DOI https://doi.org/10.15407/usim.2019.01.022
УДК 004.852

О.Г. РУДЕНКО, д-р техн. наук, проф., зав. кафедрой информационных систем,
Харьковский национальный экономический ун-т имени Семена Кузнеца,
Харьков, 61166, пр. Науки, 9-А, Украина,
oleg.rudenko@hneu.net

А.А. БЕССОНОВ, д-р техн. наук, проф. кафедры информационных систем,
Харьковский национальный экономический ун-т имени Семена Кузнеца,
Харьков, 61166, пр. Науки, 9-А, Украина,
oleksandr.bezsonov@hneu.net

РЕГУЛЯРИЗОВАННЫЙ АЛГОРИТМ ОБУЧЕНИЯ АДАЛИНЫ В ЗАДАЧЕ ОЦЕНИВАНИЯ НЕСТАЦИОНАРНЫХ ПАРАМЕТРОВ

В статье исследуются свойства модифицированного регуляризированного алгоритма Качмажа. Определены условия его сходимости при оценивании нестационарных параметров при наличии помех измерения. Получены неасимптотические и асимптотические оценки для уточнения входящих в алгоритмы параметров.

Ключевые слова: *регуляризованный алгоритм, сходимость, асимптотические оценки, рекуррентная процедура, закон дрейфа.*

Введение

Многие задачи управления, прогнозирования, распознавания образов и т.д. связаны с построением модели вида:

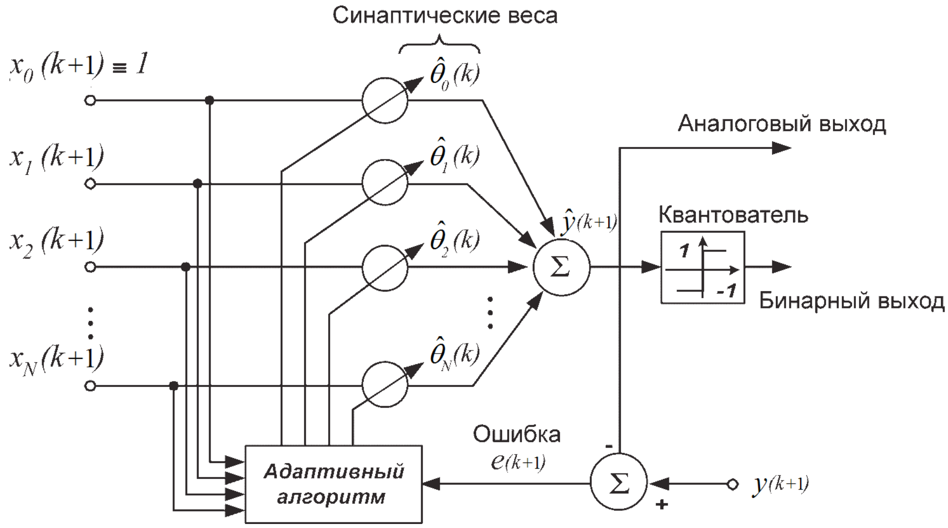
$$y(k+1) = \theta^{*T} x(k+1) + \xi(k+1), \qquad (1)$$

где $y(k+1)$ — наблюдаемый выходной сигнал; $x(k+1) = (x_1(k+1), x_2(k+1), \ldots, x_N(k+1))^T$ — вектор входных сигналов $N \times 1$; $\theta^* = \left(\theta_1^*, \theta_2^*, \ldots, \theta_N^*\right)^T$ — вектор искомых параметров $N \times 1$; $\xi(k+1)$ — помеха; k — дискретное время, и сводятся к минимизации некоторого наперед выбранного функционала качества (критерия идентификации). Наиболее широко используемый на практике квадратичный функционал приводит к различным алгоритмам идентификации, позволяющим получить оценки искомого вектора θ^* при нормальных распределениях помехи, т.е. $\xi(k+1) \sim N\left(0, \sigma_\xi^2\right)$.

Однако задача идентификации существенно усложняется, если параметры θ изменяются (дрейфуют) во времени, т.е. $\theta^*(k) = \mathrm{var}$.

Результаты большинства работ, посвященные этой проблеме, носят в основном рекомендательный характер и слабо связаны с характером нестационарности исследуемого объекта. Для оценки нестационарных параметров обычно применяются модифицированные алгоритмы МНК (использование скользящего окна либо экспоненциального сглаживания и т.д.), алгоритм Качмажа (и его динамический вариант), алгоритмы динамической адаптации и др.

Отсутствие достаточно общих рекомендаций по выбору алгоритмов оценивания нестационарных параметров обусловлено, по-видимому, трудностями теоретического исследования свойств этих алгоритмов в нестационарных условиях.



ADALINE

Эффективность применения того или иного алгоритма для оценки дрейфующих параметров существенно зависит от объема априорной информации о характере дрейфа. В соответствии с этим адаптивные алгоритмы идентификации нестационарных параметров можно разделить на два класса:

- алгоритмы оценивания параметров при известном законе их дрейфа;
- алгоритмы оценивания параметров при неизвестном законе дрейфа.

Первое направление берет начало с работы Дупача [1]. В этой работе предложено обобщение метода стохастической аппроксимации на случай, когда дрейф корня уравнения регрессии близок к линейному. В большинстве последующих работ рассматривался случай, когда дрейф либо в среднем близок к линейному, либо затухает. В [2] рассматривался случай, когда дрейф аппроксимировался, и задача идентификации сводилась к оцениванию неизвестного параметра $\theta^*(k)$ многомерным методом стохастической аппроксимации.

Среди наиболее простых в вычислительном отношении одношаговых алгоритмов идентификации наиболее эффективны алгоритмы Качмажа и Нагумо-Ноды [3, 4]. Однако их быстродействие оказывается зачастую недостаточным при оценивании нестационарных параметров. Знание (или аппроксимация) закона дрейфа позволяет получать эффективные алгоритмы отслеживания нестационарных параметров. В [5] рассмотрен алгоритм, названный динамическим алгоритмом Качмажа, в котором используется некоторая априорная модель модель дрейфа. Следует отметить, что ошибки в описании закона изменения параметров могут привести к потере свойств сходимости алгоритма.

Отсутствие информации о характере дрейфа требует разработки алгоритмов идентификации, использующих минимальный объем информации о $\theta^*(k)$ и сохраняющих работоспособность в широком диапазоне варьирования $\theta^*(k)$. Поэтому более естественным является изучение динамических свойств конкретных алгоритмов и определение максимально достижимой точности слежения, что в конечном итоге позволит определить области наиболее эффективного использования алгоритмов и разработать рекомендации по их практическому применению.

Естественной платой за отсутствие априорной информации о характере дрейфа параметров является запаздывание в оценке изменения параметров, а, следовательно, меньшая точность слежения, что обусловлено инерционностью алгоритмов.

Отметим, что в ряде случаев достаточно эффективным при решении данной задачи оказывается нейросетевой подход, основанный на использовании наиболее простой искусственной нейронной сети (ИНС) типа АДАЛИНА (*ADALINE*).

ADALINE имела структуру, представленную на рисунке [6, 7]. Данная ИНС, являющаяся первой линейной нейронной сетью, была предложена Б. Уидроу и М.Е. Хоффом в работе [6] в 1960 г. Разработанная ими модель была названа адаптивным нейроном, а затем *ADALINE (ADAptive LInear NEuron)*. С течением времени свою популярность в качестве ИНС она утратила и теперь *ADALINE* — это *ADAptive LINear Element*. На рисунке $\hat{y}(k+1)$ — выходной сигнал нейросетевой модели; $\hat{\theta}(k)=(\hat{\theta}_1(k),\hat{\theta}_2(k),\ldots,\hat{\theta}_N(k))^T$ — вектор весовых параметров $N\times 1$; $\hat{y}(k+1)=\hat{\theta}^T(k)x(k+1)$.

Основное отличие АДАЛИНЫ от персептрона заключалось в используемом алгоритме обучения. Хотя обучение АДАЛИНЫ также состояло в коррекции ее весовых коэффициентов на основании предъявленных обучающих пар и сравнения реальных выходных сигналов модели $\hat{y}(k+1)$ с желаемыми (требуемыми) $y(k=1)$, Уидроу и Хофф, основываясь на работах Н. Винера, связанных с исследованием фильтров, использовали для этого градиентный алгоритм

$$\hat{\theta}(k+1)=\hat{\theta}(k)+\frac{e(k+1)x(k+1)}{\left\| x(k+1)\right\|^2}, \qquad (2)$$

где $e(k+1)=y(k+1)-\hat{y}(k+1)=y(k+1)-\hat{\theta}^T(k)x(k+1)$, γ — некоторый параметр, влияющий на скорость сходимости алгоритма (длительность процесса обучения модели); $\left\|\bullet\right\|$ — евклидова норма. Хотя данный алгоритм в литературе по ИНС называется алгоритмом Уидроу–Хоффа, значительно раньше он был предложен в работе [3] для решения систем линейных алгебраических уравнений, а впоследствии, начиная с [8, 9], с успехом стал применяться для решения задачи идентификации при построении модели типа (1). Оценки скорости сходимости данного алгоритма при идентификации стационарных объектов (1) впервые были получены в [8-12].

Данный алгоритм применяется не только в системах идентификации, но и получил широкое распространение при решении задач фильтрации [13-15]. В иностранной литературе этот алгоритм более известен как нормализованный алгоритм наименьших квадратов (*NLMS — normilized least-mean-square*).

Постановка задачи

Для повышения вычислительной устойчивости (2) В.М. Чадеевым [8, 9] предложена его модификация — регуляризованный алгоритм

$$\hat{\theta}(k+1)=\hat{\theta}(k)+\frac{e(k+1)x(k+1)}{\left\| x(k+1)\right\|^2+\delta}, \qquad (3)$$

где $\delta>0$ — параметр регуляризации.

В названных выше работах были получены оценки скорости сходимости алгоритма для стационарного случая, т.е. для случая $\theta^*=$ const.

В работе исследуются свойства модифицированного регуляризованного алгоритма

$$\hat{\theta}(k+1)=\hat{\theta}(k)+\gamma\frac{e(k+1)x(k+1)}{\left\| x(k+1)\right\|^2+\delta}, \qquad (4)$$

где $\gamma>0$ — некоторый параметр, влияющий на скорость сходимости алгоритма, в задаче оценивания нестационарных параметров, представляющих собой стационарные случайные процессы, которые могут быть описаны модифицированной Марковской моделью первого порядка [12, 16]

$$\theta^*(k+1)=\varepsilon\theta^*(k)+\sqrt{1-\varepsilon}\,S(k+1), \qquad (5)$$

где $S(k+1)=(S_1(k+1),S_2(k+1),\ldots,S_N(k+1))^T$ — вектор случайной последовательности $N\times 1$; $S_i\sim N\left(0,\sigma_s^2\right)$; $0<<\varepsilon<1$ — параметр (при $\varepsilon=1$ имеем стационарную модель).

Исследование сходимости алгоритма

Введем в рассмотрение ошибку оценивания

$$\tilde{\theta}(k+1)=\theta^*(k)-\hat{\theta}(k). \qquad (6)$$

С учетом (5) выражение для $e(k+1)$ можно записать следующим образом:

$$\begin{aligned} e(k+1)&=\tilde{\theta}^T(k)x(k+1)+\xi(k+1)=\\ &=\tilde{\theta}^T(k)x(k+1)+(\varepsilon-1)\theta^{*T}(k)x(k+1)+\\ &+\sqrt{1-\varepsilon}\,S^T(k+1)x(k+1)+\xi(k+1). \end{aligned} \qquad (7)$$

В предположении нормальности помехи можно записать

$$M\left\{e^2(k+1)\right\}=\sigma_\xi^2+\sigma_x^2 M\left\{\left\|\tilde\theta(k+1)\right\|^2\right\}, \quad (8)$$

где М — символ математического ожидания.

Принимая во внимание (5), представим ошибку оценивания (6) так:

$$\tilde\theta(k+1)=\tilde\theta(k)+(\varepsilon-1)\theta^*(k)+\sqrt{1-\varepsilon^2}S(k+1)-$$
$$-\gamma\frac{x(k+1)}{\left\|x(k+1)\right\|^2+\delta}\left[\xi(k+1)+\tilde\theta^T(k)x(k+1)\right] \quad (9)$$

Предполагается, что компоненты вектора ошибки оценивания $\tilde\theta(k)$ подчиняются нормальному закону распределения с $\tilde{\tilde\theta}_i(k)=M\left\{\tilde\theta_i(k)\right\}$ и дисперсией $\sigma_i^2(k)$ [17].

Вычислим $M\left\{\tilde\theta(k+1)\right\}$.

Учитывая [10, 18]

$$M\left\{\frac{1}{\left\|x(k+1)\right\|^2+\delta}\right\}=$$
$$=\frac{1}{(N-2)\sigma_x^2+\delta}\left(1-\frac{2\delta\sigma_x^2}{\left((N-2)\sigma_x^2+\delta\right)^2}\right); \quad (10)$$

$$M\left\{\frac{x(k+1)x^T(k+1)}{\left\|x(k+1)\right\|^2+\delta}\right\}=\frac{1}{N}I-M\left\{\frac{\delta}{\left\|x(k+1)\right\|^2+\delta}\right\}I, \quad (11)$$

где I — единичная матрица $N\times N$, получим

$$M\left\{\tilde\theta(k+1)\right\}=$$
$$=M\left\{\tilde\theta(k)\right\}+(\varepsilon-1)M\left\{\theta(k)\right\}+$$
$$+\sqrt{1-\varepsilon^2}M\left\{S(k+1)\right\}-\gamma M\left\{\frac{x(k+1)\xi(k+1)}{\left\|x(k+1)\right\|^2+\delta}\right\}-$$
$$-\gamma\frac{1}{N}M\left\{\tilde\theta(k)\right\}+\gamma M\left\{\frac{\delta}{\left\|x(k+1)\right\|^2+\delta}\right\}M\left\{\tilde\theta(k)\right\}=$$
$$=M\left\{\tilde\theta(k)\right\}+(\varepsilon-1)M\left\{\theta(k)\right\}+ \quad (12)$$
$$+\sqrt{1-\varepsilon^2}M\left\{S(k+1)\right\}-\gamma\frac{1}{N}M\left\{\tilde\theta(k)\right\}+$$
$$+\frac{1}{(N-2)\sigma_x^2+\delta}\left(1-\frac{2\delta\sigma_x^2}{\left((N-2)\sigma_x^2+\delta\right)^2}\right)M\left\{\tilde\theta(k)\right\}=$$
$$=1-\gamma\frac{1}{N}\left(1-\frac{\delta N}{(N-2)\sigma_x^2+\delta}\left(1-\frac{2\delta\sigma_x^2}{\left((N-2)\sigma_x^2+\delta\right)^2}\right)\right)M\left\{\tilde\theta(k)\right\}.$$

Из полученного выражения (12) следует, что для сходимости алгоритма (4) в среднем необходимо выполнение неравенства

$$\left|1-\gamma\frac{1}{N}\left(1-\frac{\delta N}{(N-2)\sigma_x^2+\delta}\left(1-\frac{2\delta\sigma_x^2}{\left((N-2)\sigma_x^2+\delta\right)^2}\right)\right)\right|<1 \quad (13)$$

или

$$0<\gamma\frac{1}{N}\frac{\left((N-2)\sigma_x^2+\delta\right)^3-\delta N\left((N-2)\sigma_x^2+\delta\right)^2+2N\delta\sigma_x^2}{\left((N-2)\sigma_x^2+\delta\right)^3}<2$$

т.е.

$$0<\gamma<\frac{2N\left((N-2)\sigma_x^2+\delta\right)^3}{\left((N-2)\sigma_x^2+\delta\right)^3-\delta N\left((N-2)\sigma_x^2+\delta\right)^2+2N\delta\sigma_x^2}. \quad (14)$$

При выполнении (14) алгоритм (4) сходится в среднем, т.е.

$$\lim_{k\to\infty}M\left\{\tilde\theta(k+1)\right\}=0.$$

Из неравенства (14) для классического алгоритма Качмажа ($\delta=0$) следует известное условие $0<\gamma<2N$.

Для исследования сходимости в среднеквадратичном умножим (9) слева на $\tilde\theta^T(k+1)$.

$$\left\|\tilde\theta(k+1)\right\|^2=\left\|\tilde\theta(k)\right\|^2+$$
$$+2\tilde\theta^T(k)\theta(k)(\varepsilon-1)+(\varepsilon-1)^2\left\|\tilde\theta(k)\right\|^2+$$
$$+2\tilde\theta^T(k)S(k+1)\sqrt{1-\varepsilon^2}+$$
$$+2(\varepsilon-1)\sqrt{1-\varepsilon}\,\theta^T(k)S(k)+(1-\varepsilon^2)\left\|S(k+1)\right\|^2-$$
$$-2\gamma\frac{\tilde\theta^T(k)x(k+1)}{\left\|x(k+1)\right\|^2+\delta}\left[\xi(k+1)+\tilde\theta^T(k)x(k+1)\right]-$$
$$-2\gamma\frac{(\varepsilon-1)\theta^T(k)x(k+1)}{\left\|x(k+1)\right\|^2+\delta}\left[\xi(k+1)+\tilde\theta^T(k)x(k+1)\right]-$$
$$-2\gamma\frac{\sqrt{1-\varepsilon^2}x^T(k+1)S(k)}{\left\|x(k+1)\right\|^2+\delta}\left[\xi(k+1)+\tilde\theta^T(k)x(k+1)\right]+$$
$$+\gamma^2\frac{\left\|x(k+1)\right\|^2}{\left(\left\|x(k+1)\right\|^2+\delta\right)^2}\left[\xi(k+1)+\tilde\theta^T(k)x(k+1)\right]^2.$$

Вычислим $M\left\{\left\|\tilde\theta(k+1)\right\|^2\right\}$.

С учетом статистических свойств сигналов и помех имеем

$$M\left\{\left\|\tilde{\theta}(k+1)\right\|^2\right\}=M\left\{\left\|\tilde{\theta}(k)\right\|^2\right\}+2(1-\varepsilon)\sigma_S^2-2\gamma M\left\{\frac{(\tilde{\theta}^T(k)x(k+1))^2}{\|x(k+1)\|^2+\delta}\right\}+$$
$$+\gamma^2 M\left\{\frac{\|x(k+1)\|^2}{(\|x(k+1)\|^2+\delta)^2}\xi^2(k+1)\right\}+\gamma^2 M\left\{\frac{(\tilde{\theta}^T(k)x(k+1))^2\|x(k+1)\|^2}{(\|x(k+1)\|^2+\delta)^2}\right\}. \quad (15)$$

Воспользовавшись соотношениями (10), (11), получаем

$$M\left\{\frac{(\tilde{\theta}^T(k)x(k+1))^2}{\|x(k+1)\|^2+\delta}\right\}=\left[\frac{1}{N}I-\frac{\delta}{(N-2)\sigma_x^2+\delta}\left(1-\frac{2\sigma_x^2\delta}{\left((N-2)\sigma_x^2+\delta\right)^2}\right)\right]M\left\{\left\|\tilde{\theta}(k)\right\|^2\right\};$$

$$M\left\{\frac{\|x(k+1)\|^2}{(\|x(k+1)\|^2+\delta)^2}\xi^2(k+1)\right\}=\frac{(N-2)\sigma_x^2\sigma_\xi^2}{\left((N-2)\sigma_x^2+\delta\right)^2};$$

$$M\left\{\frac{\|x(k+1)\|^4}{(\|x(k+1)\|^2+\delta)^2}\right\}=-2\delta\left(\frac{1}{(N-2)\sigma_x^2+\delta}\left(1-\frac{2\sigma_x^2\delta}{\left((N-2)\sigma_x^2+\delta\right)^2}\right)-\frac{\delta^2}{\left((N-2)\sigma_x^2+\delta\right)^2}\right)=$$

$$=-2\delta\frac{\left((N-2)\sigma_x^2+\delta\right)^2-2\sigma_x^2\delta}{\left((N-2)\sigma_x^2+\delta\right)^3}+\frac{\delta^2}{\left((N-2)\sigma_x^2+\delta\right)^2}=\frac{\delta^2\left((N-2)\sigma_x^2+\delta\right)-2\delta\left((N-2)\sigma_x^2+\delta\right)^2+4\delta^2\sigma_x^2}{\left((N-2)\sigma_x^2+\delta\right)^3}.$$

После подстановки данных выражений в (15) и несложных преобразований имеем

$$M\left\{\left\|\tilde{\theta}(k+1)\right\|^2\right\}=\left\{1-\frac{\gamma}{N}\left[(2-\gamma)-\frac{2(\gamma-1)\delta\left((N-2)\sigma_x^2+\delta\right)^2+2\sigma_x^2\delta}{\left((N-2)\sigma_x^2+\delta\right)^3}-\right.\right.$$
$$\left.\left.-\frac{\gamma\delta^2}{\left((N-2)\sigma_x^2+\delta\right)^2}\right]\right\}M\left\{\left\|\tilde{\theta}(k)\right\|^2\right\}+\gamma^2\frac{(N-2)\sigma_x^2\sigma_\xi^2}{\left((N-2)\sigma_x^2+\delta\right)^2}+2(1-\varepsilon)\sigma_S^2. \quad (16)$$

Из соотношения (16) можно получить выражение для асимптотической ошибки оценивания

$$M\left\{\left\|\tilde{\theta}(\infty)\right\|^2\right\}=\frac{\gamma^2N(N-2)\sigma_x^2\sigma_\xi^2\left((N-2)\sigma_x^2+\delta\right)+2N(1-\varepsilon)\sigma_S^2\left((N-2)\sigma_x^2+\delta\right)^3}{\gamma[(2-\gamma)\left((N-2)\sigma_x^2+\delta\right)^3-2(\gamma-1)\delta\left((N-2)\sigma_x^2+\delta\right)^2+2\sigma_x^2\delta^2-\gamma\delta^2\left((N-2)\sigma_x^2+\delta\right)]}. \quad (17)$$

Учитывая (17), можно записать следующее выражение для асимптотической ошибки идентификации (8):

$$M\left\{e^2(\infty)\right\}=\sigma_\xi^2+\frac{\gamma^2N(N-2)\sigma_x^2\sigma_\xi^2\left((N-2)\sigma_x^2+\delta\right)+2N(1-\varepsilon)\sigma_S^2\left((N-2)\sigma_x^2+\delta\right)^3}{\gamma[(2-\gamma)\left((N-2)\sigma_x^2+\delta\right)^3-2(\gamma-1)\delta\left((N-2)\sigma_x^2+\delta\right)^2+2\sigma_x^2\delta^2-\gamma\delta^2\left((N-2)\sigma_x^2+\delta\right)]}.$$

Из полученных соотношений следуют неасимптотические и асимптотические оценки для алгоритма Качмажа (Уидроу-Хоффа). Для стационарного случая:

- $\gamma=1$, $\delta=0$, $\varepsilon=1$

$$M\left\{\left\|\tilde{\theta}(k+1)\right\|^2\right\}=\left(1-\frac{1}{N}\right)M\left\{\left\|\tilde{\theta}(k)\right\|^2\right\}+\frac{\sigma_\xi^2}{(N-2)\sigma_x^2},$$

$$M\left\{\left\|\tilde{\theta}(\infty)\right\|^2\right\}=\frac{N\sigma_\xi^2}{(N-2)\sigma_x^2};$$

- $\gamma \neq 1, \delta = 0, \varepsilon = 1$

$$M\left\{\left\| \tilde{\theta}(\infty) \right\|^2\right\} = \frac{\gamma N \sigma_\xi^2}{(2-\gamma)(N-2)\sigma_x^2}.$$

Для нестационарного случая:

- $\gamma = 1, \delta = 0, \varepsilon \neq 1$

$$M\left\{\left\| \tilde{\theta}(k+1) \right\|^2\right\} = \left(1 - \frac{1}{N}\right) M\left\{\left\| \tilde{\theta}(k) \right\|^2\right\} + \\ + \frac{\sigma_\xi^2}{(N-2)\sigma_x^2} + 2(1-\varepsilon)\sigma_S^2;$$

$$M\left\{\left\| \tilde{\theta}(\infty) \right\|^2\right\} = \frac{N\sigma_\xi^2}{(N-2)\sigma_x^2} + 2(1-\varepsilon)\sigma_S^2.$$

- $\gamma \neq 1, \delta = 0, \varepsilon \neq 1$

$$M\left\{\left\| \tilde{\theta}(k+1) \right\|^2\right\} = \left(1 - \gamma(2-\gamma)\frac{1}{N}\right) M\left\{\left\| \tilde{\theta}(k) \right\|^2\right\} + \\ + \frac{\gamma^2\sigma_\xi^2}{(N-2)\sigma_x^2} + 2(1-\varepsilon)\sigma_S^2;$$

$$M\left\{\left\| \tilde{\theta}(\infty) \right\|^2\right\} = \frac{\gamma N \sigma_\xi^2}{(2-\gamma)(N-2)\sigma_x^2} + \frac{2(1-\varepsilon)\sigma_S^2}{\gamma(2-\gamma)}.$$

Заключение

Как показали результаты исследований, использование регуляризирующей добавки в алгоритмах идентификации, улучшая устойчивость алгоритмов, приводит к некоторому замедлению процесса построения модели. Определены условия сходимости регуляризованного алгоритма Качмажа при оценивании нестационарных параметров при наличии помех измерения.

Полученные неасимптотические и асимптотические оценки достаточно общие и зависят как от степени нестационарности объекта, так и от статистических характеристик помех. Если же эти параметры не известны, следует воспользоваться какой-либо рекуррентной процедурой их оценивания и использовать получаемые оценки для уточнения входящих в алгоритмы параметров.

ЛИТЕРАТУРА

1. *Dupac V.* A dynamic stochastic approximation method. Ann. Math. Statist., 1965, 36, pp. 1695–1702.
2. *Цыпкин Я.З.*, *Каплинский А.И.*, *Ларионов К.А.* Алгоритмы адаптации и обучения в нестационарных условиях. Изв. АН СССР. Техн. Кибернетика, 1970, 5, с. 9–21.
3. *Kaczmarz S.* Angenherle Auflsung von Systemen linearer Gleichungen. Bull. Int. Acad. Polon. Sci. Lett., C 1, Sci. Math. Nat., Ser. A-1937, pp. 355–357.
4. *English* translation: Kaczmarz S. Approximate solution of systems of linear equations, Int. J. of Control, 1993, 57, p. 1269–1271.
5. *Nagumo I.*, *Noda A.* A learning method for system identification, IEEE Trans. Autom. Control, 1967, AC-12, N3, p. 282–287.
6. *Округ А.И.* Динамический алгоритм Качмажа, Автоматика и телемеханика, 1981, 1, С. 74–79.
7. *Widrow B.*, *Hoff M.* Adaptive switching circuits, IRE WESCON Convention Record. Part 4. New York: Institute of Radio Engineers, 1960, pp. 96–104.
8. *Widrow B.*, *Lehr M.* 30 years of adaptive neural networks: perceptron, madaline, and backpropagation, Proc. IEE, 1990, vol. 78, 9, pp. 1415–1442.
9. *Чадеев В.М.* Определение динамических характеристик объектов в процессе их нормальной эксплуатации для целей самонастройки, Автоматика и телемеханика, 1964, Т. 25, 9, С. 1302–1306.
10. *Райбман Н.С.*, *Чадеев В.М.* Адаптивные модели в системах управления, М.: Сов. Радио, 1966, 156 с.
11. *Руденко О.Г.* Оценка скорости сходимости одношаговых устойчивых алгоритмов идентификации, Доклады АН УССР, Сер. А. Физ.-мат. и техн. науки, 1982, №1, С.64–66.
12. *Ciochina S.*, *Paleologu C.*, *Benesty J.* An optimized NLMS algorithm for system identification, Signal Processing, 2016, 118, pp. 115–121
13. *Khong A.W.H., Naylor P.A.* Selective-tap adaptive filtering with performance analysis for identification of timevarying systems, IEEE Trans. Audio, Speech, and Language Processing,2007, vol. 15, pp. 1681–1695.
14. *Benesty J., Paleologu C.*, *Ciochin S.* On regularization in adaptive filtering, IEEE Trans. Audio, Speech, Language Process, 2011, vol. 19, pp. 1734–1742.
15. *Paleologu C.*, *Ciochin S.*, *Benesty J.*, *Grant S.* An overview on optimized NLMS algorithms for acoustic echo cancellation, EURASIP J. Adv. Sig. Proc., 2015, 97, 19 p.
16. *Bershad N.J.*, *McLaughlin S.*, *Cowan C.F.N.* Performance comparison of RLS and LMS algorithms for tracking a first order Markov communications channel, Proc. IEEE Int. Symp. on Circuits and Syst., 1990, pp. 266–270.

17. *Wagner K.*, *Doroslovacki M.* Towards analytical convergence analysis of proportion-type NLMS algorithms, Proc. IEEE Int. Conf. Acoustics Speech Signal Processing, 2008, p.3825–3828.
18. *Либероль Б.Д.*, *Руденко О.Г.*, *Бессонов А.А.* Исследование сходимости одношаговых адаптивных алгоритмов идентификации, Проблемы управления и информатики, 2018, 5, с. 19–32.

Поступила 09.01.2019

REFERENCES

1. Dupac, V., 1965. “A dynamic stochastic approximation method”. Ann. Math. Statist., 36, pp. 1695–1702.
2. Tsypkin, Ya.Z., Kaplinskiy, A.I., Larionov, K.A., 1970. “Algoritmy adaptatsii i obucheniya v nestatsionarnykh usloviyakh”, Izv. AN SSSR. Tekhn. Kibernetika, 5, pp. 9–2. (In Russian).
3. Kaczmarz, S. Angenherle Auflsung von Systemen linearer Gleichungen. Bull. Int. Acad. Polon. Sci. Lett., C 1, Sci. Math. Nat., Ser. A-1937, pp. 355–357.
4. English translation: Kaczmarz, S., 1993. Approximate solution of systems of linear equations, Int. J. of Control, 57, pp. 1269–1271.
5. Nagumo, I., Noda, A., 1967. "A learning method for system identification", IEEE Trans. Autom. Control, AC-12, N3, pp. 282–287.
6. Okrug, A.I., 1981. "Dinamicheskiy algoritm Kachmazha", Avtomatika i telemekhanika, 1, pp. 74—79. (In Russian).
7. Widrow, B., Hoff, M., 1960. "Adaptive switching circuits", IRE WESCON Convention Record. Part 4. New York: Institute of Radio Engineers, pp. 96–104.
8. Widrow, B., Lehr, M., 1990. "30 years of adaptive neural networks: perceptron, madaline, and backpropagation", Proc. IEE, 78 (9), pp. 1415–1442.
9. Chadeyev, V.M., 1964. "Opredeleniye dinamicheskikh kharakteristik obektov v protsesse ikh normal’noy ekspluatatsii dlya tseley samonastroyki", Avtomatika i telemekhanika, 25 (9), pp. 1302–1306. (In Russian).
10. Raybman, N.S., Chadeyev, V.M., 1966. Adaptivnyye modeli v sistemakh upravleniya, M.: Sov. Radio, 156 p. (In Russian).
11. Rudenko, O.G., 1982. "Otsenka skorosti skhodimosti odnoshagovykh ustoychivykh algoritmov identifikatsii". Doklady AN USSR, Ser. A. Fiz-mat i tekhn. nauki, 1, pp.64–66. (In Russian).
12. Ciochina, S., Paleologu, C., Benesty, J., 2016. "An optimized NLMS algorithm for system identification", Signal Processing, 118, pp. 115–121.
13. Khong, A.W.H., Naylor, P.A., 2007. "Selective-tap adaptive filtering with performance analysis for identification of timevarying systems", IEEE Trans. Audio, Speech, and Language Processing, 15, pp. 1681–1695.
14. Benesty, J., Paleologu, C., Ciochin, S., 2011. "On regularization in adaptive filtering", IEEE Trans. Audio, Speech, Language Process, 19, pp. 1734–1742.
15. Paleologu, C., Ciochin, S., Benesty, J., Grant, S., 2015. An overview on optimized NLMS algorithms for acoustic echo cancellation, EURASIP J. Adv. Sig. Proc., 97, 19 p.
16. Bershad, N.J., McLaughlin, S., Cowan, C.F.N., 1990. "Performance comparison of RLS and LMS algorithms for tracking a first order Markov communications channel", Proc. IEEE Int. Symp. on Circuits and Syst., pp. 266–270.
17. Wagner, K., Doroslovacki, M., 2008. "Towards analytical convergence analysis of proportion-type NLMS algorithms", Proc. IEEE Int. Conf. Acoustics Speech Signal Processing, pp.3825–3828.
18. Liberol', B.D., Rudenko, O.G., Bessonov, A.A., 2018. "Issledovaniye skhodimosti odnoshagovykh adaptivnykh algoritmov identifikatsii". Problemy upravleniya i informatiki., 5, pp. 19–32. (In Russian).

Received 09.01.2019

O.G. Rudenko, Doctor of Technical Sciences, Professor, Head of Information Systems Departament
at Simon Kuznets Kharkiv National University of Economics,
Nauky ave., 9-A, Kharkiv, 61166, Ukraine,
oleg.rudenko@hneu.net

O.O. Bezsonov, Doctor of Technical Sciences, Assosciate Professor, Professor of Information Systems Departament
at Simon Kuznets Kharkiv National University of Economics,
Nauky ave., 9-A, Kharkiv, 61166, Ukraine,
oleksandr.bezsonov@hneu.net

THE REGULARIZED ADALINE LEARNING ALGORITHM
FOR THE PROBLEM OF EVALUATION OF NON-STATIONARY PARAMETERS

Introduction. Among the most simple in terms of calculation of one-step algorithms for identification, the most effective are Kachmazh and Nagumo-Noda algorithms. However, their speed is often not sufficient in the evaluation of non-stationary parameters. Knowledge (or approximation) of the drift law allows us to obtain an effective algorithms for tracking non-stationary parameters. It should, however, be noted that the errors in the task of the parameters change can lead to the properties loss of the algorithm convergence. The lack of information about the nature of drift requires the development of identification algorithms that use the minimum amount of information and maintain efficiency over a wide range of variation of parameters. Therefore, it is natural to study the dynamic properties of specific algorithms and determine the maximum achievable accuracy of tracking, which will ultimately determine the most effective areas to use the algorithms and develop the recommendations for their practical application.

Purpose is to study the properties of a modified regularized Kachmazh algorithm, develop recommendations for its practical application.

Methods. Research methods are based on the theory of identification. On its basis, the properties of the modified regularized Kachmazh algorithm are investigated. Also, methods of simulation is used, which allow to confirm the effectiveness of the obtained results and to develop the recommendations for their practical use.

Results. The conditions of the regularized Kachmazh algorithm convergence are determined in the estimation of nonstationary parameters in the presence of noise disturbances. Non-asymptotic and asymptotic estimates are rather general and depend on both the degree of non-stationary object and the statistical characteristics of the noise.

Conclusion. As the results of the research have shown, the use of the regularization application in identification algorithms improves the stability of the algorithms, but leads to some slowdown in the process of the model constructing. The convergence conditions of the regularized Kachmazh algorithm are determined in the estimation of the nonstationary parameters in the presence of the noise disturbances. Non-asymptotic and asymptotic estimates are rather general and depend on both the degree of non-stationary object and the statistical characteristics of the noise. If these parameters are not known, then it is necessary to use some recursive procedure for their evaluation and use the obtained estimates to clarify the parameters included in the algorithms.

***Keywords**: regularized algorithm, convergence, asymptotic estimates, recurrent procedure, law of drift.*

О.Г. Руденко, д-р технічних наук, проф., зав. кафедрою інформаційних систем,
Харківський національний економічний ун-т ім. Семена Кузнеця,
Харків, 61166, просп. Науки, 9-А, Україна,
oleg.rudenko@hneu.net

А.А. Безсонов, д-р технічних наук, доцент, проф. кафедри інформаційних систем,
Харківський національний економічний ун-т ім. Семена Кузнеця,
Харків, 61166, пр. Науки, 9-А, Україна,
oleksandr.bezsonov@hneu.net

РЕГУЛЯРИЗОВАНИЙ АЛГОРИТМІ НАВЧАННЯ АДАЛІНИ В ЗАДАЧІ ОЦІНЮВАННЯ НЕСТАЦІОНАРНИХ ПАРАМЕТРІВ

Вступ. Серед простих в обчислювальному відношенні однокрокових алгоритмів ідентифікації найбільш ефективними є алгоритми Качмажа і Нагумо-Ноди. Однак їх швидкодія виявляється найчастіше недостатньою при оцінюванні нестаціонарних параметрів. Знання (або апроксимація) закону дрейфу дозволяє отримувати ефективні алгоритми відстеження нестаціонарних параметрів. Однак слід відзначити, що помилки в описі закону зміни параметрів можуть призвести до втрати властивостей збіжності алгоритму. Відсутність інформації про характер дрейфу вимагає розробки алгоритмів ідентифікації, що використовують мінімальний обсяг інформації і зберігають працездатність в широкому діапазоні варіювання параметрів. Тому природним є вивчення динамічних властивостей конкретних алгоритмів і визначення максимально досяжної точності стеження, що дозволить визначити області найбільш ефективного використання алгоритмів і розробити рекомендації щодо їх практичного застосування.

Мета статті — дослідження властивостей модифікованого регуляризованого алгоритму Качмажа, розробка рекомендацій щодо його практичного застосування.

Методи дослідження ґрунтуються на теорії ідентифікації. На їх основі було досліджено властивості модифікованого регуляризованого алгоритму Качмажа. Також використано методи імітаційного моделювання, що дозволило підтвердити ефективність отриманих результатів та розробити рекомендації щодо їх практичного використання.

Результат. Визначено умови збіжності регуляризованого алгоритму Качмажа при оцінюванні нестаціонарних параметрів при наявності завад вимірів. Отримані неасимптотичні та асимптотичні оцінки є досить загальними і залежать як від ступеня нестаціонарності об'єкта, так і від статистичних характеристик завад.

Висновок. Як показали результати досліджень, використання в алгоритмах ідентифікації додатку, що регуляризує, покращує стійкість алгоритмів, але призводить до деякого уповільнення процесу побудови моделі. Визначено умови збіжності регуляризованого алгоритму Качмажа при оцінюванні нестаціонарних параметрів при наявності завад вимірів. Отримані неасимптотичні та асимптотичні оцінки є досить загальними і залежать як від ступеня нестаціонарності об'єкта, так і від статистичних характеристик завад. Якщо ж ці параметри не відомі, слід скористатися будь-якою рекурентною процедурою їх оцінювання і використовувати одержані оцінки для уточнення параметрів, що входять в алгоритми.

***Ключові слова**: регуляризований алгоритм, збіжність, асимптотичні оцінки, рекурентна процедура, закон дрейфу.*