

УДК 62-50

Б.А. Скороход, профессор, д-р техн. наук*Севастопольский национальный технический университет,**ул. Университетская 33, г. Севастополь, Украина, 99053**E-mail: b.skorohod@mail.ru***ДИФFUЗНЫЙ ФИЛЬТР КАЛМАНА В ЗАДАЧЕ НАСТРОЙКИ ПАРАМЕТРОВ НЕЧЕТКИХ СИСТЕМ***Рассматривается задача оценки параметров нечеткой системы Сугено.****Ключевые слова:** фильтр Калмана, диффузный фильтр Калмана, система нечеткого логического вывода, оценка параметров нечетких систем*

Введение. Опубликованы различные методы оптимизации параметров функций принадлежности (ФП) и параметров, входящих в следствия правил для нечетких систем (НС). Так в [1, 2] разработан алгоритм обратного распространения ошибки (АОРО) и гибридная процедура обучения системы Сугено (ANFYS), интегрирующая АОРО и рекуррентный метод наименьших квадратов (РМНК). В [3] используется расширенный фильтр Калмана (РФК) для обучения системы Мамдани с треугольными ФП. Алгоритмы кластерного анализа и эволюционного программирования предложены в [2, 4]. Во многих случаях, все эти альтернативные методы имеют серьезные проблемы обусловленные отсутствием сходимости или сходимостью за неприемлемо большое время и к тому же только к локальному минимуму. При этом известно, что сходимость зависит как от используемого алгоритма, так и от того насколько удачно выбрано начальное приближение. Вместе с тем, структура НС Сугено такова, что априорная информация о части оцениваемых параметрах отсутствует и, следовательно, неясно, как обоснованно инициализировать выбранный алгоритм. В работе предложены алгоритмы оценивания параметров НС Сугено, основанные на использовании нелинейных модификаций диффузного фильтра Калмана [5], ориентированного на решение подобных задач. Показано, что они обладают рядом важных свойств, отсутствующих у его классического прототипа – РФК. Более точно, его итерационной версии с матрицей ковариации начальных условий, часть элементов которой, учитывающая отсутствие априорной информации о начальных условиях пропорциональна большому параметру.

1. Задание априорной информации для неизвестных параметров нечетких систем

Рассмотрим нелинейную систему

$$y = F(x) + \xi, \quad (1)$$

где $x \in R^p$, $y \in R^1$ – вход и выход ее, соответственно, $\xi \in R^1$ – случайная величина с $E(\xi) = 0$, $\text{var}(\xi) = \sigma_\xi^2$, $\sigma_\xi^2 > 0$ – задано, $F(x): R^p \rightarrow R^1$ – неизвестная функция, определенная на интервале $U = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_p, b_p] \subset R^p$, $E(.)$ – математическое ожидание, $\text{var}(.)$ – дисперсия, R^n – пространство векторов размерности n . Система может использоваться для описания как статических, так и динамических объектов. При учете динамики y является выходом в заданный момент времени t , а компоненты x его значениями в некоторые моменты предшествующие t и входными воздействиями.

Требуется по последовательности, измеренных значений входов и выходов системы $\{x_t, y_t\}$, $t \in T = \{a, a+1, \dots, N\}$ оценить параметры, входящие в представление $F(x)$ системой нечеткого логического вывода Сугено [1]

$$F(x) = F(x, \beta, \alpha) = \frac{\sum_{l=1}^k f_l(x, \alpha^l) (\prod_{j=1}^p \mu_l^j(x^j, \beta^{lj}))}{\sum_{l=1}^k (\prod_{j=1}^p \mu_l^j(x^j, \beta^{lj}))} = \frac{\sum_{l=1}^k f_l(x, \alpha^l) \mu_l(x, \beta^l)}{\sum_{l=1}^k \mu_l(x, \beta^l)}, \quad (2)$$

где $x = (x^1, \dots, x^p)^T$, $\mu_l^j(x^j, \beta^{lj})$ – ФП, параметризованные вектором $\beta^{lj} \in R^{m_{lj}}$, $f_l(x, \alpha^l) = \alpha_0^l + \alpha_1^l x_1 + \dots + \alpha_p^l x_p$, $\alpha^l = (\alpha_0^l, \dots, \alpha_p^l)^T$, k – количество правил, $\beta = ((\beta^{11})^T, \dots, (\beta^{1p})^T, \dots, (\beta^{k1})^T, \dots, (\beta^{kp})^T)^T \in R^m$, $\alpha = ((\alpha^1)^T, \dots, (\alpha^k)^T)^T \in R^q$, T – операция транспонирования.

Начнем с анализа выражения (2). Два его свойства имеют для нас принципиальное значение с точки зрения конструирования алгоритмов оценивания неизвестных параметров ФП β и параметров следствий α . Во-первых, это выражение линейно зависит от α и нелинейно от β , т.е. имеет структуру

вида $F(x, \beta, \alpha) = \varphi(x, \beta)\alpha$. Во-вторых, для параметров ФП часто имеются некоторые экспертные знания, связанные со спецификой решаемой содержательной задачи. Например, для гауссовских ФП

$$\mu_l^j(x^j) = \exp(-(x^j - m_{lj})^2 / 2\sigma_{lj}^2), l = 1, 2, \dots, k, j = 1, 2, \dots, p \quad (3)$$

могут быть определены до оптимизации значения $\bar{m}_{lj}$ и $\bar{\sigma}_{lj}$. Естественным требованием после оптимизации является сохранение первоначального порядка лингвистических значений $m_{lj+1} > m_{lj}$ и $\sigma_{lj} > 0$, накладывающее ограничения на величину разброса относительно выбранных $\bar{m}_{lj}$ и $\bar{\sigma}_{lj}$. С учетом этого, параметры ФП до оптимизации могут интерпретироваться, как случайные величины с математическими ожиданиями $\bar{m}_{lj}$, $\bar{\sigma}_{lj}$ и достаточно малыми дисперсиями $S_\beta = E((\beta - m_\beta)(\beta - m_\beta)^T) = \text{diag}(s_1, \dots, s_k)$, где $s_i \in [0, \varepsilon]$ – равномерно распределенные случайные числа, $\varepsilon > 0$ – малый параметр. Аналогичные соображения справедливы для любых параметризованных ФП. В отличие от β , относительно α отсутствует какая-либо априорная информация.

Учитывая сказанное, введем модель пространства состояния

$$\beta_{t+1} = \beta_t, \alpha_{t+1} = \alpha_t, y_t = F(x_t, \beta_t, \alpha_t) + \xi_t, t \in T, \quad (4)$$

где β_0 случайный вектор с известными характеристиками

$$m_\beta = E\beta_0, S_\beta = E((\beta_0 - m_\beta)(\beta_0 - m_\beta)^T), \quad (5)$$

а статистические характеристики α_0 неизвестны.

2. Диффузные алгоритмы оценки параметров. Рассмотрим линейную регрессию

$$y_t = c_t x + \xi_t, \quad t \in T = \{a, a+1, \dots\}, \quad (6)$$

где $x = (\beta^T, \alpha^T)^T$, $c_t^T \in R^n$, β и α – некоррелированные векторные случайные величины, удовлетворяющие условиям

$$E(\alpha) = 0, \text{var}(\alpha) = \lambda I_q, \quad (7)$$

где $\lambda > 0$ – большой параметр, $I_q \in R^{q \times q}$ – единичная матрица, $R^{k \times l}$ – пространство матриц размерности $k \times l$, соответственно.

Введем модель пространства состояния, соответствующую (6)

$$x_{t+1} = x_t, y_t = c_t x_t + \xi_t, \quad t \in T = \{a, a+1, \dots\}. \quad (8)$$

При сделанных предположениях, линейная несмещенная оценка состояния системы (8) с минимальной среднеквадратической ошибкой определяется фильтром Калмана

$$\hat{x}_{t+1} = \hat{x}_t + K_t(y_t - c_t \hat{x}_t), \hat{x}_a = (m_a^T, 0_q^T)^T, t \in T = \{a, a+1, \dots\}, \quad (9)$$

где

$$K_t = P_t c_t^T N_t^{-1}, \quad (10)$$

$$P_{t+1} = P_t - P_t c_t^T N_t^{-1} c_t P_t, P_a = \text{blockdiag}(S_a, \lambda M_a), \quad (11)$$

$m_a = \{m_a^i\} \in R^m, S_a = \{s_a^{ij}\} \in R^{m \times m}$, $\text{blockdiag}(A, B)$ – блочно-диагональная матрица, 0_n – вектор размерности n с нулевыми элементами, $N_t = c_t P_t c_t^T + \sigma_t^2$.

Сформулируем утверждение, являющееся следствием результатов, полученных в [5].

Теорема 1. Пусть $\bar{T} = \{a, a+1, \dots, b\}$ произвольное, ограниченное множество. Тогда:

1. Справедливы равномерные по $t \in \bar{T}$ асимптотические представления

$$P_t = R_t(I_{n-q} - W_{a,t} W_{a,t}^+) R_t^T \lambda + S_t + R_t W_{a,t}^+ R_t^T + O(\lambda^{-1}), \quad \lambda \rightarrow \infty, \quad (12)$$

$$K_t = \begin{cases} 0, & c_t = 0 \\ K_t^d + O(\lambda^{-1}), & \lambda \rightarrow \infty, c_t \neq 0 \end{cases} \quad (13)$$

где

$$W_{a,t+1} = W_{a,t} + R_t^T C_t^T N_{1t}^{-1} C_t R_t, W_{a,a} = 0_{k \times k}, \quad (14)$$

$$S_{t+1} = S_t - S_t C_t^T N_{1t}^{-1} C_t S_t, \tilde{S}_a = \text{block diag}(S_a, 0_{k \times k}), \quad (15)$$

$$R_{t+1} = A_{1t} R_t, A_{1t} = I - S_t C_t^T N_{1t}^{-1} C_t, R_a = (e_{r+1}, \dots, e_{k+r}), \quad (16)$$

$$K_t^d = (S_t + A_{1t} R_t W_{a,t}^+ R_t^T) C_t^T N_{1t}^{-1}, N_{1t} = C_t S_t C_t^T + \sigma_t^2, \quad (17)$$

A^+ – псевдообратная матрица Пенроуза A , $e^i \in R^n$ – i -единичный вектор.

2. Для любого $\varepsilon > 0$ равномерно по $t \in \bar{T}$ выполняется условие

$$P(\|\hat{x}_t - \hat{x}_t^d\| \geq \varepsilon) = O(\lambda^{-1}), \quad \lambda \rightarrow \infty, t \in \bar{T}, \quad (18)$$

где $\hat{x}_t^d$ определяется с помощью ДФК

$$\hat{x}_{t+1}^d = \hat{x}_t^d + K_t^d (y_t - c_t \hat{x}_t^d), \hat{x}_a^d = (m_a^T, 0_k^T)^T. \quad (19)$$

Соотношения (12) – (19) определяют диффузный ФК (ДФК) для оценки параметров линейной регрессии (6).

Теорема 1 неприменима для оценки состояния (4) из-за нелинейности модели измерений. Предлагаемый далее подход состоит в линеаризации нелинейности в окрестности последней оценки и использовании ДФК из теоремы 1 в итерационной форме. Специфика рассматриваемой задачи приводит к алгоритму, обладающему рядом важных свойств, отсутствующих у его классического прототипа – РФК. Более точно, его итерационной версии с матрицей ковариации начальных условий пропорциональной большому параметру, учитывающей отсутствие априорной информации о начальных условиях. Будет удобно сформулировать, полученный результат в виде утверждения и нескольких следствий к нему.

Теорема 2. Расширенный ДФК (РДФК) для оценки состояния (4) определяется выражениями:

$$\bar{\beta}_{t+1} = \bar{\beta}_t + K_t^\beta (y_t - F(x_t, \bar{\alpha}_t, \bar{\beta}_t)), \quad \bar{\beta}_a = \hat{\beta}_i, \quad (20)$$

$$\bar{\alpha}_{t+1} = \bar{\alpha}_t + K_t^\alpha (y_t - F(x_t, \bar{\alpha}_t, \bar{\beta}_t)), \quad \bar{\alpha}_a = \hat{\alpha}_i, \quad (21)$$

$$K_t = ((K_t^\beta)^T, (K_t^\alpha)^T)^T = (\text{blockdiag}(\bar{P}_t, 0_{q \times q}) + A_{1t} \bar{R}_t \bar{M}_{a,t+1}^+ \bar{R}_t^T) C_t^T N_{1t}^{-1}, \quad (22)$$

$$\bar{P}_{t+1} = \bar{P}_t - \bar{P}_t (C_t^\beta)^T N_{1t}^{-1} C_t^\beta \bar{P}_t, \quad \bar{P}_a = P_i, \quad (23)$$

$$\bar{M}_{a,t+1} = \bar{M}_{a,t} + \bar{R}_t^T C_t^T N_{1t}^{-1} C_t \bar{R}_t, \quad \bar{M}_{a,a} = M_{a,i}, \quad (24)$$

$$\bar{R}_t = (\bar{R}_{1t}^T, I_q)^T, \quad \bar{R}_{1t+1} = (I_m - \bar{P}_t (C_t^\beta)^T N_{1t}^{-1} C_t^\beta) \bar{R}_{1t} - \bar{P}_t (C_t^\beta)^T N_{1t}^{-1} C_t^\alpha, \quad \bar{R}_{1a} = R_{1i}, \quad (25)$$

$$N_{1t} = C_t^\beta \bar{P}_t (C_t^\beta)^T + \sigma_t^2, \quad A_{1t} = I_{m+q} - \begin{pmatrix} \bar{P}_t (C_t^\beta)^T N_{1t}^{-1} C_t^\beta & \bar{P}_t (C_t^\beta)^T N_{1t}^{-1} C_t^\alpha \\ 0_{q \times m} & 0_{q \times q} \end{pmatrix}, \quad (26)$$

$$C_t = (C_{1t}^\beta(\bar{\alpha}_t, \bar{\beta}_t), \dots, C_{qt}^\beta(\bar{\alpha}_t, \bar{\beta}_t), C_{1t}^\alpha(\bar{\beta}_t), \dots, C_{kt}^\alpha(\bar{\beta}_t)) = (C_t^\beta, C_t^\alpha), \quad (27)$$

$$C_{kt}^\beta(\bar{\alpha}_t, \bar{\beta}_t) = \partial F(x_t, \bar{\alpha}_t, \bar{\beta}_t) \partial \bar{\beta}_{kt}, k = 1, 2, \dots, m, \quad C_{jt}^\alpha(\bar{\alpha}_t, \bar{\beta}_t) = \partial F(x_t, \bar{\alpha}_t, \bar{\beta}_t) \partial \bar{\alpha}_{jt}, j = 1, 2, \dots, q, \quad (28)$$

$$\hat{\beta}_{i+1} = \bar{\beta}_N, \quad \hat{\alpha}_{i+1} = \bar{\alpha}_N, \quad \bar{P}_{i+1} = \bar{P}_N, \quad \bar{M}_{a,i+1} = \bar{M}_{a,N} \quad i = 1, 2, \dots, M \quad (29)$$

с инициализацией при $i = 1$

$$\bar{\beta}_a = \hat{\beta}_1 = m_\beta, \quad \bar{\alpha}_a = \hat{\alpha}_1 = 0, \quad \bar{P}_a = S_\beta, \quad \bar{M}_{a,a} = 0, \quad R_a = (e^{m+1}, \dots, e^{m+q}).$$

Доказательство. Следует из анализа соотношений для линеаризованного в окрестности текущей оценки ДФК. Действительно, теорема 1 дает

$$\bar{\beta}_{t+1} = \bar{\beta}_t + K_t^\beta (y_t - F(x_t, \bar{\alpha}_t, \bar{\beta}_t)), \quad \bar{\beta}_a = m_\beta, \quad (30)$$

$$\bar{\alpha}_{t+1} = \bar{\alpha}_t + K_t^\alpha (y_t - F(x_t, \bar{\alpha}_t, \bar{\beta}_t)), \quad \hat{\alpha}_a = 0, \quad (31)$$

$$K_t = ((K_t^\beta)^T, (K_t^\alpha)^T)^T = (S_t + A_{1t} R_t M_{a,t+1}^+ R_t^T) C_t^T N_{1t}^{-1}, \quad (32)$$

$$S_{t+1} = S_t - S_t C_t^T N_{1t}^{-1} C_t S_t, \tilde{S}_a = \text{blockdiag}(S_\beta, 0_{q \times q}), \quad (33)$$

$$M_{a,t+1} = M_{a,t} + R_t^T C_t^T N_{1t}^{-1} C_t R_t, M_{a,a} = 0, \quad (34)$$

$$R_{t+1} = A_{1t} R_t, R_a = (e^{m+1}, \dots, e^{m+q}), \quad (35)$$

$$A_{1t} = I_{m+q} - S_t C_t^T N_{1t}^{-1} C_t, \quad N_{1t} = C_t S_t C_t^T + \sigma_t^2, \quad t \in T. \quad (36)$$

С учетом того, что решение матричного уравнения (33) имеет вид $S_t = \text{blockdiag}(P_t, 0_{q \times q})$, получаем

$$N_{1t} = C_t^\beta P_t (C_t^\beta)^T + \sigma_t^2, \quad A_{1t} = I_{m+q} - \begin{pmatrix} P_t (C_t^\beta)^T N_{1t}^{-1} C_t^\beta & P_t (C_t^\beta)^T N_{1t}^{-1} C_t^\alpha \\ 0_{q \times m} & 0_{q \times q} \end{pmatrix},$$

$$R_t = (R_{1t}^T, I_q)^T, \quad R_{1t+1} = (I_m - P_t (C_t^\beta)^T N_{1t}^{-1} C_t^\beta) R_{1t} - P_t (C_t^\beta)^T N_{1t}^{-1} C_t^\alpha, \quad R_{1a} = 0_{m \times q},$$

что влечет соотношения (20) – (29).

Замечание 1. В [6] показано, что при больших λ даже в линейном случае классической регрессии возможна расходимость рекуррентных алгоритмов. Кроме того, в этой же работе показано, что свойства сходимости улучшаются при диффузной инициализации линейной регрессии – использованием предельного алгоритма при $\lambda \rightarrow \infty$. Для модели (4) имеет место аналогичная ситуация, усложненная необходимостью оценивать дополнительно нелинейно входящий векторный параметр β с заданными статистическими характеристиками (5).

Следствие 1. Пренебрегая в матрице A_{1t} недиагональным блоком, приходим к упрощенной версии РДФК

$$\bar{\beta}_{t+1} = \bar{\beta}_t + K_t^\beta (y_t - F(x_t, \bar{\alpha}_t, \bar{\beta}_t)), \quad \bar{\beta}_a = \hat{\beta}_i, \quad (37)$$

$$\bar{\alpha}_{t+1} = \bar{\alpha}_t + K_t^\alpha (y_t - F(x_t, \bar{\alpha}_t, \bar{\beta}_t)), \quad \bar{\alpha}_a = \hat{\alpha}_i, \quad (38)$$

$$K_t^\beta = P_t (C_t^\beta)^T (N_t)_{1t}^{-1}, \quad K_t^\alpha = M_t^+ (C_t^\alpha)^T (N_t)_{1t}^{-1}, \quad (39)$$

$$\bar{P}_{t+1} = \bar{P}_t - \bar{P}_t (C_t^\beta)^T (N_t)_{1t}^{-1} C_t^\beta \bar{P}_t, \quad \bar{P}_a = P_i \quad (40)$$

$$\bar{M}_{a,t+1} = \bar{M}_{a,t} + (C_t^\alpha)^T N_{1t}^{-1} C_t^\alpha, \quad \bar{M}_{a,a} = M_{a,i}, \quad (41)$$

$$\hat{\beta}_{i+1} = \bar{\beta}_N, \quad \hat{\alpha}_{i+1} = \bar{\alpha}_N, \quad \bar{P}_{i+1} = \bar{P}_N, \quad \bar{M}_{a,i+1} = \bar{M}_{a,N} \quad i = 1, 2, \dots, M \quad (42)$$

с инициализацией при $i = 1$ $\bar{\beta}_a = \hat{\beta}_1 = m_\beta$, $\bar{\alpha}_a = \hat{\alpha}_1 = 0$, $\bar{P}_a = S_\beta$, $\bar{M}_{a,a} = 0$. Численная реализация этого алгоритма требует существенно меньшего количества операций по сравнению с РФК. Действительно, вместо решения матричного уравнения для $S_t \in R^{(p+q) \times (p+q)}$ с начальным условием $\tilde{S}_a = \text{blockdiag}(S_\beta, \lambda I_q)$, где $\lambda > 0$ – большой параметр, задача сводится к решению двух уравнений меньшего порядка, для $P_t \in R^{m \times m}$ и $M_t \in R^{q \times q}$. Помимо этого отпадает необходимость подбора величины λ , а сам фильтр будет робастным по отношению к ошибкам округления [6].

Следствие 2. Алгоритм, описываемый выражениями (37)–(42) может быть получен, также, с использованием идеи двух стадийной оптимизации. Действительно, при фиксированном значении β_t , алгоритм представляют собой РМНК с диффузной инициализацией [6] для оценки α_t , а при фиксированном α_t – расширенный фильтр Калмана для оценки β_t .

Следствие 3. Из теоремы 2 вытекают выражения для алгоритмов типа ANFIS [1, 2], в соответствии, с которыми оценка параметров ищется из условия минимума критерия качества $J(\alpha, \beta) = 1/(2N) \sum_{t=1}^N (y_t - F(x_t, \beta, \alpha))^2$ по α при фиксированном β с помощью РМНК и по β при фиксированном α , используя градиентный алгоритм. Действительно, полагая в (66) $P_t = \gamma_i I_m$, где $\gamma_i > 0$ – малый параметр, получаем гибридный алгоритм (градиентный алгоритм в приращениях [7] для оценки β и РМНК с диффузной инициализацией [6] для оценки α)

$$\bar{\beta}_{t+1} = \bar{\beta}_t + \gamma_i (C_t^\beta)^T (y_t - F(x_t, \bar{\alpha}_t, \bar{\beta}_t)), \quad \bar{\beta}_a = \hat{\beta}_i, \quad (43)$$

$$\bar{\alpha}_{t+1} = \bar{\alpha}_t + K_t^\alpha (y_t - F(x_t, \bar{\alpha}_t, \bar{\beta}_t)), \quad \bar{\alpha}_a = \hat{\alpha}_i, \quad (44)$$

$$K_t^\alpha = M_t^+ (C_t^\alpha)^T / \sigma_t^2, \quad (45)$$

$$\bar{M}_{a,t+1} = \bar{M}_{a,t} + (C_t^\alpha)^T C_t^\alpha / \sigma_t^2, \quad \bar{M}_{a,a} = M_{a,i}, \quad (46)$$

$$\hat{\beta}_{i+1} = \bar{\beta}_N, \quad \hat{\alpha}_{i+1} = \bar{\alpha}_N, \quad \bar{M}_{a,i+1} = \bar{M}_{a,N} \quad i = 1, 2, \dots, M, \quad (47)$$

с инициализацией при $i = 1$ $\bar{\beta}_a = \hat{\beta}_1 = m_\beta$, $\bar{\alpha}_a = \hat{\alpha}_1 = 0$, $\bar{M}_{a,a} = 0$.

Полагая $\bar{\beta}_t = \hat{\beta}_i$, приходим к упрощенной его версии

$$\hat{\beta}_{i+1} = \hat{\beta}_i - \gamma_i \partial J(\hat{\alpha}_{i+1}, \hat{\beta}_i) / \partial \hat{\beta}_i, \quad \hat{\beta}_1 = m_\beta, \quad i = 1, 2, \dots, M,$$

$$\bar{\alpha}_a = \hat{\alpha}_i, \quad \bar{\alpha}_{t+1} = \bar{\alpha}_t + K_t^\alpha (y_t - F(x_t, \bar{\alpha}_t, \hat{\beta}_i)), \quad (48)$$

$$K_t^\alpha = M_t^+ (C_t^\alpha)^T / \sigma_t^2, \quad M_{a,t+1} = M_{a,t} + (C_t^\alpha)^T C_t^\alpha / \sigma_t^2, \quad M_{a,a} = 0, \quad t \in T, \quad (49)$$

где $\bar{\alpha}_a = \hat{\alpha}_1 = 0$ для $i = 1$.

Замечание 2. Матрицу C_t нетрудно определить в предположении, что ФП гауссовские, по выражениям:

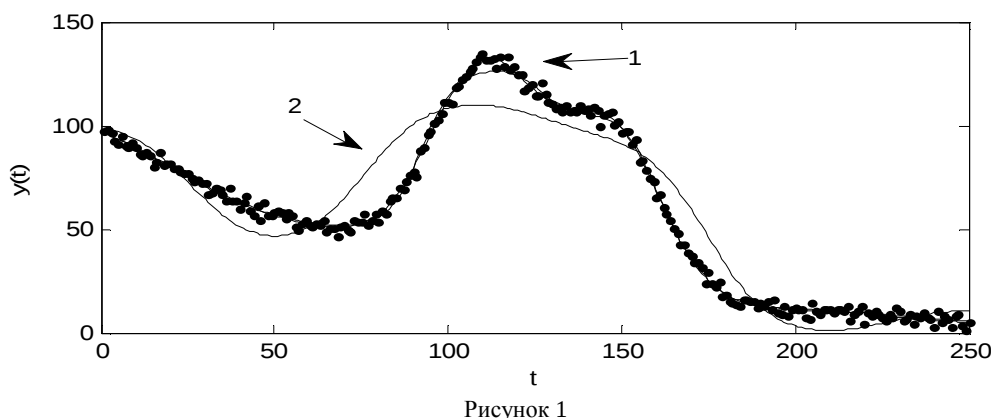
$$\begin{aligned} \partial F(x, \alpha, \beta) \partial m_{lj} &= f_l(x) \partial \mu_l(x) \partial m_{lj} [1/B(x) - \mu_l(x)/(B(x))^2] = \\ &= (\mu_l(x)/B(x)) (f_l(x) - F(x, \alpha, \beta)(x^j - m_{lj})/\sigma_{lj}^2), \end{aligned} \quad (50)$$

$$\partial F(x, \alpha, \beta) \partial \sigma_{lj} = (\mu_l(x)/B(x)) (f_l(x) - F(x, \alpha, \beta)(x^j - m_{lj})^2/\sigma_{lj}^3), \quad (51)$$

$$\partial F(x, \alpha, \beta) \partial \alpha_{il} = x_i \mu_l(x)/B(x), \quad i = 0, 1, \dots, p, \quad x_0 = 0, \quad (52)$$

$$B(x) = \sum_{l=1}^k \mu_l(x), \quad l = 1, 2, \dots, k, \quad j = 1, 2, \dots, p. \quad (53)$$

3. Численный пример. Проиллюстрируем эффективность, полученных в работе результатов, используя численный пример. Рассмотрим задачу идентификации объекта по вход – выходным данным [8], показанных на рисунке 1 точками, используя систему нечеткого логического вывода Сугено. Имеются два пика, наблюдаемые на фоне затухающего тренда и широкополосного шума. Использовались 6 ФП и система Сугено нулевого порядка. Значения математических ожиданий центров m_{lj} задавались, используя стандартную методику начального (до оптимизации) расположения параметров ФП [9], показанную на рисунке 2. Начальные разбросы σ_{lj} полагались равными нулю. Величина q оценивалась по разбросу данных и полагалась равной 1,69. В матричное уравнение (40) было добавлено аддитивное слагаемое $V_{2t} = \varepsilon I_n$, где $\varepsilon = 3 \times 10^{-5}$ – стабилизирующий параметр, выбранный в процессе моделирования. Результаты идентификации приведены на рисунках 1, 3, 4. На рисунках 1, кривые 1, 2 построены с помощью соотношений (43)–(47) и системы ANFIS [9], соответственно, после 40 итераций (эпох). Всего неизвестных параметров $n = 18$, из них $q = 12$ входят в описание ФП. Видно, что в отличие от предложенного в работе алгоритма, ANFIS не позволяет за 40 эпох получить приемлемый результат. Этот вывод можно уточнить, опираясь на графики, показанные на рисунок 3 – зависимости среднеквадратической ошибки оценивания от количества эпох. Действительно, для получения примерно одинаковой точности, предложенному алгоритму требуется 40 эпох, ANFIS – более 180. На рисунке 4 показаны ФП после оптимизации. Видно, что, несмотря на значительные изменения, их первоначальный лингвистический порядок (рисунок 2) сохранился, а все оптимизированные значения σ_j , $j = 1, \dots, 6$ положительны. И наконец, отметим, что использование ДРФК из теоремы 2 не приводит к уменьшению среднеквадратической ошибки оценивания, увеличивая продолжительность вычислений примерно на 44 %.



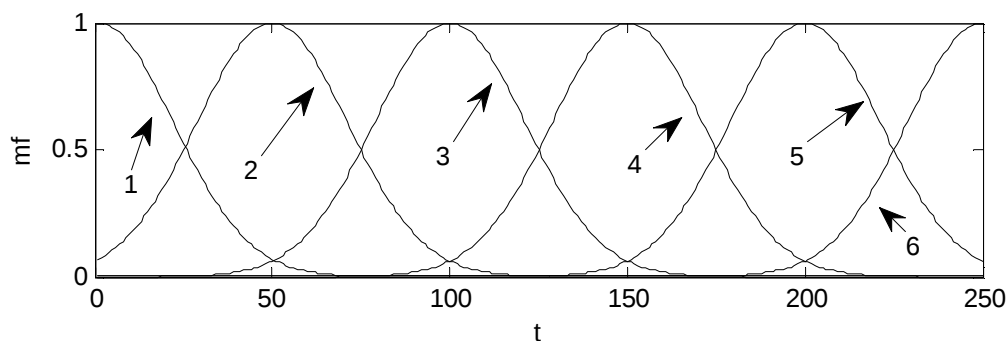


Рисунок 2

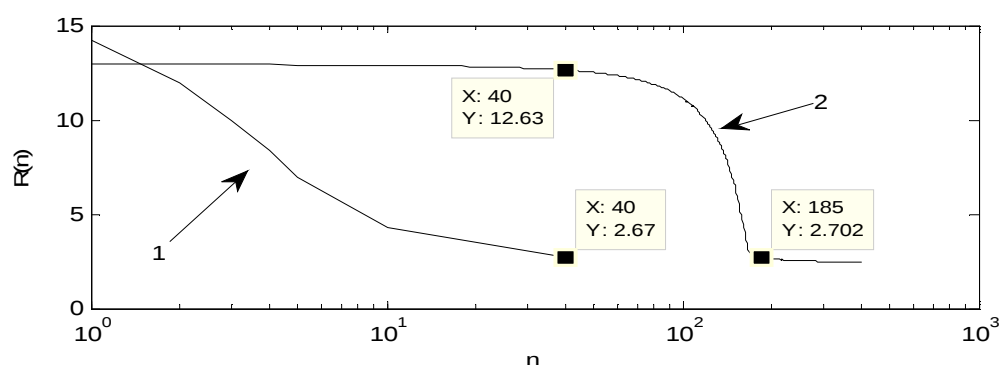


Рисунок 3

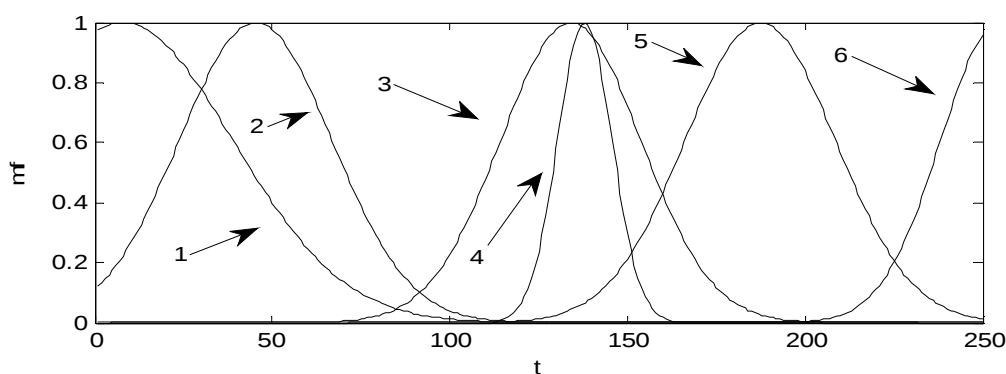


Рисунок 4

Заключення. Предложены нелинейные алгоритмы настройки, нечеткой системы Сугено, обладающие рядом важных свойств, отсутствующих у его классического прототипа – РФК. Более точно, его итерационной версии с матрицей ковариации начальных условий, часть элементов которой, учитывающая отсутствие априорной информации о начальных условиях, пропорциональна большому параметру. Дальнейшие исследования будут направлены на изучение асимптотических свойств, предложенных алгоритмов оценки.

Бібліографічний список використаної літератури

1. Jang R. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence / R. Jang, C. Sun, E. Mizutani. — N.Y.: Prentice Hall, 1997. — 601 p.
2. Jang R. Neuro-fuzzy modeling and control / R. Jang, C. Sun // Proceeding of the IEEE. — Vol. 83. — No.3, March 1995. — P. 378–406.
3. Simon D. Training fuzzy systems with the extended Kalman filter / D. Simon // Fuzzy Sets and Systems. — December 2002. — Vol. 132. — Issue 2. — P. 189–199.
4. Goddard J. Rule learning in fuzzy systems using evolutionary programs. / J. Goddard, R. Parrazales, I. Lopez, A de Luca // IEEE Midwest Symp. on Circuits and Systems, Ames, Iowa. — 1996. — P. 703–709.
5. Скороход Б.А. Диффузная инициализация фильтра Калмана // Пробл. упр. и информатики. — 2011. — № 2. — С. 78–91.
6. Скороход Б. А. Асимптотика линейной рекуррентной регрессии при диффузной инициализации / Б.А. Скороход // Пробл. упр. и информатики. — 2009. — № 3. — С. 98–107.

7. Bertsekas D. A new class of incremental gradient methods for least squares problems / D. Bertsekas // SIAM J. on Optimization. — 1997. — № 7. — P. 913–926.
8. Curve Fitting Toolbox 3. MathWorks, Inc.
9. Fuzzy Logic Toolbox. The MathWorks, Inc.

Поступила в редакцію 31.08.2011 г.

Скороход Б.А. Дифузний фільтр Калмана в задачі настроювання параметрів нечітких систем

Розглядається задача оцінки параметрів нечіткої системи Сугено.

Ключові слова: фільтр Калмана, дифузний фільтр Калмана, система нечіткого логічного виводу.

Skorohod B.A. The diffuse Kalman filter in the problem of setting the fuzzy system parameters

The problem of estimating the parameters of the Sugeno fuzzy system is considered.

Keywords: Kalman filter, the diffuse Kalman filter, fuzzy system, parameters estimation.