УДК 681.513

ОБУЧЕНИЕ ИСКУССТВЕННЫХ ВСПЛЕСК-НЕЙРОННЫХ СЕТЕЙ ПРИ ОБРАБОТКЕ НЕСТАЦИОНАРНЫХ СТОХАСТИЧЕСКИХ СИГНАЛОВ

БОДЯНСКИЙ Е.В., ВИНОКУРОВА Е.А.

Рассматривается группа адаптивных алгоритмов обучения всплеск-нейронных сетей, предназначенных для обработки нестационарных стохастических сигналов. Вводятся алгоритмы, обладающие как фильтрующими, так и следящими свойствами, анализируется их сходимость и показываются преимущества перед известными процедурами. Отмечается, что ряд известных алгоритмов обучения является частным случаем предложенных конструкций.

1. Введение

В настоящее время при обработке сигналов различной природы (прогнозирование, фильтрация, сглаживание, сегментация, сжатие, диагностика и т.п.) применяются технологии, основанные на использовании базисов всплесков (вейвлетов, wavelets) [1-4], кроме всего благодаря их кратномасштабным свойствам. Всплески имеют вид коротких, локализованных во времени (или пространстве) волновых пакетов с нулевым значением интеграла; обладают возможностью сдвига по оси времени; способны к масштабированию (сжатию-растяжению); имеют локальный частотный спектр.

Сочетание кратномасштабных свойств всплесков с универсальными аппроксимирующими свойствами искусственных нейронных сетей (ИНС) привело к появлению всплеск-нейронных сетей (wavelet neural networks) [5-10], представляющих собой по сути ИНС с прямой передачей информации и архитектурой, подобной широко распространенным радиально-базисным нейронным сетям [11-13].

Всплеск-нейронная сеть (ВНС) реализует нелинейное отображение вектора входов

$$x(k) = (x_1(k), x_2(k), \ldots, x_n(k))^T \in R^n$$

в вектор выходов $y(k) = (y_1(k), y_2(k), \ldots, y_m(k))^T \in R^m$:

$$y_j(k) = F_j(x(k)) = \sum_{i=1}^{h} w_{ji}\varphi_i(x(k)) = w_j^T\varphi(k),$$

$$j = 1,2,\ldots,m,$$

где $F_j(\bullet)$ — некоторый оператор, связывающий пространство входов с пространством выходов; $w_j = (w_{j1}, w_{j2}, \ldots, w_{jh})^T$ — $(h \times 1)$-вектор настраиваемых синаптических весов;

$\varphi(k) = (\varphi_1(x(k)), \varphi_2(x(k)), \ldots, \varphi_h(x(k)))^T$ — вектор функций активации нейронов скрытого слоя на основе всплесков; $n << h$ — число нейронов скрытого слоя; k — текущее дискретное время.

В качестве функций активации могут использоваться различные всплески, такие как всплеск Sinc, Хаара (wavelet Haar), Симлета (wavelet Symlet), всплески Добеши (wavelets Daubechies), Гаусса (wavelet Gaussian), Морлета (wavelet Morlet), “Мексиканская шляпа” (wavelet “Mexican hat”).

2. Алгоритмы обучения радиально-базисных ИНС

Для настройки синаптических весов радиально-базисных ИНС, являющихся прототипом ВНС, наиболее широко применяются алгоритмы, минимизирующие критерий обучения, связанный с квадратами рассогласований между внешним обучающим сигналом $d_j(k)$ и выходом сети $y_j(k)$.

Простейшей из таких процедур является алгоритм Уидроу-Хоффа [14-16], минимизирующий на каждом такте локальный критерий

$$E_j(k) = \frac{1}{2}(d_j(k) - y_j(k))^2 = \frac{1}{2}e_j^2(k)$$

и имеющий вид

$$w_j(k+1) = w_j(k) + \frac{d_j(k+1) - w_j^T(k)\varphi(k+1)}{\|\varphi(k+1)\|^2} \times$$

$$\times \varphi(k+1) = w_j(k) + \frac{e_j(k+1)\varphi(k+1)}{\|\varphi(k+1)\|^2}. \qquad (1)$$

Обладая высоким быстродействием, а следовательно, и выраженными следящими свойствами, этот алгоритм плохо работает в условиях помех, искажающих обрабатываемые сигналы, в связи с чем при работе со стохастическими процессами чаще применяется метод наименьших квадратов в различных модификациях [17-23].

Алгоритмы, связанные с методом наименьших квадратов, минимизируют критерий

$$E_j^k = \sum_{p=1}^{k} E_j(p) = \frac{1}{2}\sum_{p=1}^{k} e^2(p), \qquad (2)$$

заданный на всей обучающей выборке, и фактически обеспечивают решение системы нормальных уравнений

$$\nabla_w E_j^k = -r_j(k) + R(k)w_j(k) = 0, \qquad (3)$$

имеющее вид

$$w_j(k) = R^{-1}(k)r_j(k),$$

где $r_j(k) = \sum_{p=1}^{k} d_j(p)\varphi(p)$, $R(k) = \sum_{p=1}^{k} \varphi(p)\varphi^T(p)$.

Чтобы реализовать алгоритм в реальном времени, запишем искомый вектор синаптических весов для $(k+1)$-го такта времени в виде

$$w_j(k+1) = R^{-1}(k+1)r_j(k+1) =$$

$$= (R_j(k) + \varphi(k+1)\varphi^T(k+1))^{-1} \times$$

$$\times (r_j(k) + d_j(k+1)\varphi(k+1)),$$

после чего, применяя формулу Шермана-Моррисона

$$R^{-1}(k+1)=R^{-1}(k)-$$
$$-\frac{R^{-1}(k)\varphi(k+1)\varphi^T(k+1)R^{-1}(k)}{1+\varphi^T(k+1)R^{-1}(k)\varphi(k+1)}, \quad (4)$$

получим рекуррентную форму

$$w_j(k+1)=w_j(k)+R^{-1}(k)\times$$
$$\times\frac{d_j(k+1)-w_j^T(k)\varphi(k+1)}{1+\varphi^T(k+1)R^{-1}(k)\varphi(k+1)}\varphi(k+1)=$$
$$=w_j(k)+\frac{R^{-1}(k)e_j(k+1)\varphi(k+1)}{1+\varphi^T(k+1)R^{-1}(k)\varphi(k+1)}. \quad (5)$$

Соотношения (4), (5) описывают стандартный рекуррентный метод наименьших квадратов. Это наиболее распространенная процедура, используемая для обучения радиально-базисных ИНС, основним недостатком которой является численная громоздкость, проявляющаяся при работе в реальном времени с высокой частотой квантования сигналов и при большом числе нейронов скрытого слоя h.

По своим сглаживающим свойствам к рекуррентному методу наименьших квадратов близок алгоритм Гудвина-Рэмеджа-Кэйнеса [24, 25], существенно более простой в вычислительном отношении, однако имеющий низкую скорость сходимости. В принятых здесь обозначениях этот алгоритм может быть записан в форме

$$\begin{cases} w_j(k+1)=w_j(k)+(\mathrm{Tr}\,R(k+1))^{-1}(d_j(k+1)- \\ \quad -w_j^T(k)\varphi(k+1))\varphi(k+1), \\ \mathrm{Tr}\,R(k+1)=\mathrm{Tr}\,R(k)+\|\varphi(k+1)\|^2 . \end{cases} \quad (6)$$

По сути он является процедурой стохастической аппроксимации.

3. Модифицированные алгоритмы обучения

Используя для минимизации критерия (2) процедуру Гаусса-Ньютона, записываем

$$w_j(k+1)=w_j(k)-H^{-1}(k+1)\nabla_W E_j^{k+1}=$$
$$=w_j(k)+H^{-1}(k+1)\sum_{p=1}^{k+1}(d_j(p)-w_j^T(k)\varphi(p))\varphi(p)= \quad (7)$$
$$=w_j(k)+H^{-1}(k+1)(r_j(k+1)-R(k+1)w_j(k)),$$

где $H(k+1)=\sum_{p=1}^{k+1}\varphi(p)\varphi^T(p)=R(k+1)$ – гессиан принятого критерия обучения. Заметив, что выражение

$$r_j(k+1)-R(k+1)w_j(k)=r_j(k)+d_j(k+1)\varphi(k+1)-$$
$$-R(k)w_j(k)-\varphi(k+1)\varphi^T(k+1)w_j(k)$$

с учетом (3) превращается в мгновенную разность:

$$r_j(k+1)-R(k+1)w_j(k)=d_j(k+1)\varphi(k+1)-$$
$$-\varphi(k+1)\varphi^T(k+1)w_j(k),$$

можно записать еще одну форму рекуррентного метода наименьших квадратов:

$$\begin{cases} R^{-1}(k+1)=R^{-1}(k)- \\ \quad -\dfrac{R^{-1}(k)\varphi(k+1)\varphi^T(k+1)R^{-1}(k)}{1+\varphi^T(k+1)R^{-1}(k)\varphi(k+1)}, \\ w_j(k+1)=w_j(k)+R^{-1}(k+1)(d_j(k+1)- \\ \quad -w_j^T(k)\varphi(k+1))\varphi(k+1). \end{cases} \quad (8)$$

Используя в (7) вместо обратного гессиана скалярный параметр скорости обучения, можно ввести модифицированный алгоритм Гудвина-Рэмеджа-Кэйнеса:

$$\begin{cases} w_j(k+1)=w_j(k)+\eta_j(k+1)(r_j(k+1)- \\ \quad -R(k+1)w_j(k))=w_j(k)+(\mathrm{Tr}R(k+1))^{-1}\times \\ \quad \times(r_j(k+1)-R(k+1)w_j(k)), \\ R(k+1)=R(k)+\varphi(k+1)\varphi^T(k+1), \\ r_j(k+1)=r_j(k)+d_j(k+1)\varphi(k+1). \end{cases} \quad (9)$$

Низкая скорость обучения, присущая алгоритму (9), ставит задачу отыскания оптимальных значений параметра шага $\eta_j(k+1)$ вместо сугубо эмпирического $(\mathrm{Tr}\,R(k+1))^{-1}$.

Запишем первое соотношение (9) в виде

$$w_j(k+1)=w_j(k)+\eta_j(k+1)(r_j(k+1)-R(k+1)w_j(k))=$$
$$=w_j(k)+\eta_j(k+1)\sum_{p=1}^{k+1}(d_j(p)-w_j^T(k)\varphi(p))\varphi(p)$$

или, что то же самое,

$$\tilde{w}_j(k+1)=\tilde{w}_j(k)-\eta_j(k+1)\sum_{p=1}^{k+1}(d_j(p)-$$
$$-w_j^T(k)\varphi(p))\varphi(p)=\tilde{w}_j(k)- \quad (10)$$
$$-\eta_j(k+1)\sum_{p=1}^{k+1}\tilde{w}_j^T(k)\varphi(p)\varphi(p),$$

где $\tilde{w}_j(k)=w_j^*-w_j(k)-(h\times1)$ -вектор уклонений текущих значений синаптических весов $w_j(k)$ от своих гипотетических оптимальных значений w_j^*.

Вводя в рассмотрение норму $\|\tilde{w}_j(k+1)\|$ и соотношение, описывающее ее изменение:

$$\|\tilde{w}_j(k+1)\|^2=\|\tilde{w}_j(k)\|^2-$$
$$-2\eta_j(k+1)\sum_{p=1}^{k+1}\tilde{w}_j^T(k)\tilde{w}_j^T(k)\varphi(p)\varphi(p)+$$
$$+\eta_j^2(k+1)\sum_{p=1}^{k+1}(d_j(p)-w_j^T(k)\varphi(p))\varphi^T(p)\times \quad (11)$$
$$\times\sum_{p=1}^{k+1}(d_j(p)-w_j^T(k)\varphi(p))\varphi(p),$$

с учетом того, что

$$\sum_{p=1}^{k+1}(\tilde{w}_j^T(k)\varphi(p))^2 = \sum_{p=1}^{k+1} e_j^2(p) = \bar{e}_j^2(k+1) \qquad (12)$$

есть накопленная сумма мгновенных ошибок обучения, и решая дифференциальное уравнение

$$\partial\left\|\tilde{w}_j(k+1)\right\|^2 \Big/ \partial\eta_j = 0, \qquad (13)$$

находим оптимальное значение параметра скорости обучения в виде

$$\eta_j(k+1) = \bar{e}_j^2(k+1)\left\|r_j(k+1) - R(k+1)w_j(k)\right\|^{-2}.$$

Тогда алгоритм настройки синаптических весов ВНС может быть записан в форме

$$\begin{cases} w_j(k+1) = w_j(k) + \\ + \dfrac{\bar{e}_j^2(k+1)(r_j(k+1) - R(k+1)w_j(k))}{\left\|r_j(k+1) - R(k+1)w_j(k)\right\|^2}, \\ \bar{e}_j^2(k+1) = \bar{e}_j^2(k) + e_j^2(k+1), \end{cases} \qquad (14)$$

представляющей собой модификацию адаптивных алгоритмов одновременного действия, введенных в [26, 27].

4. Обучение в нестационарных условиях

В условиях, когда характеристики обрабатываемых сигналов изменяются во времени, в качестве критерия обучения достаточно часто используется взвешенная сумма квадратов ошибок

$$E_j^k = \sum_{p=1}^{k}\alpha^{k-p}E_j(p) = \frac{1}{2}\sum_{p=1}^{k}\alpha^{k-p}e_j^2(p), \qquad (15)$$

где $0 \le \alpha \le 1$ — параметр забывания устаревшей информации.

Минимизация (15) с помощью процедуры Гаусса-Ньютона приводит к экспоненциально-взвешенному рекуррентному методу наименьших квадратов:

$$\begin{cases} w_j(k+1) = w_j(k) + R^{-1}(k) \times \\ \times \dfrac{d_j(k+1) - w_j^T(k)\varphi(k+1)}{\alpha + \varphi^T(k+1)R^{-1}(k)\varphi(k+1)}\varphi(k+1) = \\ = w_j(k) + \dfrac{R^{-1}(k)e_j(k+1)\varphi(k+1)}{\alpha + \varphi^T(k+1)R^{-1}(k)\varphi(k+1)}, \\ R^{-1}(k+1) = \dfrac{1}{\alpha}\Big(R^{-1}(k) - \\ - \dfrac{R^{-1}(k)\varphi(k+1)\varphi^T(k+1)R^{-1}(k)}{\alpha + \varphi^T(k+1)R^{-1}(k)\varphi(k+1)}\Big), \quad 0 < \alpha \le 1, \end{cases} \qquad (16)$$

склонному к неустойчивости («взрыв параметров») при малых α и больших h.

В качестве альтернативы можно использовать либо экспоненциально-взвешенный алгоритм Гудвина-Рэмеджа-Кэйнеса [28]

$$\begin{cases} w_j(k+1) = w_j(k) + (\mathrm{Tr}\,R(k+1))^{-1} \times \\ \times (d_j(k+1) - w_j^T(k)\varphi(k+1))\varphi(k+1), \\ \mathrm{Tr}\,R(k+1) = \alpha\,\mathrm{Tr}\,R(k) + \left\|\varphi(k+1)\right\|^2, \quad 0 \le \alpha \le 1, \end{cases} \qquad (17)$$

либо версию процедуры (9) в виде

$$\begin{cases} w_j(k+1) = w_j(k) + (\mathrm{Tr}\,R(k+1))^{-1} \times \\ \times (r_j(k+1) - R(k+1)w_j(k)), \\ R(k+1) = \alpha R(k) + \varphi(k+1)\varphi^T(k+1), \\ r_j(k+1) = \alpha r_j(k) + d_j(k+1)\varphi(k+1), \quad 0 \le \alpha \le 1. \end{cases} \qquad (18)$$

Повторяя выкладки, аналогичные (10)-(13), можно получить экспоненциально-взвешенный алгоритм обучения со скалярным параметром скорости:

$$\begin{cases} w_j(k+1) = w_j(k) + \\ + \dfrac{\bar{e}_j^2(k+1)(r_j(k+1) - R(k+1)w_j(k))}{\left\|r_j(k+1) - R(k+1)w_j(k)\right\|^2 + \eta_j}, \\ \bar{e}_j^2(k+1) = \alpha\,\bar{e}_j^2(k) + e_j^2(k+1), \\ R(k+1) = \alpha R(k) + \varphi(k+1)\varphi^T(k+1), \\ r_j(k+1) = \alpha r_j(k) + d_j(k+1)\varphi(k+1), \\ 0 \le \alpha \le 1, \end{cases} \qquad (19)$$

где параметр $\eta_j \ge 0$ вводится в целях защиты от деления на ноль в окрестности оптимального вектора w_j^*.

Заметим также, что в отличие от процедуры (16), где параметр забывания α не может принимать нулевых значений, алгоритмы (17)-(19) при $\alpha = 0$ автоматически превращаются в одношаговый алгоритм Уидроу-Хоффа (1).

5. Исследование сходимости

Запишем первое соотношение (19) относительно вектора уклонений $\tilde{w}_j(k)$ при $\eta_j = 0$:

$$\tilde{w}_j(k+1) = \tilde{w}_j(k) - \frac{\tilde{w}_j^T(k)R(k+1)\tilde{w}_j(k)}{\left\|R(k+1)\tilde{w}_j(k)\right\|^2} \times \\ \times R(k+1)\tilde{w}_j(k) \qquad (20)$$

(здесь $R(k+1) = \sum_{p=1}^{k+1}\alpha^{k+1-p}\varphi(p)\varphi^T(p)$) и введем квадрат нормы

$$\left\|\tilde{w}_j(k+1)\right\|^2 = \left\|\tilde{w}_j(k)\right\|^2 - \frac{(\tilde{w}_j^T(k)R(k+1)\tilde{w}_j(k))^2}{\left\|R(k+1)\tilde{w}_j(k)\right\|^2}. \qquad (21)$$

Поскольку конструкция

$$(\tilde{w}_j^T(k)R(k+1)\tilde{w}_j(k))^2\left\|R(k+1)\tilde{w}_j(k)\right\|^{-2}$$

всегда неотрицательна, норма уклонений $\left\|\tilde{w}_j(k+1)\right\|$ в процессе обучения не может возрастать, при этом если $\alpha = 0$, (21) превращается в

$$\left\|\tilde{w}_j(k+1)\right\|^2 = \left\|\tilde{w}_j(k)\right\|^2 - \frac{(\tilde{w}_j^T(k)\varphi(k+1))^2}{\left\|\varphi(k+1)\right\|^2},$$

что совпадает с известным результатом для алгоритма Уидроу-Хоффа. Для случая же независимых центрированных входов ($R(k+1) = I$ — единичная матрица)

$$\frac{(\tilde{w}_j^T(k)I\tilde{w}_j(k))^2}{\left\|I\tilde{w}_j(k)\right\|^2} = \left\|\tilde{w}_j(k)\right\|^2 =$$

$$= \frac{\left\|\tilde{w}_j(k)\right\|^2 \left\|\varphi(k+1)\right\|^2}{\left\|\varphi(k+1)\right\|^2} \geq \frac{(\tilde{w}_j^T(k)\varphi(k+1))^2}{\left\|\varphi(k+1)\right\|^2},$$

т.е. алгоритм (19) по быстродействию всегда не хуже (1).

Рассмотрим процесс обучения в нестационарных условиях, что означает дрейф оптимального вектора синаптических весов:

$$w_j^*(k+1) = w_j^*(k) + \Delta w^*.$$

Тогда

$$\left\|\tilde{w}_j(k+1) + \Delta w^*\right\|^2 = \left\|\tilde{w}_j(k)\right\|^2 - (\tilde{w}_j^T(k) \times$$
$$\times R(k+1)\tilde{w}_j(k))^2 \left\|R(k+1)\tilde{w}_j(k)\right\|^{-2},$$

или

$$\left\|\tilde{w}_j(k)\right\|^2 - \left\|\tilde{w}_j(k+1)\right\|^2 = (\tilde{w}_j^T(k)R(k+1)\tilde{w}_j(k))^2 \times$$
$$\times \left\|R(k+1)\tilde{w}_j(k)\right\|^{-2} + 2\Delta w^{*T}\tilde{w}_j(k+1) + \left\|\Delta w^*\right\|^2. \qquad (22)$$

Подставляя (20) в (22), получаем условие сходимости в виде неравенства

$$\frac{(\tilde{w}_j^T(k)R(k+1)\tilde{w}_j(k))^2}{\left\|R(k+1)\tilde{w}_j(k)\right\|^2} > 2\frac{\tilde{w}_j^T(k)R(k+1)\tilde{w}_j(k)}{\left\|R(k+1)\tilde{w}_j(k)\right\|^2} \times$$
$$\times \tilde{w}_j^T(k)R(k+1)\Delta w^* - 2\tilde{w}_j^T(k)\Delta w^* - \left\|\Delta w^*\right\|^2,$$

с выполнением которого связано устойчивое слежение за изменяющимися характеристиками обрабатываемого сигнала.

Рассмотрим далее ситуацию, когда сигнал ошибки обучения $e_j(k)$ наблюдается на фоне помехи $\zeta_j(k)$ с ограниченным вторым моментом. Тогда, расписывая квадрат нормы (21) с учетом того, что

$$e_j(k+1) + \zeta_j(k+1) = \tilde{w}_j^T(k)\varphi(k+1) + \zeta_j(k+1),$$

после несложных, но громоздких преобразований получаем условие сходимости в виде

$$2M\left\{\tilde{w}_j^T(k)\sum_{p=1}^{k+1}\alpha^{k+1-p}(\tilde{w}_j^T(k)\varphi(p) + \zeta_j(p))\varphi(p)\right\} >$$
$$> M\left\{\sum_{p=1}^{k+1}\alpha^{k+1-p}(\tilde{w}_j^T(k)\varphi(p) + \zeta_j)^2\right\}$$

(здесь $M\{\bullet\}$ — символ математического ожидания) или с учетом некоррелированности $\varphi(k+1)$ и $\zeta_j(k+1)$ —

$$M\left\{\tilde{w}_j^T(k)R(k+1)\tilde{w}_j(k)\right\} > M\left\{\sum_{p=1}^{k+1}\alpha^{k+1-p}\zeta^2(p)\right\}. \qquad (23)$$

Таким образом, при наличии помех в канале обучения алгоритм обеспечивает сходимость в область, определяемую условием (23), размер которой определяется отношением сигнал / шум.

В ситуации, когда помехи проникают в скрытый слой со входа сети, т.е. вместо вектора $\varphi(k)$ обрабатывается сигнал $\psi(k) = \varphi(k) + \xi(k)$, условие сходимости может быть записано в виде

$$M\left\{\tilde{w}_j^T(k)R_\psi(k+1)\tilde{w}_j(k)\right\} > w_j^{*T}P_\xi(k+1)w_j^*, \qquad (24)$$

где $R_\psi(k+1) = M\left\{\sum_{p=1}^{k+1}\alpha^{k+1-p}\psi(p)\psi^T(p)\right\}$,

$$P_\xi(k+1) = M\left\{\sum_{p=1}^{k+1}\alpha^{k+1-p}\xi(p)\xi^T(p)\right\},$$

и, наконец, если в сети присутствуют помехи ζ_j и ξ, получаем

$$M\left\{\tilde{w}_j^T(k)R_\psi(k+1)\tilde{w}_j(k)\right\} >$$
$$> w_j^{*T}P_\xi(k+1)w_j^* + M\left\{\sum_{p=1}^{k+1}\alpha^{k+1-p}\zeta^2(p)\right\}. \qquad (25)$$

Оценки (22)-(25) показывают, что процедура обучения (19) по своим свойствам близка к алгоритму Уидроу-Хоффа, однако превосходит его в скорости, а кроме того, подобно методу наименьших квадратов, обладает фильтрующими свойствами.

6. Заключение

Рассмотренные алгоритмы обучения всплеск-нейронных сетей предназначены для настройки синаптических весов в условиях нестационарности и зашумленности обрабатываемых сигналов, обладают фильтрующими и следящими свойствами, просты в реализации и устойчивы при любых значениях параметра забывания.

Литература: 1. *Vatterli M., Kovacevic J.* Wavelets and Subband Coding. Englewood Cliffs, NJ: Prentice Hall, 1995. 489p. 2. *Mallat S.A.* Wavelet Tour of Signal Processing. San Diego: Academic Press, 1998. 635p. 3. *Daubechies I.* Ten Lectures on Wavelets. Philadelphia, PA: SIAM, 1992. 228p. 4. *Kugarajah T., Zhang Q.* Multidimensional wavelet frames // IEEE Trans. on Neural Networks. 1995. 6. P.1552-1556. 5. *Echauz J., Vachtsevanos G.* Elliptic and radial wavelet neural networks // Proc. Second World Automation Congress (WAC'96), Montpellier, France, 1996. 5. P.365-374. 6. *Zhang Q., Benveniste A.* Wavelet networks // IEEE Trans. on Neural Networks. 1992. 3. № 6, P.889-898. 7. *Katic D., Vukobratovic M.* Wavelet neural network approach for control of non-contact and contact robotic tasks // Proc. IEEE Symposium on Intelligent Control, 16-18 July 1997, Istambul. P.245-250. 8. *Liangyue Cao, Yiguang Hong, Haiping Fang.* Predicting chaotic time series with wavelet networks // Physica D. 1995. P.225-238. 9. *Rao S., Kumthekar B.* Recurrent wavelet networks // Neural networks for signal processing III: proceeding of the 1993 IEEE-SP Workshop IEEE. 1993. P.3143-3147. 10. *БодянскийЕ.В., Винокурова Е.А., Плисс И.П.* Алгоритм обучения вплеск-нейронной сети // Сб. науч. тр. 1-го Международного радиоэлектронного форума «Прикладная радиоэлектроника. Состояние и перспективы развития» - МРФ-2002. Часть 2. Харьков: АН ПРЭ, ХНУРЭ, 2002. С.87-89. 11. *Moody J., Darken C.J.* Fast learning in networks of locally-tuned processing units // Neural Computation.1989.1.P.281-294. 12. *Park J., Sandberg I.W.* Universal approximation using radial-basis-function networks // Neural Computation. 1991. 3. P.246-257. 13. *Radial* Basis Function Networks. Recent Developments in Theory and Applications / Eds. by R.J.Howlett, L.C.Jain. Berlin: Springer, 2001. 318p. 14. *Widrow B., Hoff Jr.M.E.* Adaptive switching circuits // 1960 IRE Western Electric Show and Connection Record. 1960. Part 4. P.96-104. 15. *Widrow B., Lee M.* 30 years of adaptive neural networks: perceptron, adaline and backpropagation // Proc. IEEE. 1990. 78. № 9. P.1415-1442. 16. *Ham F.M., KostanicI.* Principles of Neurocomputing for Science & Engineering. N.Y.: Mc Graw-Hill, Inc., 2001. 642p. 17. *Chen S., Billings S.A., Cowan C.F.N., Grant P.M.* Non-linear system identification using radial basis function // Int. J. Syst. Sci. 1990. 21. № 12. P.2513-2539. 18. *Chen S., Cowan C.F.N., Grant P.M.* Orthogonal least squares learning algorithm for radial basis function networks // IEEE. Trans. on Neural Networks. 1991. 2. № 12. P.302-308. 19. *Chen S., Billings S.A., Grant P.M.* Recursive hybrid algorithm for nonlinear system identification using radial basis function // Int. J. Contr. 1992. 55. № 5. P.1051-1070. 20. *Shah S., Palmieri F., Datum M.* Optimal filtering algorithms for fast learning in feedforward neural networks // Neural Networks. 1992. № 5. P.779-787. 21. *Kasparian V., Batur C., Zhang H., Padovan J.* Davidon least squares-based learning algorithm for feedforward neural networks // Neural Networks. 1994. 7. № 4. P.661-670. 22. *Sherstinsky A., Picard R.W.* On the efficiency of the orthogonal least squares training method for radial basis function networks // IEEE Trans. on Neural Networks. 1996. 7. №1. P.195-200. 23. *Fung C.F., Billings S.A., Luo W.* On-line supervised adaptive training using radial basis function networks // Neural Networks. 1996. 9. №.9. P.1597-1617. 24. *Goodwin G.C., Ramadge P.J., Caines P.E.* Discrete time stochastic adaptive control // SIAM J. Control and Optimization. 1981. 19. № 6. P.829-853. 25. *Goodwin G.C., Ramadge P.J., Caines P.E.* A globally convergent adaptive predictor // Automatica. 1981. 17. №1. P.135-140. 26. *Бодянский Е.В., Плисс И.П.* Об одном модифицированном алгоритме одновременного действия для идентификации объектов управления. Харьков, 1981. 18с. Рук. деп. в ВИНИТИ 15.09.1981, № 4474 – 81 Деп. 27. *Бодянский Е.В., Плисс И.П.* Об одном многошаговом адаптивном алгоритме идентификации нестационарных объектов. Харьков, 1984.8с. Рук. деп. в УкрНИИНТИ 03.02.1984, № 183 Ук–Д 84. 28. *Бодянский Е.В., Плисс И.П., Соловьева Т.В.* Многошаговые оптимальные упредители многомерных нестационарных стохастических процессов // Докл. АН УССР. 1986. Сер.А. № 12. С. 47-49.

Поступила в редколлегию 25.12.2002

Рецензент: д-р техн. наук, проф. Любчик Л.М..

Бодянский Евгений Владимирович, д-р техн. наук, профессор кафедры искусственного интеллекта, научный руководитель проблемной НИЛ АСУ ХНУРЭ, член IEEE, WSES. Научные интересы: нейро-фаззи-системы. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 40-98-90. E-mail: bodya@kture.kharkov.ua

Винокурова Елена Анатольевна, аспирант кафедры искусственного интеллекта, младший научный сотрудник ПНИЛ АСУ ХНУРЭ. Научные интересы: искусственные нейронные сети, всплески (вейвлеты, wavelet). Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел.: 40-98-90. E-mail: elena_vinokurova@yahoo.com.

УДК 519.81

ИСПОЛЬЗОВАНИЕ МЕТОДА ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КОМПАРАТОРНОЙ ИДЕНТИФИКАЦИИ МОДЕЛИ МНОГОФАКТОРНОГО ОЦЕНИВАНИЯ

ПЕТРОВ Э.Г., БУЛАВИН Д.А.

Рассматривается постановка и основанный на применении генетических алгоритмов метод решения задачи идентификации структуры модели индивидуального многофакторного оценивания.

1. Введение

Всякое новое научное направление возникает на базе уже известных и связано с другими направлениями. Несомненна связь проблемы принятия решений с исследованием операций, кибернетикой, искусственным интеллектом. В то же время принятие решений имеет свои, отличные от прочих направлений задачи и свою логику развития. Одной из актуальных проблем общей теории принятия решений является формализация процессов выбора решений в условиях многокритериальности. Конструктивное решение этой проблемы связано с идентификацией модели формирования скалярной многофакторной оценки качества (эффективности) допустимых альтернативных решений X вида:

$$P(x) = F[\lambda_i, k_i(x)], \qquad (1)$$

где $k_i(x)$, $i = \overline{1,n}$ – частные критерии (характеристики), однозначно определенные для каждого $x \in X$; λ_i – коэффициенты изоморфизма, приводящие разнородные частные критерии к единой размерности (или безразмерному виду), одинаковому интервалу изменения и учитывающие различную их значимость (вес) в обобщенной оценке $P(x)$.

В общем случае проблема идентификации модели (1) требует решения задач структурной и параметрической идентификации, т.е. соответственно определения вида оператора F и значений параметров λ_i. При этом классические методы идентификации непригодны для идентификации моделей интеллектуальной деятельности. Перспективным для этих целей является использование метода компараторной идентификации [1,3].

2. Метод компараторной идентификации модели многофакторного оценивания

Решение задачи структурной идентификации любой математической модели связано с необходимостью принятия некоторой гипотезы о характере взаимосвязи входных и выходных переменных. В