей между нейронами, то она может быть решена сетью без какого-либо дополнительного анализа.

## 4. НЕЙРОУПРАВЛЕНИЕ

Управление динамическими объектами и системами в условиях структурной и параметрической неопределенности является важным направлением в приложениях нейронных сетей.

Нейросетевые системы управления, или *нейрорегуляторы* имеют следующие преимущества перед обычными системами автоматического управления:

1. Возможности обработки больших объемов сенсорной информации;
2. Высокое быстродействие, достигаемое благодаря распараллеливанию вычислений;
3. Возможность работы с существенно нелинейными объектами, о структуре и параметрах которых практически ничего неизвестно.

Нейрорегуляторы являются развитием адаптивных систем управления и могут быть разбиты на два класса: *прямые* и *непрямые* системы.

### Прямая система управления

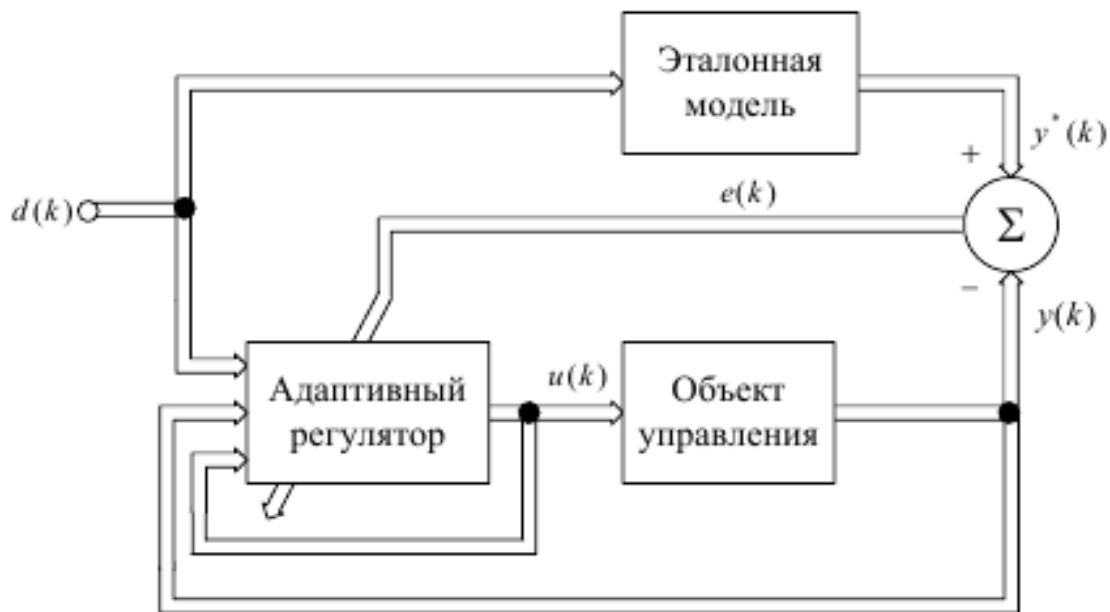


Рисунок 2 – Схема прямого адаптивного управления

На вход системы управления поступает внешний задающий сигнал  $d(k)$ , который подается параллельно на регулятор и эталонную модель, определяющую желаемое поведение объекта управления. Наличие эталонной модели не является обязательным. Цель управления определяется требованием устойчивого и точного слежения выходом объекта  $y(k)$  за задающим сигналом  $d(k)$ .

Адаптивная прямая система в своей работе использует ошибку

$$e(k) = y^*(k) - y(k),$$

где  $y(k)$  – выходной сигнал реального объекта,

$y^*(k)$  – выходной сигнал эталонной модели.

В процессе работы такой системы настраиваются параметры регулятора. Вырабатываемые регулятором управляющие воздействия  $u(k)$  определяются путем аналитической минимизации принятого критерия управления, являющегося функцией от ошибки управления  $e(k)$ .

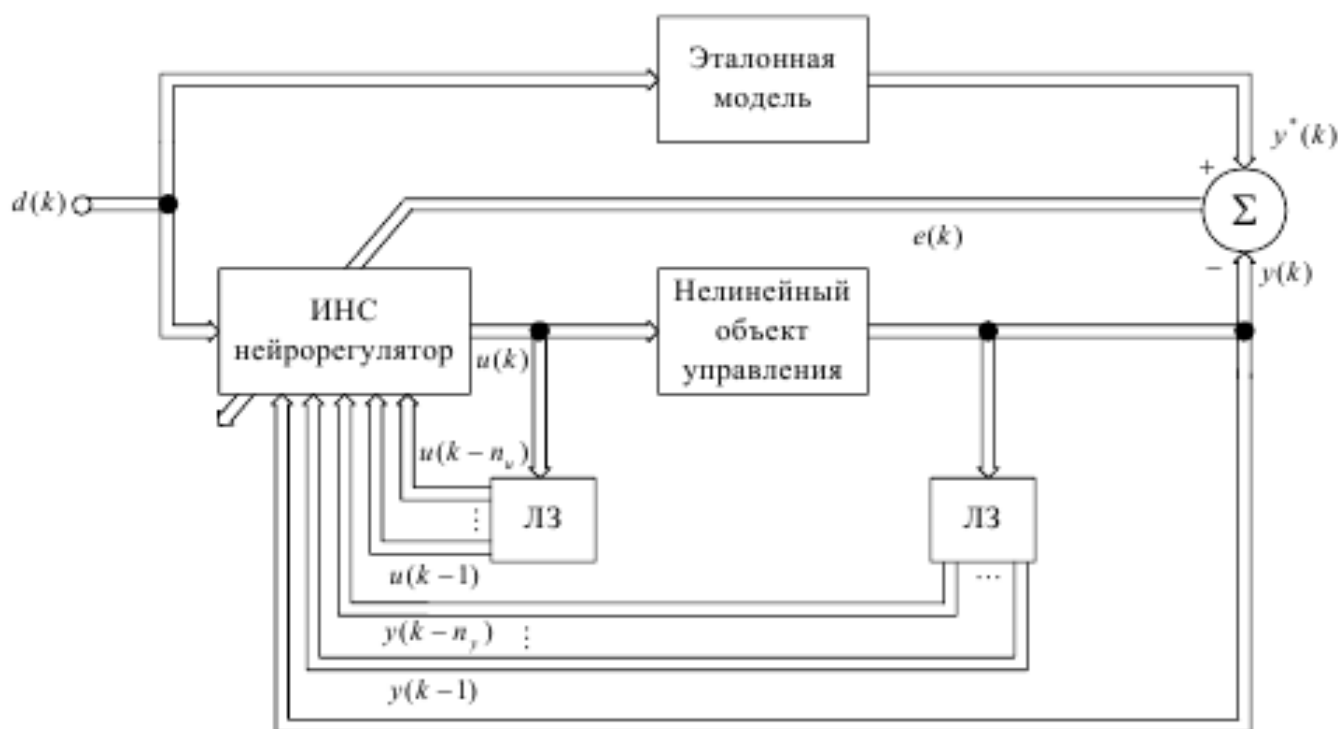


Рисунок 3 – Структура прямого нейрорегулятора

Прямой нейрорегулятор содержит нейросеть (например, двуслойный персептрон), обеспечивающую оптимизацию принятого критерия управления по данным наблюдений за входами и выходами нелинейного объекта.

Элементы ЛЗ, представляют собой линии элементов задержки, необходимые для учета динамики нелинейного объекта и процесса управления.

*Преимущество:* простота.

### Непрямая система управления.

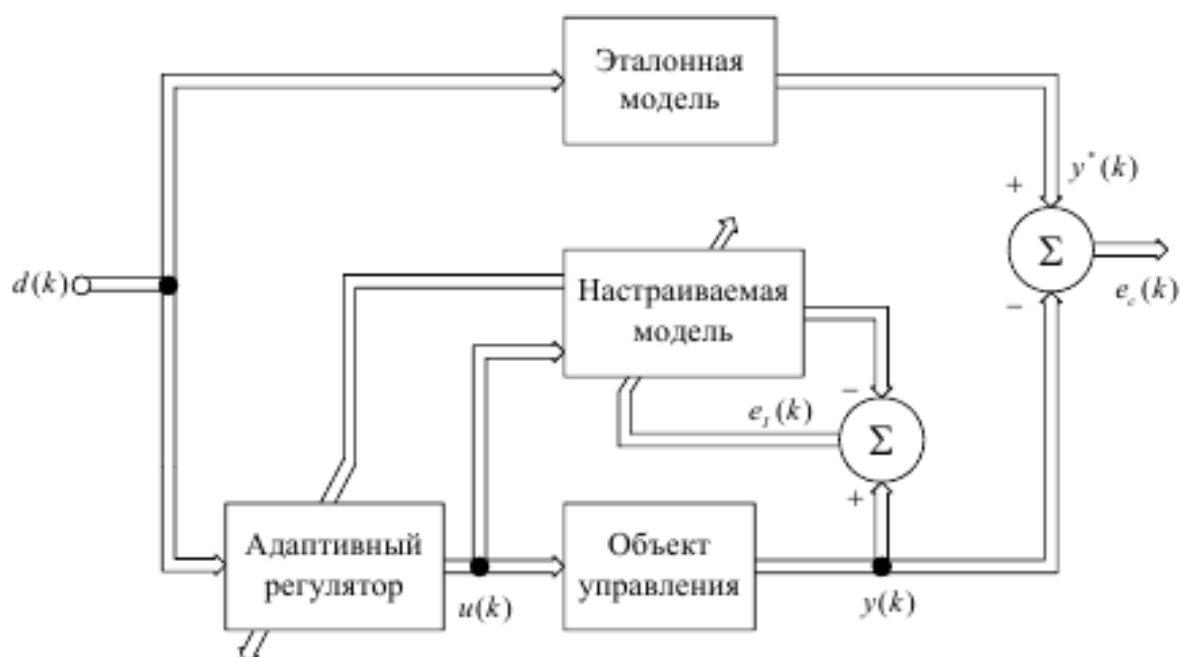


Рисунок 4 – Схема непрямого адаптивного управления

В адаптивной непрямой системе управления параллельно объекту подключена настраиваемая модель, параметры которой непрерывно уточняются с помощью какого-либо алгоритма адаптивного оценивания, минимизирующего в реальном времени целевую функцию от ошибки идентификации

$$e_I(k) = y(k) - y(k),$$



где  $y(k)$  – выход настраиваемой модели.

Получаемые параметры настраиваемой модели используются регулятором в качестве оценок параметров объекта управления. Ошибка управления определяется как  $e_c(k) = y^*(k) - y(k)$ . Чем точнее настраиваемая модель отслеживает поведение реального объекта, тем меньше ошибка идентификации  $e_I(k)$  и тем точнее выход объекта следует за желаемой траекторией  $y^*(k)$ . Качество работы адаптивной не прямой системы полностью определяется эффективностью процесса идентификации, поскольку  $e_c(k) = y^*(k) - y(k) - e_I(k)$  и процесс управления сводится к отслеживанию адаптивной моделью поведения эталонной модели.

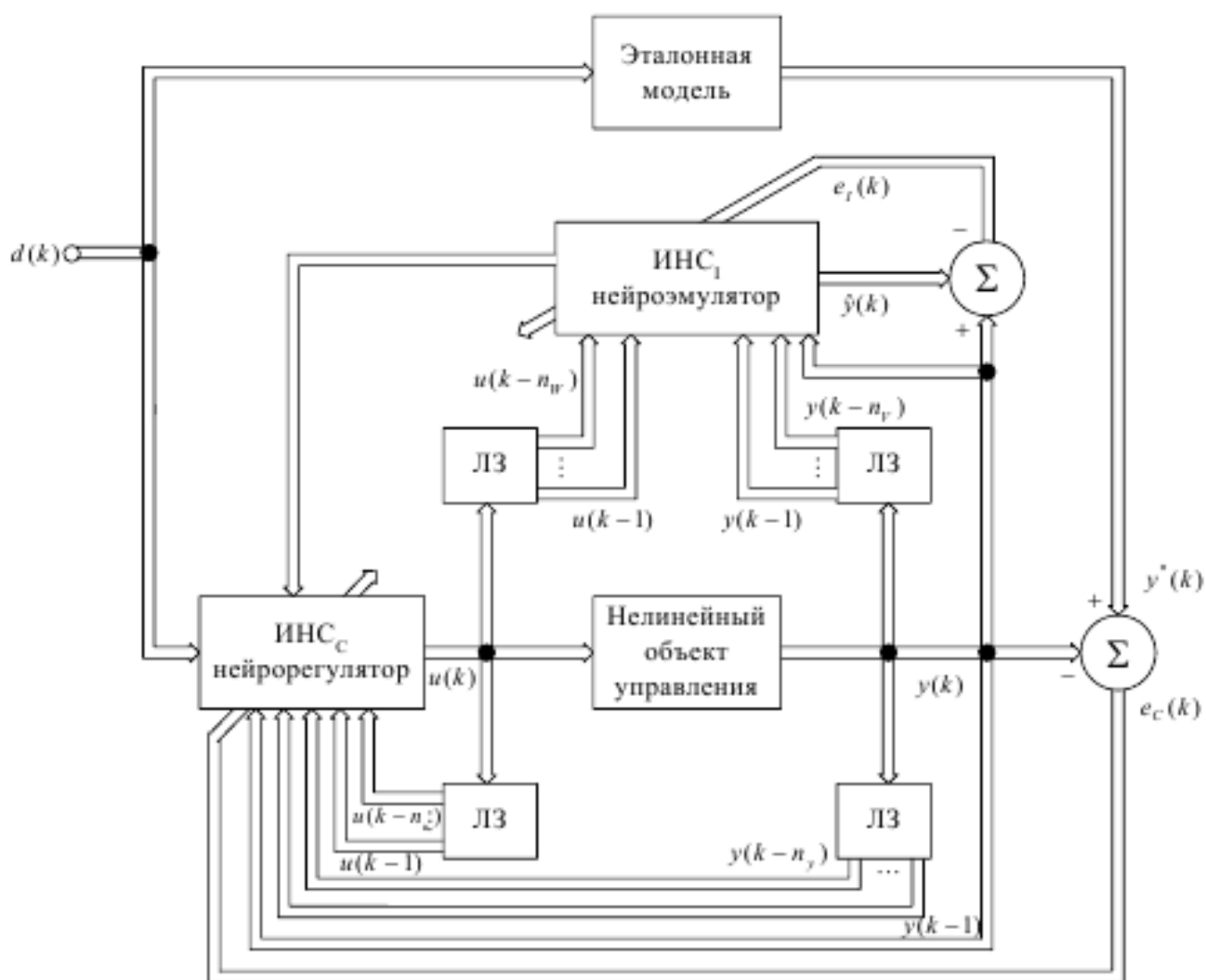


Рисунок 5 – Структура непрямого нейрорегулятора

Непрямой нейрорегулятор содержит две нейросети ИНС<sub>1</sub> и ИНС<sub>С</sub>, исполняющих функции нейроэмулятора и нейрорегулятора, обеспечивающего в процессе непрерывного обучения минимизацию принятого критерия управления.

*Преимущество:* высокое быстродействие и помехоустойчивость.

## СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

- 1 Гітіс В. Б. Теорія і практика застосування нейронних мереж: навчальний посібник. – Краматорськ: ДДМА, 2012. – 96 с
- 2 Ковалевский С.В., Гитис В.Б. Создание и применение нейронных сетей для решения прикладных задач: Учебно-методическое пособие для студентов специальностей «Технология машиностроения» и «Экономическая кибернетика». – Краматорск: ДГМА, 2005. – 80 с.
- 3 Методические указания к выполнению лабораторных работ и самостоятельной работе по курсу «Нейросетевые технологии» для студентов направления «Системный анализ» дневной формы обучения / Сост. Гитис В.Б. – Краматорск: ДГМА, 2013. – 20 с.
- 4 Дубровін В. І., Субботін С. О. Методи оптимізації та їх застосування в задачах навчання нейронних мереж: Навчальний посібник. – Запоріжжя: ЗНТУ, 2003. – 136 с.
- 5 Ковалевский С.В., Гитис В.Б. Создание и применение нейронных сетей для решения прикладных задач: Учебно-методическое пособие для студентов специальностей «Технология машиностроения» и «Экономическая кибернетика». – Краматорск: ДГМА, 2005. – 80 с.
- 6 Комашинский В. И., Смирнов Д. А. Нейронные сети и их применение в системах управления и связи. – М.: Горячая линия – Телеком, 2003. – 94 с.
- 7 Круглов В. В., Борисов В. В. Искусственные нейронные сети. Теория и практика. – 2-е изд., стереотипное. – М.: Горячая линия – Телеком, 2002. – 382 с.
- 8 Медведев В. С., Потемкин В. Г. Нейронные сети. MATLAB 6 / Под общ. ред. к. т. н. В. Г. Потемкина. – М.: ДИАЛОГ-МИФИ, 2002. – 496 с.
- 9 Ахапкин Ю. К., Барцев С. И., Всеволодов Н. Н. и др. Биотехника – новое направление компьютеризации. – М.: Наука, 1990. – 144 с.
- 10 Боровиков В. П. Нейронные сети. Statistica neural networks: методология и технологии современного анализа данных. – М.: Горячая линия-Телеком. – 2008. – 392 с.

- 11 Вороновский Г. К. и др. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности. – Х.: ОСНОВА, 1997. – 112 с.
- 12 Галушкин А. И. Нейрокомпьютеры. Кн. 3: Учеб. пособие для вузов / Общая ред. А. И. Галушкина. – М.: ИПРЖР, 2000. – 528 с.
- 13 Галушкин А. И. Теория нейронных сетей. Кн. 1: Учеб. пособие для вузов / Общая ред. А. И. Галушкина. – М.: ИПРЖР, 2000. – 416 с.
- 14 Горбань А. Н. Обучение нейронных сетей. – М.: СП ПараГраф, 1991. – 394 с.
- 15 Евтихий Н. Н., Оныкий Б. Н., Перепелица В. В., Щербаков И. Б. Математические модели и оптические реализации многослойных и полиномиальных нейронных сетей. – М.: Препринт/МИФИ, 1994. – 32 с.
- 16 Заенцев И. В. Нейронные сети: основные модели. Учебное пособие. – Воронеж: ВГУ, 1999. – 76 с.
- 17 Миркес Е. М. Нейроинформатика: Учебное пособие. – Красноярск: КГТУ, 2002. – 120 с.
- 18 Руденко О. Г., Бодянський Є. В. Штучні нейронні мережі: Навчальний посібник. – Харків: ТОВ «Компанія СМІТ», 2006. – 404 с.
- 19 Уоссермен Ф. Нейрокомпьютерная техника: теория и практика. – М.: Мир, 1992. – 538 с.