

## МОДЕЛИ НЕЙРОННЫХ СЕТЕЙ, ИСПОЛЬЗУЮЩИЕ МНОЖИТЕЛИ ЛАГРАНЖА, В ЗАДАЧЕ ОЦЕНИВАНИЯ СОСТОЯНИЯ ЭЭС

М.В. Хохлов

Оценивание состояния электроэнергетических систем заключается в расчете (оценке) параметров режима ЭЭС по данным телеметрических измерений, поступающих от энергообъектов в оперативные информационно-управляющие комплексы АО-энерго, ОДУ и ЦДУ, и относится к классу нелинейных многопараметрических задач на экстремум [1]. Широкое использование при оперативном управлении результатов оценивания состояния предъявляет высокие требования к быстродействию. Темп решения задач реального времени очень высок. Объем телеинформации, подлежащей обработке, достигает 1000 и более аналоговых телеизмерений при цикле поступления 1-10 с. В то же время уровень методических и алгоритмических разработок достиг состояния, когда основным ограничивающим фактором является производительность ЭВМ.

В настоящее время немалое внимание уделяется разработке и исследованиям нейросетевых методов решения различного рода задач. Это объясняется, во-первых, способностью нейронных сетей (НС) к обучению и обобщению, а во-вторых, присущей им массово параллельной обработке информации, что позволяет строить вычислительные системы высокой производительности. В [2] показано, как организовать параллельные вычисления по оцениванию состояния ЭЭС путем отображения задачи на рекуррентную нейронную сеть. Предложена модель НС, использующая множители Лагранжа для обработки ограничений в форме равенств и неравенств. В данной работе рассматриваются вопросы устойчивости такой НС и предлагается расширенная нейронная сеть, обладающая предельной точкой сходимости при ослабленных требованиях к виду целевой функции оценивания.

**Постановка задачи.** При оценивании состояния ЭЭС минимизируется некоторая функция невязки между измеренными значениями параметров режима и их оценками относительно вектора состояния  $x$ , компонентами которого обычно выступают модули и фазы узловых напряжений:

$$E(x) = \sum_{i=1}^m \rho \left( \frac{v_i(x) - \bar{v}_i}{\sigma_i} \right) \rightarrow \min \quad (1)$$

при ограничениях

$$c_j(x) = 0, \quad j = 1, 2, \dots, p, \quad (2)$$

$$g_k(x) \leq 0, \quad k = 1, 2, \dots, q, \quad (3)$$

где  $\rho(r)$  – произвольная функция невязки, в частном случае – строго выпуклая квадратичная  $\rho(r) = r^2/2$ ;  $\bar{v}_i$  и  $\sigma_i^2$  – значение и дисперсия ошибки  $i$ -го измерения;  $v_i(x)$ ,  $c_j(x)$  и  $g_k(x)$  – нелинейные функции, связывающие параметры режима законами электрических цепей.

Наложение ограничений (2) необходимо для выдерживания точных измерений, например нулевых инъекций в транзитных узлах, при сохранении численной устойчивости алгоритмов оценивания. Обработка точных измерений введением их в (1) с малой дисперсией чревата ухудшением обусловленности задачи. Ограничения (3) отражают технически возможные пределы изменения измеряемых и неизменяемых параметров режима: производительности электростанций, диапазонов регулирования коэффициента трансформации трансформаторов, значений нагрузок и др. Введение подобных ограничений целесообразно при наличии в расчетной схеме ЭЭС областей, не обеспе-

ченных в достаточном объеме телеизмерениями. Устойчивость нейронной сети для задачи (1) показана в [3]. Ниже исследуются свойства НС для решения задач (1) – (3).

**Отображение на нейронную сеть.** Поиск оценки  $\mathfrak{E}$ , доставляющей минимум функции (1) при ограничениях (2) и (3), сводится к нахождению седловой точки  $(\lambda, \lambda^*, \pi)$  функции Лагранжа

$$L(x, \lambda, \pi) = E(x) + \lambda^T c(x) + \pi^T g(x), \quad (4)$$

где  $\lambda$  и  $\pi$  – вектора неопределенных множителей Лагранжа, причем  $\pi_k \geq 0$ ,  $k = 1, 2, \dots, q$ . Алгоритм, предложенный в [2], состоит в решении последовательности подзадач для линеаризованной модели ЭЭС. Аппроксимируем зависимости  $v(x)$ ,  $c(x)$  и  $g(x)$  разложением в ряд Тейлора 1-го порядка в точке исходного приближения  $x(s)$ . Тогда функция Лагранжа принимает вид

$$L'(\Delta x, \lambda, \pi) = \sum_{i=1}^m \rho \left[ \sigma_i^{-1} \left( \frac{\partial v_i}{\partial x} \Delta x + b_i \right) \right] + \sum_{j=1}^p \lambda_j \left( \frac{\partial c_j}{\partial x} \Delta x + c_j(x) \right) + \sum_{k=1}^q \pi_k \left( \frac{\partial g_k}{\partial x} \Delta x + g_k(x) \right), \quad \pi_k \geq 0 \quad (5)$$

где  $b_i = v_i(x) - \bar{v}_i$ . Если  $\Delta x^*(s)$  есть решение линеаризованной подзадачи, минимизирующее функцию Лагранжа (5), а  $\lambda^*(s)$ ,  $\pi^*(s)$  – отвечающие этому решению множители, то очередное приближение определяется как

$$x(s+1) = x(s) + \mu \Delta x^*(s), \quad \mu > 0,$$

$$\lambda(s+1) = \lambda^*(s),$$

$$\pi(s+1) = \pi^*(s)$$

и является исходным для следующей линеаризации.

Минимизацию  $L'(\Delta x, \lambda, \pi)$  по  $\Delta x$  можно рассматривать как минимизацию энергетической функции рекуррентной нейронной сети (рис.1) в процессе ее конвергенции [2]. Синаптические веса НС являются соответствующими компонентами матриц частных производных:  $R^{-1/2} \frac{\partial v}{\partial x}$ ,  $\frac{\partial c}{\partial x}$  и  $\frac{\partial g}{\partial x}$ , где  $R = \text{diag}(\sigma^2)$  – ковариационная матрица ошибок телеизмерений. Выведенная из состояния равновесия НС стремится к устойчивому состоянию, при этом вектор, характеризующий состояние НС представляет собой вектор искомым переменных  $\{\Delta x, \lambda, \pi\}$  линеаризованной задачи оценивания состояния ЭЭС. Динамика нейронной сети описывается системой дифференциальных уравнений:

$$\begin{cases} \frac{\partial \Delta x}{\partial t} = -\nabla_{\Delta x} L'(\Delta x, \lambda, \pi), \\ \frac{\partial \lambda}{\partial t} = \nabla_{\lambda} L'(\Delta x, \lambda, \pi), \\ \frac{\partial \pi}{\partial t} = \nabla_{\pi} L'(\Delta x, \lambda, \pi), \end{cases} \quad (6)$$

которые представляют непрерывный аналог градиентного метода, приложенного к функции Лагранжа (5).

**Устойчивость нейронной сети.** В [4] доказано, что если матрица вторых частных производных функции Лагранжа  $\nabla_{\Delta x \Delta x}^2 L'(\Delta x^*, \lambda^*, \pi^*)$  положительно определена, то решение системы (6) приводит к седловой точке функции Лагранжа (5). Иными словами, нейронная сеть (6) сходится к решению  $\Delta x^*$ , если функция невязки  $\rho(r)$  строго выпук-

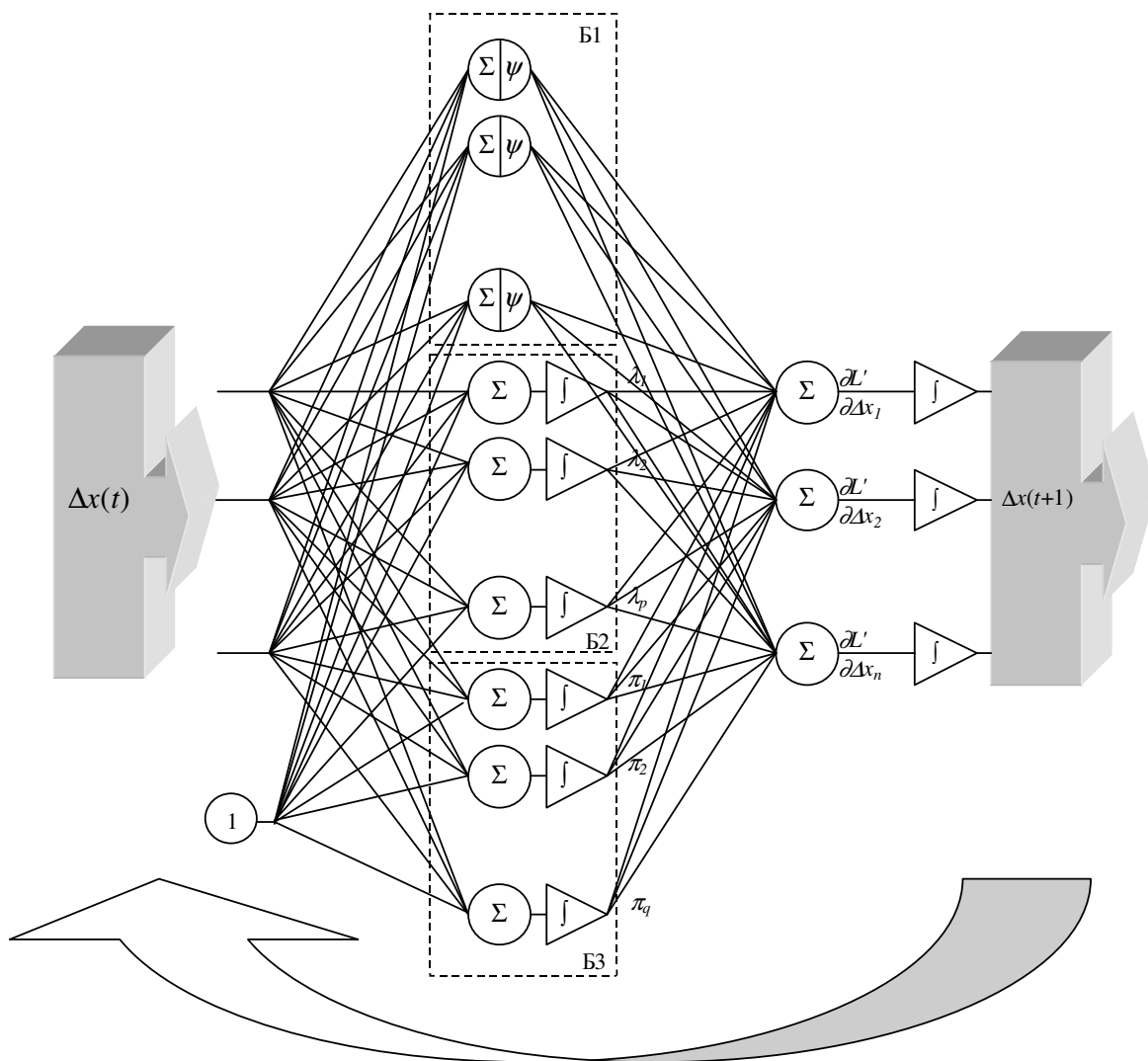


Рис.1. Структура рекуррентной нейронной сети для функции Лагранжа (5):  
 Б1 – слой обработки невязок, Б2 – слой оценки множителей  $\lambda$ , Б3 – слой оценки множителей  $\pi$ .

лая, или функция активации нейронов  $\psi = \rho'$  строго возрастающая. Среди различных  $\rho$ -функций практическое применение при оценивании состояния ЭЭС получила квадратичная  $\rho(r) = r^2/2$ , обеспечивающая статистически оптимальные несмещенные оценки параметров режима при нормальном распределении ошибок телеизмерений. Нейронная сеть в таком случае имеет линейную функцию активации  $\psi(r) = r$  и всегда обладает точкой равновесия, являющейся седловой точкой  $(\Delta x^*, \lambda^*, \pi^*)$  функции Лагранжа (5).

Наибольший интерес представляет задача оценивания состояния ЭЭС по неквадратичным критериям. Использование  $\rho$ -функций, первая производная которых  $\psi = \rho'$  ограничена, позволяет получать оценки, устойчивые к грубым ошибкам в телеизмерениях. Причинами грубых ошибок являются сбои в системе сбора и передачи телеметрической информации и т.п., что приводит к значительному смещению оценок наименьших квадратов. Примером неквадратичной функции может служить функция Хьюбера [5]:

$$\rho(r) = \begin{cases} r^2/2 & \text{при } |r| < a, \\ a|r| - a^2/2 & \text{при } |r| \geq a, \end{cases} \quad (7)$$

где  $a$  – положительная константа, определяющая робастные свойства оценок. Соответствующая функция активации нейронов имеет вид

$$\psi(r) = \begin{cases} r & \text{при } |r| < a, \\ \text{sgn}(r) \cdot a & \text{при } |r| \geq a. \end{cases} \quad (8)$$

Устойчивость НС с  $\psi$ -функцией (8) не гарантируется. Решением системы (6) в общем случае является предельный цикл, состоящий из одной точки лишь при определенных условиях. Это означает, что нейронная сеть может не достигнуть состояния равновесия.

**Нейронная сеть с  $u$ -нейронами.** Для обеспечения сходимости градиентного процесса в задачах с выпуклой целевой функцией и линейными ограничениями в форме неравенств в [4] предлагается изменить функцию Лагранжа. В этом случае линейризованные

ограничения  $\frac{\partial g_k}{\partial x} \Delta x + g_k(x) \leq 0$  следует заменить эквивалентными

$u_k \left( \frac{\partial g_k}{\partial x} \Delta x + g_k(x) \right) \leq 0$ , где  $u_s(z)$  есть строго возрастающая, строго выпуклая функция

одного переменного с  $u_k(0) = 0$ . Пример такой функции  $u_k(z) = e^{\eta_k z} - 1$ , где  $\eta_k > 0$ . Доказано, что градиентный процесс с  $u$ -модификацией сходится к предельной точке [4].

Не трудно показать, что соответствующие изменения, вносимые в нейронную сеть, минимальные. Достаточно заменить нейроны скрытого слоя, имеющие структуру "сумматор-интегратор", на  $u$ -нейроны со структурой "сумматор – функция  $y(z) = u(z)$  – интегратор – функция  $y(z) = du/dz$ ". Как показали экспериментальные расчеты с различными  $u$ -функциями, такой подход имеет множество недостатков. Помимо медленной сходимости с плохого начального приближения отметим сложности в дальнейшей реализации  $u$ -нейронов при аппаратном исполнении НС.

**Расширенная нейронная сеть.** Наиболее эффективным оказался иной подход. Синтез устойчивой нейронной сети для линейризованной задачи оценивания состояния ЭЭС с ограничениями в форме равенств и неравенств будем осуществлять на основании модифицированной функции Лагранжа [6]:

$$L''(\Delta x, \lambda, \pi) = L'(\Delta x, \lambda, \pi) + \frac{\alpha}{2} \sum_{j=1}^p \left( \frac{\partial c_j}{\partial x} \Delta x + c_j(x) \right)^2 + \frac{\beta}{2} \sum_{k=1}^q \max \left( 0, \frac{\partial g_k}{\partial x} \Delta x + g_k(x) \right)^2, \quad (9)$$

где  $\alpha, \beta$  – положительные параметры штрафа. Добавление штрафных слагаемых изменяет свойства матрицы вторых производных функции Лагранжа, не нарушая при этом стационарности точки оптимума  $\Delta x^*$ . Существуют конечные значения  $\bar{\alpha}$  и  $\bar{\beta}$ , такие, что для любых  $\alpha > \bar{\alpha}$  и  $\beta > \bar{\beta}$  матрица  $\nabla_{\Delta x}^2 L''(\Delta x^*, \lambda^*, \pi^*)$  будет положительно определенной [6]. Это гарантирует устойчивость НС при использовании неквадратичной функции оценивания.

Структура расширенной нейронной сети, которая реализует градиентный метод, приводящий к седловой точке функции  $L''(\Delta x, \lambda, \pi)$ , представлена на рис.2. Отличие от предыдущей модели на рис.1 состоит в появлении дополнительных блоков нейронов Б4 и Б5 в скрытом слое НС. Один (Б4) соответствует штрафному слагаемому для огра-

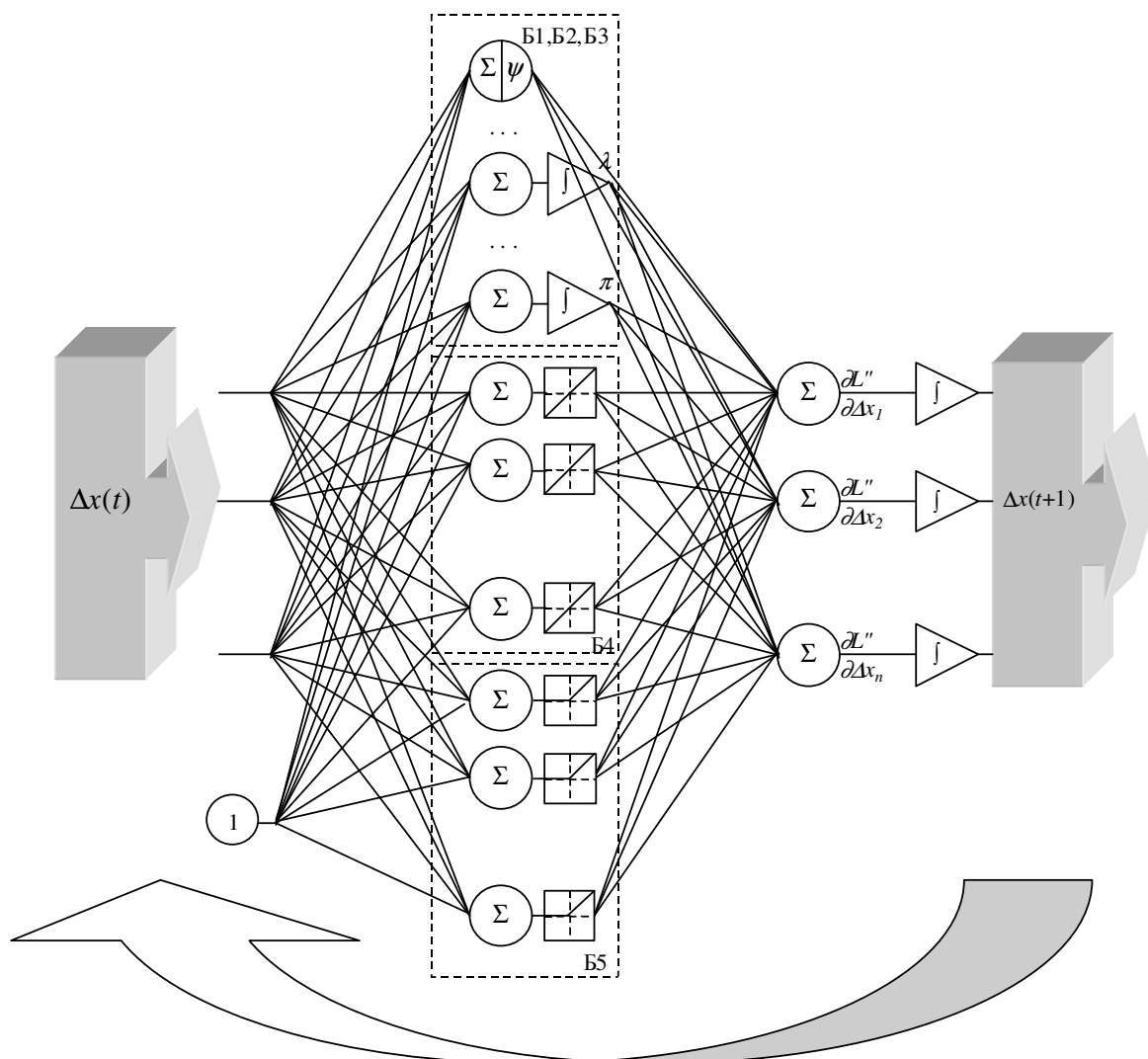


Рис.2. Структура расширенной нейронной сети для модифицированной функции Лагранжа (9).

ничений в форме равенств и состоит из нейронов со структурой "сумматор – линейная функция  $y(z) = \alpha z$ ". Другой (Б5) содержит нейроны со структурой "сумматор – функция  $y(z) = \beta \max\{0, z\}$ " и отвечает за ограничения в форме неравенств. Заметим, что нейроны этих блоков используют те же входные сигналы, что и соответствующие нейроны блоков оценки множителей Б2 и Б3. При реализации НС каждая такая пара может быть представлена одним комплексным нейроном.

Экспериментальные расчеты показали, что расширенная нейронная сеть сходится и с невыпуклыми  $\rho$ -функциями. Применение последних эффективнее для подавления влияния грубых ошибок в телеизмерениях, чем функция (7). Однако наличие множества локальных минимумов целевой функции предполагает их использование только при наличии хорошего начального приближения, полученного с использованием критерия (7).

**Иллюстративный пример.** Работу нейросетевых методов покажем на примере оценивания состояния трехузловой ЭЭС (рис.3) по неквадра-

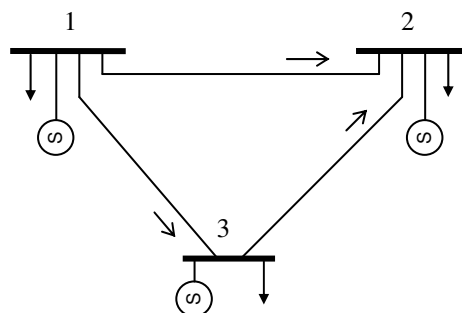


Рис.3. Схема трехузловой ЭЭС.

тичному критерию (7). При этом будем рассматривать упрощенную активную модель схемы на рис. 3, в которой активные потери в линиях приняты нулевыми. В таком случае уравнения, связывающие параметры режима, линейные и расчет оценок будет осуществляться лишь с помощью нейронной сети. Вектор измеряемых параметров  $v$  образуют перетоки мощности в линиях  $P_{1-2}$ ,  $P_{1-3}$ ,  $P_{2-3}$ . В качестве переменных  $x$  выступают фазы напряжений  $\delta_2$  и  $\delta_3$  (узел 1 – базисный). Дисперсии ошибок всех измерений приняты равными единице  $\sigma^2 = \sigma_i^2 = 1$ .

Рассмотрим ситуацию, когда два измерения  $P_{1-2}$  и  $P_{1-3}$  являются неверными. Для этого изменим знак их значений на противоположный. Такая ситуация встречается в практике оценивания состояния ЭЭС. Причина – неправильное подключение измерительных устройств на обратную полярность, либо нереверсивность используемых датчиков. Для приведенной схемы ЭЭС характерна низкая избыточность измерений. Поскольку любые два измерения образуют критическую пару, трех измерений недостаточно, чтобы идентифицировать неверные измерения. В моделируемой ситуации невозможно установить и сам факт сбоя в данных. Это видно по результатам выполненного оценивания состояния: остаточные невязки малы и не превышают величины  $3\sigma$  (см. таблицу).

Таблица

Результаты оценивания состояния

Параметр $v_i$	Остаточные невязки ( $\bar{v}_i - v_i$ )	
	без учета ограничений	с учетом ограничений
$P_{1-2}$	-1,92	-14,0
$P_{1-3}$	0,8	-17,1
$P_{2-3}$	0,4	6,0

Допустим, что известны пределы на значения инъекций в узлах 2 и 3:

$$P_i^\Sigma \leq \bar{P}_i^\Sigma, \quad i=2,3,$$

где  $P_i^\Sigma$  – суммарная мощность, выдаваемая из  $i$ -го узла в систему. Как дополнительная информация, эти ограничения увеличивают избыточность данных, что уже позволяет усомниться в достоверности измерений. Из таблицы видно, что оценки параметров режима, полученные в результате ограничения области решения, не соответствуют в точности действительным значениям. В частности, оценка достоверного измерения  $P_{2-3}$  оказывается смещенной. Однако это позволяет диагностировать сбой в измерениях. Остаточные невязки превышают величину  $3\sigma$ , сигнализируя о плохих данных в измерениях.

На рис.4 показаны траектории сходимости НС, использующих множители Лагранжа, в пространстве переменных состояния  $\pi_2$ ,  $\delta_2$  и  $\delta_3$  (ось множителя  $\pi_3$ , соответ-

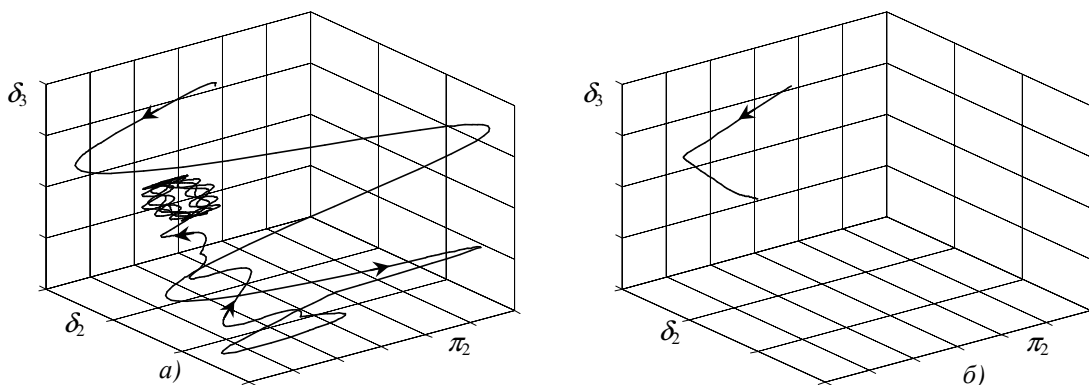


Рис.4. Траектории сходимости нейронных сетей: а) НС на базе функции Лагранжа (5), б) НС на базе модифицированной функции Лагранжа (9).

вующего второму ограничению, по понятным причинам не показана). Решение получено численным интегрированием системы дифференциальных уравнений (6), описывающих динамику НС. Использование функции Лагранжа (5) приводит к устойчивому колебательному процессу в нейронной сети. Сеть сходится к периодическому решению (рис. 4а). В то же время расширенная нейронная сеть сходится к седловой точке модифицированной функции Лагранжа (9) – решению задачи оценивания состояния с учетом ограничений (рис. 4б).

### Заключение

Разработанные модели НС, использующие множители Лагранжа, позволяют решать задачу оценивания состояния ЭЭС с учетом ограничений в форме равенств и неравенств. Устойчивость рекуррентных нейронных сетей, являющаяся необходимым условием сходимости вектора искомых переменных к решению, зависит от применяемого критерия оценивания. В случае критерия наименьших квадратов устойчивая НС конструируется исходя из функции Лагранжа (5). В задаче оценивания состояния по неквадратичным критериям НС предлагается строить на базе модифицированной функции Лагранжа (9). Не менее важным является вопрос о скорости сходимости и быстродействия НС. Это требует проведения дальнейших исследований. Окончательный вывод об эффективности их применения может быть сделан после разрешения проблем, связанных с аппаратной реализацией нейросетевых моделей.

### ЛИТЕРАТУРА

1. **Оценивание состояния** в электроэнергетике / А.З. Гамм, Л.Н. Герасимов, И.Н. Колосок и др. – М.: Наука, 1983. – 320 с.
2. **Новые информационные технологии** в задачах оперативного управления электроэнергетическими системами / Н.А. Манов, Ю.Я. Чукреев, М.И. Успенский и др. – Екатеринбург: УрО РАН, 2002. – 205 с.
3. **Хохлов М.В.** Устойчивость динамической нейронной сети для оценивания состояния электроэнергетических систем // Информационные технологии в электротехнике и электроэнергетике: Материалы IV всерос. науч.-техн. конф. – Чебоксары: Изд-во Чуваш. ун-та, 2002. – С. 166-169.
4. **Эрроу К.Дж., Гурвиц Л. Удзава Х.** Исследования по линейному и нелинейному программированию: Пер. с англ. – М.: ИЛ, 1962. – 335 с.
5. **Хьюбер Дж.П.** Робастность в статистике: Пер. с англ. – М.: Мир, 1984. – 304 с.
6. **Гилл Ф., Мюррей У., Райт М.** Практическая оптимизация: Пер. с англ. – М.: Мир, 1985. – 509 с.

### ДИСКУССИЯ

**П.И. Бартоломей.** Вывод о том, что эффективность предлагаемого подхода определит именно аппаратная реализация нейронных сетей, может быть, не совсем точен. Дело в методе Лагранжа, а не в нейронных сетях. Нейронные сети это, наверное, вещь перспективная и дело будущего. Вся проблема как раз в методе Лагранжа, который туда закладывается. Ведь метод Лагранжа – это седловая точка, и любые градиентные методы от седловой точки нас уводят в сторону, они никак не притягивают даже в какую-то область. Чем объясняется работоспособность предложенного метода? И второй вопрос: метод Лагранжа предполагает, что Вы перешли от неравенств к равенствам. Значит, Вы либо должны знать активные ограничения заранее, либо, если Вы все ограничения превратите в активные, они могут быть несовместными.

**М.В. Хохлов.** Способность градиентного метода получить седловую точку объясняется особенностью его применения к функции Лагранжа. В седловой точке функции Лагранжа  $L$  (рис.5) достигается минимум по  $\Delta x$  и максимум по  $\lambda$ , поэтому предлагается

искать седловую точку двигаясь по антиградиенту относительно изменений  $\Delta x$  и по градиенту (!) относительно изменений  $\lambda$ . Иными словами, одновременно с минимизацией функции  $L$  по  $\Delta x$  осуществлять ее максимизацию по  $\lambda$ :

$$\max_{\lambda} \min_{\Delta x} L(\Delta x, \lambda).$$

Градиентный метод, реализующий этот процесс и описываемый системой (6), можно рассматривать как исходно-двойственный, в том смысле, что он действует одновременно в пространстве исходных переменных  $\Delta x$  и в пространстве множителей  $\lambda$ . Метод был предложен для задачи с ограничениями в форме равенств в [4], где доказано, что при линейных ограничениях и строго выпуклой целевой функции процесс сходится к седловой точке.

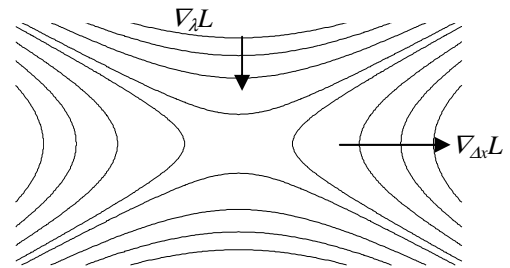


Рис.5. Седловая точка функции Лагранжа.

На рис. 6 демонстрируется работа метода при минимизации неквадратичной функции невязок  $E(x)$  для линейной системы с двумя неизвестными  $x$  и двумя ограничениями в форме неравенств  $g(x) \leq 0$ . Для наглядности используется функция Лагранжа  $L(x) = E(x) + \pi^T g(x)$  без штрафного слагаемого. В допустимой области  $\nabla_{\pi} L = g(x) \leq 0$ , следовательно, множители  $\pi(t) = \max(0, \pi(t)) = 0$  и второе слагаемое функции Лагранжа  $\pi^T g(x) = 0$ . Направление движения определяется лишь градиентом  $\nabla E$ . В зоне нарушения ограничений, поскольку  $\nabla_{\pi} L > 0$ , происходит рост соответствующих множителей  $\pi$  и в определение направления движения включается второе слагаемое  $\pi^T g(x)$ , влияние которого увеличивается до тех пор, пока процесс не вернется в допустимую область. На рис.6 хорошо наблюдается колебательный процесс сходимости множителя  $\pi_1$  к его оценке, определяющей минимизацию целевой функции  $E$  вдоль границы  $g_1(x) = 0$ .

Метод тестировался на большом количестве задач с ограничениями в форме равенств и неравенств, в том числе для оценивания состояния 24-узловой ЭЭС. Для нахождения квадратичных оценок использовалась функция Лагранжа, в то время как для неквадратичных оценок использовалась модифицированная функция Лагранжа. Во

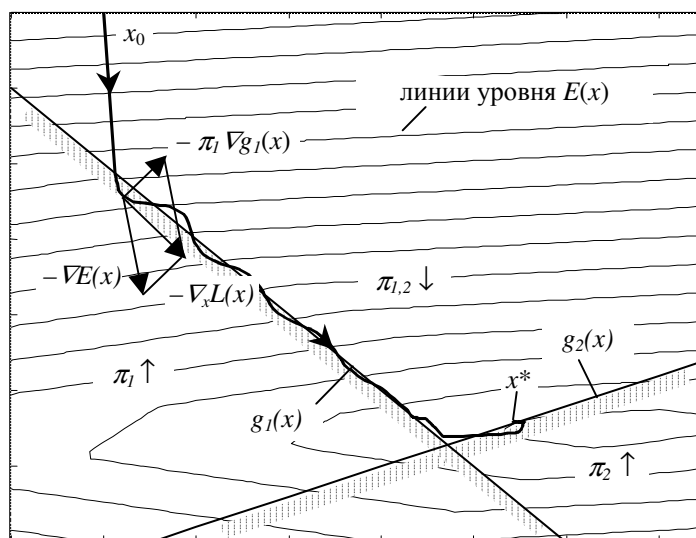


Рис.6. К методу градиента для функции Лагранжа.



всех случаях вычислительный процесс сходился к минимуму целевой функции, удовлетворяющему ограничениям.

Проблема несовместности ограничений в работе не рассматривалась. В проводимых исследованиях ограничения накладывались на независимые переменные. А именно ограничения вида  $P_i = 0$  и  $Q_i = 0$  для нулевых инъекций транзитных узлов и ограничения вида  $0 \leq P_j \leq P_j^{\max}$  и  $Q_j^{\min} \leq Q_j \leq Q_j^{\max}$  для мощностей нагрузочных узлов. В этом случае все ограничения совместны.

В общем случае, если задать также пределы на значения перетоков мощности по линиям и т.п., возникновение несовместности ограничений может приводить в процессе сходимости к неограниченному росту множителей Лагранжа. Чтобы избежать этого, для них, видимо, следует задать пределы изменения.