

МЕТОДЫ УПРАВЛЕНИЯ И ОЦЕНИВАНИЯ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

УДК 004.852

Б.Д. Либероль, О.Г. Руденко, А.А. Бессонов

ПСЕВДОПРОЕКЦИОННЫЕ АЛГОРИТМЫ ОЦЕНИВАНИЯ, ИСПОЛЬЗУЮЩИЕ ВЗВЕШИВАНИЕ ИНФОРМАЦИИ

Ключевые слова: алгоритм Качмажа, проекционный алгоритм, рекуррентная процедура, асимптотическая оценка, точность идентификации.

Введение

В основе многих задач обработки информации (обработка и фильтрация сложных сигналов, идентификация объектов и управления ими, прогнозирования временных последовательностей и т.д.) лежит задача построения модели вида

$$y_n = c^{*T} x_n + \xi_n, \quad (1)$$

где y_n — наблюдаемый выходной сигнал; $x_n = (x_{1n}, x_{2n}, \dots, x_{Nn})^T$ — вектор входных сигналов $N \times 1$; $c^* = (c_1^*, c_2^*, \dots, c_N^*)^T$ — вектор искомых (оцениваемых) параметров $N \times 1$; ξ_n — помеха; T — символ транспонирования; n — дискретное время.

Наиболее широко используемый на практике квадратичный функционал приводит к различным алгоритмам идентификации, позволяющим получить оценки искомого вектора при нормальных распределениях помехи, т.е. $\xi(k) \sim N(0, \sigma_\xi^2)$.

Исследования алгоритмов стохастической аппроксимации [1, 2] свидетельствуют о том, что они обеспечивают получение требуемого качества оценок при наличии помех. При этом, однако, возникает вопрос повышения скорости сходимости алгоритма оценивания.

Среди наиболее простых в вычислительном отношении одношаговых алгоритмов оценивания самыми эффективными являются алгоритмы Качмажа и Нагумо–Ноды [3, 4].

Алгоритм Качмажа, имеющий вид

$$c_n = c_{n-1} + \gamma \frac{y_n - c_{n-1}^T x_n}{\|x_n\|^2} x_n, \quad (2)$$

с параметром $\gamma = 1$ предложен в работе [3] для решения систем линейных алгебраических уравнений. В (2) $\|\bullet\|$ — евклидова норма.

Впоследствии данный алгоритм стал успешно применяться для решения задачи идентификации при построении модели типа (1) [5–10].

В работах [8–10] получены неасимптотические и асимптотические оценки скорости сходимости алгоритма (2) и, в частности, регуляризованного алгоритма Качмажа, обладающего большей вычислительной устойчивостью.

В работах [11–14] рассматривалась проблема ускорения алгоритма Качмажа путем использования не одного, а ряда предыдущих измерений. В [12–16] предложены многошаговые проекционные алгоритмы, которые при построении оценки на очередном n -м шаге использовали информацию об L предыдущих шагах. Так как при реализации таких алгоритмов требуется вычисление обратной матрицы наблюдений $(X_n^L X_n^{LT})^{-1}$ размерности $L \times L$, в [15, 16] установлены свойства случайных псевдообратных матриц и матриц проецирования. Это дало возможность определить скорость сходимости данных алгоритмов, в результате было установлено, что учет в многошаговых проекционных алгоритмах информации об L предыдущих шагах равносильен в смысле скорости сходимости уменьшению размерности исходного пространства N на L . Существенный положительный эффект от использования информации о предыдущих шагах, с одной стороны, и необходимость обращения матрицы, с другой, свидетельствуют о целесообразности разработки алгоритмов, обладающих близкими к проекционным алгоритмам свойствами, но являющихся более простыми в реализации.

Один из таких подходов, использующих аппроксимацию операции ортогонального проецирования вектора ошибки оценивания на линейную оболочку L входных векторов, осуществляемую на основе одношагового проекционного алгоритма Качмажа, изучался в [17].

В данной статье рассматривается модификация алгоритма Качмажа, использующая взвешивание оценок, полученных на ряде предыдущих шагов.

Алгоритм оценивания на основе алгоритма Качмажа

Рассматриваемые здесь методы также основаны на алгоритме Качмажа и являются частным случаем обобщенного проекционного метода. Идея состоит в использовании при построении очередной оценки c_n информации не только об оценке c_{n-1} , но и о ряде предыдущих c_{n-2}, \dots, c_{n-s} , взятых с некоторыми весами $\lambda_1, \lambda_2, \dots, \lambda_s$. Таким образом, оценка c_{n-1} заменяется в методе на $\hat{c}_{n-1} = \sum_{i=1}^s \lambda_i c_{n-1}$, где $\lambda_i > 0$, $\sum_{i=1}^s \lambda_i = 1$. В этом случае, как следует из (2), одношаговый псевдопроекционный метод записывается в виде

$$c_n = \hat{c}_{n-1} + \frac{y_n - \hat{c}_{n-1}^T x_n}{\|x_n\|^2} x_n, \quad (3)$$

Утверждение 1. Если в методе (3) $\lambda_i > 0$, $i = \overline{1, s}$, то для данного метода справедлива оценка

$$\alpha_n = (1 - kN^{-1})\alpha_{n-1}, \quad (4)$$

где $\alpha_i = M\{\|\Theta_i\|^2\}$; $\theta_i = c^* - c_i$ — ошибка оценивания; $k = 1 - \left(1 - \bar{\lambda}^{-1}\right)^2$;
 $\bar{\lambda} = \lambda_1 + 2\lambda_2 + \dots + s\lambda_s$.

Доказательство. Введем ошибку оценивания $\theta_i = c^* - c_i$.

Тогда алгоритм (3) можно записать

$$\theta_n = \sum_{i=1}^s \lambda_i (I - P_n^{(1)}) \theta_{n-1}, \quad (5)$$

где $P_n^{(1)} = x_n x_n^T / \|x_n\|^2$ — матрица проецирования на вектор x_n .

Для анализа свойств данного алгоритма введем вектор $\vartheta_n^T = (\theta_n^T, \theta_{n-1}^T, \dots, \theta_{n-s+1}^T)$ размерности $NS \times 1$. Тогда (5) можно записать так: $\vartheta_n = F_n \vartheta_{n-1}$, где матрица F_n представляется в блочном ($S \times S$ блоков размерности $N \times N$ каждый) виде следующим образом:

$$F_n = \begin{pmatrix} \lambda_1(I - P_n^{(1)}) & \dots & \lambda_s(I - P_n^{(1)}) \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & I & 0 \end{pmatrix}. \quad (6)$$

Сходимость (3) будем оценивать по критерию $M\{\|\vartheta_n\|^2\} = \alpha_n$. А так как $M\{\|\vartheta_n\|^2\} = M\{\|\theta_n\|^2 + \dots + \|\theta_{n-s+1}\|^2\}$, то если α_n сходится со скоростью геометрической прогрессии со знаменателем ε , т.е. $\alpha_n \sim \varepsilon^n$, $M\{\|\vartheta_n\|^2\}$ будет сходиться с той же скоростью.

Введем матрицу ковариации ошибок $\Phi_n = M\{\vartheta_n \vartheta_n^T\}$, которая в нашем случае имеет вид

$$\Phi_n = M\{F_n \vartheta_{n-1} \vartheta_{n-1}^T F_n^T\} = M\{F_n \Phi_{n-1} F_n^T\}. \quad (7)$$

Поскольку $M\{P_n^{(1)}\} = N^{-1}I$, $(P_n^{(1)})^2 = P_n^{(1)}$, $M\{P_n^{(1)2}\} = N^{-1}I$, то, обозначив $l = 1 - N^{-1}$ и принимая во внимание, что $M\{I - P_n^{(1)}\} = lI$, $M\{(I - P_n^{(1)})^2\} = lI$, для вывода рекуррентного соотношения для Φ_n , усредненного по ϑ_n , поступим следующим образом.

Представим F_n в виде $F_n = V_n + U_n$, где

$$V_n = \begin{pmatrix} \lambda_1(I - P_n^{(1)}) & \dots & \lambda_s(I - P_n^{(1)}) \\ 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 0 \end{pmatrix}; \quad U_n = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & I & 0 \end{pmatrix}.$$

В свою очередь, матрицы V_n и U_n можно записать в виде прямых произведений [18]:

$$V_n = v \otimes (I - P_n^{(1)}), \quad U_n = u \otimes I, \quad (8)$$

где \otimes — символ прямого произведения

$$v = \begin{pmatrix} \lambda_1 & \dots & \lambda_s \\ 0 & \dots & \dots \\ 0 & \dots & 0 \end{pmatrix}; \quad u = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 \end{pmatrix}. \quad (9)$$

С учетом свойства прямого произведения $(A \otimes B)^T = A^T \otimes B^T$ и того, что $P_n^{(1)} = P_n^{(1)T}$, имеем

$$\Phi_n = M\{(u \otimes I + v^T \otimes (I - P_n^{(1)}))\Phi_{n-1}(u \otimes I - v \otimes (I - P_n^{(1)}))\}$$

или (так как u не зависит от v)

$$\begin{aligned} \Phi_n = & (u^T \otimes I)\Phi_{n-1}(u \otimes I) + Mv\{(v^T \otimes (I - P_n^{(1)}))\Phi_{n-1}((u \otimes I) + \\ & + (u^T \otimes I)\Phi_{n-1}(v \otimes (I - P_n^{(1)})))\} + M\{(v^T \otimes (I - P_n^{(1)}))\Phi_{n-1}(v \otimes (I - P_n^{(1)}))\}. \end{aligned} \quad (10)$$

Из (7) и начального условия $\Phi_0 = I$ следует, что Φ_n также будет иметь структуру прямого произведения матриц, т.е. решение (10) можно искать в виде

$$\Phi_n = \Gamma_n \otimes I, \quad (11)$$

где I — единичная матрица $N \times N$; Γ_n — матрица $S \times S$; $\Gamma_n = I_S$.

Учитывая свойства прямого произведения

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2), \quad (12)$$

после подстановки (11) в (10) получаем

$$\begin{aligned} \Gamma_n \otimes I = & (u^T \Gamma_{n-1} u) \otimes I + \\ & + M\{(v^T \Gamma_{n-1} u + u^T \Gamma_{n-1} v) \otimes (I - P_n^{(1)})\} + M\{(v^T \Gamma_{n-1} u) \otimes (I - P_n^{(1)})\} \end{aligned}$$

или

$$\Gamma_n \otimes I = (u^T \Gamma_{n-1} u + l(v^T \Gamma_{n-1} u + u^T \Gamma_{n-1} v) + lv^T \Gamma_{n-1} v) \otimes I. \quad (13)$$

Тот факт, что $M\{P_n^{(1)}\} = M\{(P_n^{(1)})^2\}$ — скалярная матрица $\sim I$, позволяет представить Φ_{n-1} в виде прямого произведения (11) и получить для Γ_n рекуррентное соотношение

$$\Gamma_n = u^T \Gamma_{n-1} u + l(v^T \Gamma_{n-1} u + u^T \Gamma_{n-1} v) + l(v^T \Gamma_{n-1} v). \quad (14)$$

Введем матрицу $W = u^T + lv^T$. Добавляя и вычитая из правой части (14) член $l^2 v^T \Gamma_{n-1} v$ и учитывая, что $l - l^2 = lN^{-1} = p$, получаем

$$\Gamma_n = W \Gamma_{n-1} W^T + pv^T \Gamma_{n-1} v, \quad (15)$$

где

$$W^T = \begin{pmatrix} l\lambda_1 & \dots & l\lambda_{s-1} & l\lambda_s \\ 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 \end{pmatrix}.$$

Матричные уравнения типа (15) можно преобразовать к более удобному для анализа виду, если рассматривать матрицу Γ_n размерности $S \times S$ как сокращенную

$$\text{запись вектора размерности } S^2: \bar{\Phi}_n = \begin{pmatrix} \Gamma_{i1}^T \\ \Gamma_{i2}^T \\ \dots \\ \Gamma_{is}^T \end{pmatrix}, \text{ где } \Gamma_{ik} \text{ — } k\text{-я строка матрицы } \Gamma_i.$$

Тогда для вектора $\bar{\Phi}_n$ рекуррентное соотношение (15) запишем так:

$$\bar{\Phi}_n = G\bar{\Phi}_{n-1}, \quad (16)$$

где $G = W \otimes W + p v^T \otimes v^T$.

Поведение $\bar{\Phi}_n$ при $n \rightarrow \infty$ определяется свойствами матрицы G , в частности, ее спектральным радиусом μ_G , поскольку $\|\bar{\Phi}_n\| = \|G^n \bar{\Phi}_0\| \sim \mu_0^n$ при $n \rightarrow \infty$, G — неотрицательная матрица (так как выбранные веса $\lambda_i > 0$). Матрица также неприводима (это следует из того, что ее направленный граф сильно связан [18]). Поэтому матрица G удовлетворяет условиям теоремы Перрона–Фробениуса, из первой части которой [18] следует, что G имеет положительное собственное значение, равное спектральному радиусу матрицы G .

Оценку снизу для спектрального радиуса можно получить, если учесть, что для положительных матриц спектральный радиус — монотонная функция матрицы [18]. А поскольку матрица $v^T \otimes v^T$ неотрицательна, имеем $\mu_G \geq \mu_{W \otimes W}$. Здесь μ_G — спектральный радиус G ; $\mu_{W \otimes W}$ — спектральный радиус $W \otimes W$.

Спектральный радиус $\mu_{W \otimes W}$ прямого произведения неотрицательных матриц W , имеющих стандартную фробениусову форму, равен квадрату спектрального радиуса W . Это следует из того, что собственные значения прямого произведения матриц есть произведение их собственных значений, а собственными векторами являются прямые произведения их собственных векторов.

Характеристическое уравнение для определения собственных значений W имеет вид

$$\mu^S - l(\lambda_1 \mu^{S-1} + \dots + \lambda_S) = 0.$$

Нетрудно показать, что это уравнение при $\sum_{i=1}^S \lambda_i = 1$ имеет единственный положительный корень μ_0 (при $l=1$, $\mu_0 = 1$). А так как у нас $l = 1 - N^{-1}$, то разлагая μ_0 по малому параметру $p = 1 - l = N^{-1}$, получаем

$$\begin{aligned} \mu_0 &= 1 - \frac{1}{N(\lambda_1 + \dots + \lambda_S)} = 1 - \frac{1}{N\lambda}, \\ \mu_G &\geq (\mu_W)^2 = 1 - \frac{2}{N\lambda} + O\left(\frac{1}{N}\right). \end{aligned} \quad (17)$$

Представим G в виде $G = G_0 + pG_1$, где $G_0 = W \otimes W$; $G_1 = v^T \otimes v^T$.

Так как $p = N^{-1} \ll 1$, второе слагаемое в правой части этого выражения можно рассматривать как возмущение. Тогда (17) будет нулевым приближением в

разложении μ_G по параметру ρ . Член первого порядка можно получить, используя формулы теории возмущений собственных значений.

Пусть матрица W имеет правый собственный вектор x_0 , соответствующий собственному значению μ_W , и левый собственный вектор y_0^T , т.е. $Wx_0 = \mu_W x_0$, $y_0^T W = \mu_W y_0^T$.

Для $G = W \otimes W$ правый и левый собственные векторы, соответствующие $\mu_0 = (\mu_W)^2$, будут равны соответственно $X_0 = x_0 \otimes x_0$ и $Y_0 = y_0^T \otimes y_0^T$. Если выбрать нормировку так, чтобы $y_0^T x_0 = 1$, то $y_0^T X_0 = (y_0^T \otimes y_0^T)(x_0 \otimes x_0) = y_0^T x_0 \otimes y_0^T x_0 = 1$.

Член первого порядка μ_1 в разложении собственных значений матрицы $G = G_0 + \rho G_1$ по $\rho = N^{-1}$ дается формулой [18] $\mu_0 = Y_0^T G_1 X_0$. Подставляя X_0 и Y_0 и используя (11), получаем

$$\mu_1 = (Y_0^T \otimes Y_0^T)(v^T \otimes v^T)(X_0 \otimes X_0) = (Y_0^T v^T X_0) \otimes (Y_0^T v^T X_0) = (Y_0^T v^T X_0)^2. \quad (18)$$

Поскольку мы ищем добавку с точностью до членов $o(N^{-1})$, а $\rho = N^{-1}$, в выражении (18) достаточно взять нулевые члены разложения по ρ для Y_0^T и X_0 . Из (14) для X_0 и Y_0^T с требуемой точностью имеем выражения для X_0 и Y_0^T из (16) получаем

$$X_0^T = (1, 1 - \lambda_1, \dots, 1 - (\lambda_1 + \dots + \lambda_{S-1}))\zeta, \quad Y_0^T = (1, \dots, 1).$$

Из условия нормировки следует

$$\zeta = \frac{1}{\lambda_1 + 2\lambda_2 + \dots + S\lambda_S} = \frac{1}{\bar{\lambda}}. \quad (19)$$

Подставляя в (18) выражение для v^T из (9), получаем $\mu_1 = (\bar{\lambda})^{-2}$.

Окончательно

$$\mu_G = \mu_0 + \rho\mu_1 = 1 - \frac{2}{N\bar{\lambda}} + \frac{1}{N(\bar{\lambda})^2} + O\left(\frac{1}{N}\right). \quad (20)$$

Введем приведенную характеристику скорости сходимости метода $\mu_G = 1 - \frac{1}{N}\delta$ или $\delta = \frac{1 - \mu_G}{1 - l}$, где $l = 1 - N^{-1}$ — скорость сходимости алгоритма Качмажа, соответствующему случаю

$$\lambda_i = \begin{cases} 1, & i = n - 1, \\ 0, & i \neq n - 1. \end{cases}$$

Очевидно, $\delta(l) = 1$. Из (20) получаем

$$\delta = \frac{1}{\bar{\lambda}} \left(2 - \frac{1}{\bar{\lambda}} \right) = 1 - \left(1 - \frac{1}{\bar{\lambda}} \right)^2. \quad (21)$$

Таким образом, доказано утверждение 1.

Как следует из (19), величина $\bar{\lambda} \geq 1$ и определяется весами λ_i ($i = \overline{1, s}$). При $\lambda_i = \text{const} = s^{-1}$ имеем $\bar{\lambda} = 0,5(S+1)$. В этом случае из (19) получаем, что при $S=1$ $\delta=1$ (алгоритм Качмажа); при $S=2$ $\delta = \frac{8}{9}$; при $S=3$ $\delta = \frac{3}{4}$; при $S=4$ $\delta = 0,64$; при $S=5$ $\delta = \frac{5}{9}$. Эти значения хорошо согласуются с вычисленным на компьютере спектральным радиусом матрицы G .

Рассмотрен случай, когда при построении c_n используется взвешивание оценок c_{n-1}, \dots, c_{n-S} . Возможна модификация метода (3), использующая взвешивание не только оценок, но и поправок.

Рассмотрим одну из таких модификаций:

$$c_n = \lambda_1 c_{n-1} + (1-\lambda_1)c_{n-2} + \lambda_2 \frac{y_n - C_{n-1}^T x_n}{\|x_n\|^2} x_n + (1-\lambda_2) \frac{y_n - C_{n-2}^T x_n}{\|x_n\|^2} x_n. \quad (22)$$

Алгоритм Качмажа отсюда следует при $\lambda_1 = \lambda_2 = 1$.

Утверждение 2. Метод (22) со скоростью, определяемой формулой (4), с $k = (1-\lambda_1)^2$, сходится к области, равной

$$\lim \alpha_n = \frac{\sigma_\xi^2}{(N-2)\sigma_x^2} \left[p \left(1 + \frac{2\lambda_1(1-\lambda_1)(1-p) - 2\lambda_1(\lambda_1 - \lambda_2)^2}{\lambda_1 + p(1-\lambda_2)} \right) \right]^{-1}. \quad (23)$$

Доказательство. Записав (22) относительно θ и введя вектор $\vartheta_n = (\theta_{n-1}^T \theta_{n-2}^T)$ размерности $2N \times 1$, получим $\vartheta_n = F_n \vartheta_{n-1} - q_n$, где

$$F_n = \begin{pmatrix} \lambda_1 I - \lambda_2 P_n^{(1)} & (1-\lambda_1)I - (1-\lambda_2)P_n^{(1)} \\ I & 0 \end{pmatrix};$$

$$q_n = \frac{\xi_n x_n}{\|x_n\|^2}.$$

Рассмотрим матрицу ковариации ошибок $\hat{O}_n = M\{\vartheta_n \vartheta_n^T\}$:

$$\begin{aligned} \hat{O}_n &= M\{\vartheta_n \vartheta_n^T\} = M\{F_n \vartheta_{n-1} \vartheta_{n-1}^T F_n^T\} + \\ &+ M\{F_n \vartheta_{n-1} q_n^T + q_n \vartheta_{n-1}^T F_n^T\} + M\{q_n q_n^T\}. \end{aligned} \quad (24)$$

Так как помехи не коррелированы с полезными сигналами и имеют нулевое математическое ожидание, полученное выражение упростится:

$$\hat{O}_n = M\{\vartheta_n \vartheta_n^T\} = M\{F_n \vartheta_{n-1} \vartheta_{n-1}^T F_n^T\} + M\{q_n q_n^T\}.$$

Представим входящие в данное выражение матрицы в виде прямых произведений матриц:

$$F_n = u \otimes I - v \otimes P_n^{(1)},$$

$$M\{q_n q_n^T\} = \eta \otimes I,$$

где

$$\eta = \frac{\sigma_{\xi}^2}{N(N-2)\sigma_x^2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

С учетом этого

$$\hat{O}_n = M\{(u \otimes I - v \otimes P_n^{(1)})\hat{O}_{n-1}(u^T \otimes I - v^T \otimes P_n^{(1)}) + \eta \otimes I\}. \quad (25)$$

Будем искать решение уравнения (25) в виде (11). По аналогии с (14) для матрицы можно записать рекуррентное соотношение

$$\Gamma_n = u\Gamma_{n-1}u^T - p(v\Gamma_{n-1}u^T + u\Gamma_{n-1}v^T) + pv\Gamma_{n-1}v^T + \eta. \quad (26)$$

Это выражение можно преобразовать к виду (ср. с (16))

$$\bar{\Phi}_n = G\bar{\Phi}_{n-1} + \bar{\eta}, \quad (27)$$

где $\bar{\Phi}_n$ и $\bar{\eta}$ — матрицы Γ_n и η , из элементов которых составлены вектор-столбцы: $\Gamma_n = \begin{pmatrix} \Gamma_{1n} & \Gamma_{1n} \\ \Gamma_{3n} & \Gamma_{3n} \end{pmatrix}$; $\bar{\Phi}_n = (\Gamma_{1n}\Gamma_{2n}\Gamma_{3n}\Gamma_{4n})^T$; $\eta = \begin{pmatrix} \eta & \eta \\ 0 & 0 \end{pmatrix}$; $\bar{\eta} = (\eta \ 0 \ 0 \ 0)^T$.

Матрица G выражается через входящие в уравнения (25) матрицы u и v следующим образом:

$$G = u \otimes u - p(v \otimes u + u \otimes v) + pv \otimes v. \quad (28)$$

Подставляя в (28) выражения для матриц u и v , после несложных преобразований получаем

$$G = \begin{pmatrix} A & B & B & C \\ F & 0 & D & 0 \\ F & D & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad (29)$$

где

$$\begin{aligned} A &= \lambda_1^2 + p(\lambda_2^2 - 2\lambda_1\lambda_2); \\ D &= \lambda_1(1 - \lambda_1) - p[\lambda_2(1 - \lambda_1) + \lambda_1(1 - \lambda_2) - \lambda_2(1 - \lambda_2)]; \\ C &= (1 - \lambda_1)^2 - p[2(1 - \lambda_1)(1 - \lambda_2) - (1 - \lambda_2)^2]; \\ D &= 1 - \lambda_1 - p(1 - \lambda_2); \\ F &= \lambda_1 - \lambda_1\lambda_2. \end{aligned} \quad (30)$$

При отсутствии помех решение (27) имеет вид $\bar{\Phi}_n = G^n \bar{\Phi}_0$, и скорость сходимости метода (22) определяется спектральным радиусом μ_G матрицы G . Оценить скорость сходимости можно, анализируя характеристическое уравнение $\det(\mu I - G) = 0$. Нетрудно видеть, что один корень $\mu_1 = -D = p(1 - \lambda_2) + \lambda_1 - 1$, а остальные три определяются из кубического уравнения

$$\mu^3 - \mu^2(A + D) + \mu(AD - C - 2BF) + CD = 0.$$

Обозначим это уравнение $L(\mu, p, \lambda_1, \lambda_2) = 0$. При $p = 0$ данное уравнение, как легко видеть, имеет корень $\mu = 1$ и разлагается на множители

$$L(\mu, 0, \lambda_1, \lambda_2) = (\mu - 1)(\mu + (1 - \lambda_1))(\mu - (1 - \lambda_1)^2),$$

откуда $\mu_1 = 1, \mu_2 = 1 - \lambda_1, \mu_3 = -(1 - \lambda_1), \mu_4 = (1 - \lambda_1)^2$ (здесь $\mu_2 = -D_{p=0} = 1 - \lambda_1$).

Отсюда видно, что в (22) нельзя брать $\lambda_1 < 0$, так как метод будет расходиться. При выборе $0 < \lambda_1 < 1$ все корни μ_i ($i = 2, 3, 4$) меньше единицы и $\mu_G = \max_i |\mu_i| = 1$.

Поэтому нас интересует поведение максимального по модулю корня μ_1 при $p = 0$. Полагая, что $p = N^{-1} \ll 1$, разложим μ_1 по степеням p и ограничимся линейным членом, т.е. $\mu_1 = 1 + kp + o(p)$. Воспользовавшись представлением элементов матрицы G в виде (30), после несложных вычислений получим, что $k = -(1 - (1 - \lambda_1)^2)$, т.е.

$$\mu_1 = 1 - N^{-1}(1 - (1 - \lambda_1)^2) + o(N^{-1}),$$

откуда следует указанная в утверждении оценка скорости сходимости метода (22). Таким образом, при отсутствии помех $\bar{\Phi}_n = G^n \bar{\Phi}_0$.

При наличии помех решение неоднородного уравнения (27) проводится стандартным способом: ищется частное решение (27) в виде постоянного вектора $\bar{\Phi}_\infty$:

$$\bar{\Phi}_\infty = G\bar{\Phi}_\infty + \bar{\eta}, \quad (31)$$

и если матрица $(I - G)$ обратима, решение (31) имеет вид $\bar{\Phi}_\infty = (I - G)^{-1}\bar{\eta}$. Тогда для $\bar{\varphi}_n = \bar{\Phi}_n - \bar{\Phi}_\infty$ получаем $\bar{\varphi}_n = G\bar{\varphi}_{n-1}$, т.е. $\bar{\varphi}_n$ сходится к нулю так же, как $\bar{\Phi}_n$ при отсутствии помех: $\bar{\varphi}_n = G^n \bar{\varphi}_0$. Окончательно для $\bar{\Phi}_n$ имеем

$$\bar{\Phi}_n = G^n \bar{\Phi}_0 + (I - G^n)\bar{\varphi}_\infty. \quad (32)$$

Первое слагаемое описывает сходимость со скоростью (4) в область, определяемую вторым слагаемым:

$$\lim_{n \rightarrow \infty} \bar{\Phi}_n = \bar{\Phi}_\infty.$$

Используя представление G в виде (29), получаем следующую систему уравнений (31) для определения $\Gamma_{1\infty}, \Gamma_{2\infty}, \Gamma_{3\infty}, \Gamma_{4\infty}$:

$$\begin{cases} A\Gamma_{1\infty} + B\Gamma_{2\infty} + B\Gamma_{3\infty} + C\Gamma_{4\infty} = \Gamma_{1\infty} - \bar{\eta}; \\ F\Gamma_{1\infty} + D\Gamma_{3\infty} = \Gamma_{2\infty}; \\ F\Gamma_{1\infty} + D\Gamma_{2\infty} = \Gamma_{3\infty}; \\ \Gamma_{1\infty} = \Gamma_{4\infty}. \end{cases} \quad (33)$$

Решение системы дает вектор $\bar{\Phi}_\infty$ в виде $\bar{\Phi}_\infty = (\Gamma_{1\infty} \Gamma_{2\infty} \Gamma_{3\infty} \Gamma_{4\infty})^T$, что соответствует теплицевой форме матрицы Φ_n

$$\bar{\Phi}_\infty = \begin{pmatrix} \Gamma_{1\infty} & \Gamma_{2\infty} \\ \Gamma_{2\infty} & \Gamma_{1\infty} \end{pmatrix}, \quad (34)$$

согласующейся с ее интерпретацией как асимптотической матрицы ковариации ошибок.

Решая (33), для $\Gamma_{1\infty} = \lim_{n \rightarrow \infty} M\{\|\theta_n\|^2\}$ получаем

$$\Gamma_{1\infty} = \frac{\eta}{1 - (A + C + \frac{2BF}{F + P})}.$$

После подстановки выражений для A, B, C, F из (30) окончательно имеем

$$\Gamma_{1\infty} = \eta \left[p \left(1 + \frac{2\lambda_1(1-\lambda_1)(1-p) + 2p(\lambda_1 - \lambda_2)^2}{\lambda_1 + p(1-\lambda_2)} \right) \right]^{-1}.$$

Поскольку по определению матрицы Φ_n

$$Sp\Phi_n = M\{\|\theta_n\|^2 + \|\theta_{n-1}\|^2\},$$

предельное значение $\lim_{n \rightarrow \infty} \alpha_n$ выражается через матрицу Φ_∞ по формуле

$$\lim_{n \rightarrow \infty} \alpha_n = \frac{1}{2} Sp\Phi_\infty.$$

Учитывая, что, как следует из (11), $\Phi_\infty = \Gamma_\infty \otimes I$, а

$$\Gamma_\infty \otimes I = \begin{pmatrix} \Gamma_\infty & 0 & \dots & 0 \\ 0 & \Gamma_\infty & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \Gamma_\infty \end{pmatrix},$$

имеем $\lim_{n \rightarrow \infty} \alpha_n = M\Gamma_{1\infty}$, откуда следует требуемое выражение для области сходимости (23).

Из (23) при $\lambda_1 = \lambda_2 = 1$ получаем известное соотношение для алгоритма Качмажа

$$\alpha_k = \lim_{n \rightarrow \infty} \alpha_n = \mu p^{-1} = \frac{N\sigma_\xi^2}{(N-2)\sigma_x^2}.$$

Используя малость параметра $p = N^{-1} \ll 1$, для (23) получаем разложение

$$\alpha_\infty = \alpha_k (3 - 2\lambda_1 - 2p\lambda_1^{-1}(1 + \lambda_2^2 - \lambda_2(1 - \lambda_1)) + o(p))^{-1}.$$

Хотя формально α_∞ будет минимальна при $\lambda_1 \sim \sqrt{p}$, скорость сходимости метода при таком λ_1 будет мала (см. (4)). При $\lambda_1 \sim 1$ область сходимости в нулевом приближении по p равна $\frac{\alpha_k}{3 - 2\lambda_1}$, а минимум этого слагаемого, достигающийся при $\lambda_2 = 0,5(1 + \lambda_1)$, дает следующее выражение для минимально возможной области сходимости при заданном λ_1 :

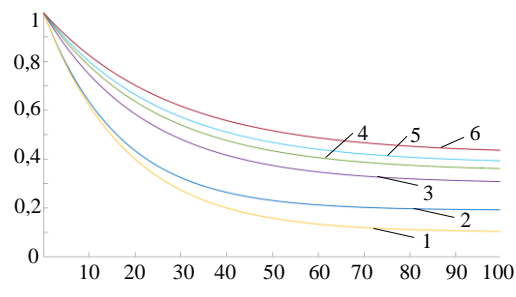
$$\alpha_\infty = \alpha_k (3 - 2\lambda_1 - p(2\lambda_1)^{-1}(4 - (1 + \lambda_1)^2) + o(p))^{-1}.$$

Таким образом, использование взвешенных оценок позволяет при незначительном уменьшении скорости сходимости уменьшить размеры области сходимости (по сравнению с алгоритмом Качмажа с $\gamma = \text{const} = 1$).

Результаты экспериментального исследования работы алгоритмов представлены на рисунке. Здесь 1 соответствует алгоритму Качмажа (2) с $\gamma = 1$; 2 — алго-

ритму Качмажа с $\gamma = 0,75$; 3 — методу (3) с $\lambda_1 = 0,5$, $\lambda_2 = 0,25$, $\lambda_3 = 0,15$, $\lambda_4 = 0,05$, $\lambda_5 = 0,05$; 4 — методу (3) с $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 0,2$; 5 — методу (3) с $\lambda_1 = \lambda_2 = 0,5$; 6 — методу (3) с $\lambda_1 = 0,75$, $\lambda_2 = 0,25$.

Анализ результатов моделирования свидетельствует о том, что используемое усложнение алгоритмов существенного выигрыша в случае стационарного объекта не дает, и кроме того, в алгоритме Качмажа можно выбрать параметр γ так, что при равной с (4), (22) скорости сходимости область его сходимости будет меньше. Этот вывод согласуется с результатами работы [19].



Заклучение

Как показали результаты исследований, использование взвешенных оценок в алгоритмах оценивания при незначительном снижении скорости сходимости алгоритмов позволяет уменьшить размеры их области сходимости (по сравнению с алгоритмом Качмажа с $\gamma = \text{const} = 1$), определяемой соотношением дисперсий полезных сигналов и помех. Анализ результатов моделирования свидетельствует о том, что используемое усложнение алгоритмов в случае идентификации стационарного объекта существенного выигрыша не дает. Более того, в алгоритме Качмажа можно выбрать параметр γ так, что при равной с рассматриваемыми алгоритмами скорости сходимости область его сходимости меньше.

Б.Д. Лібероль, О.Г. Руденко, О.О. Безсонов

ПСЕВДОПРОЕКЦІЙНІ АЛГОРИТМИ ОЦІНЮВАННЯ, ЩО ВИКОРИСТОВУЮТЬ ЗВАЖУВАННЯ ІНФОРМАЦІЇ

Алгоритм Качмажа, який в даний час широко використовується завдяки своїй простоті й ефективності при вирішенні різних технічних завдань, запропоновано в роботі [3] для вирішення систем лінійних алгебраїчних рівнянь. Аналітичні дослідження, проведені в [8–10], дозволили отримати неасимптотичні та асимптотичні оцінки швидкості збіжності цього алгоритму за наявності перешкод вимірів. У роботах [11–14] розглядалася проблема прискорення алгоритму Качмажа шляхом використання не одного, а ряду попередніх вимірювань. В [12–16] запропоновано багатокрокові проєкційні алгоритми, які при побудові оцінки на черговому n -му кроці використовували інформацію про L попередні кроки. У роботах [15, 16] визначено швидкість збіжності даних алгоритмів і показано, що використання в багатокрокових проєкційних алгоритмах інформації про L попередні кроки рівносильне в сенсі швидкості збіжності зменшенню розмірності вихідного простору N на L . Істотний позитивний ефект від використання інформації про попередні кроки, з одного боку, і необхідність обертання матриці, з іншого, свідчать про доцільність розробки алгоритмів, що володіють близькими до багатокрокових проєкційних алгоритмів властивостями, але є більш простими в реалізації. Один з таких підходів, що використовує апроксимацію операції ортогонального проєкціювання вектора помилки оці-

нювання на лінійну оболонку вхідних векторів, здійснювану на основі однокрокового проєкційного алгоритму Качмажа, вивчався в роботі [17]. У даній статті розглядається модифікація алгоритму Качмажа, що використовує зважування оцінок, отриманих на ряді попередніх кроків. Досліджено властивості псевдопроєкційних алгоритмів оцінювання, які використовують зважування інформації і побудовані на основі однокрокового адаптивного алгоритму Качмажа. Отримано оцінки швидкості збіжності алгоритмів і показано, що використання операції зважування інформації в алгоритмах оцінювання дозволяє при незначному зниженні швидкості збіжності алгоритмів зменшити розміри їх області збіжності (порівняно з алгоритмом Качмажа з $\gamma = \text{const} = 1$), яка визначається співвідношенням дисперсій сигналів та завад.

Ключові слова: алгоритм Качмажа, проєкційний алгоритм, рекурентна процедура, асимптотична оцінка, точність ідентифікації.

B.D. Liberol, O.G. Rudenko, A.A. Bezsonov

PSEUDOPROJECTION ESTIMATION ALGORITHMS THAT USE INFORMATION WEIGHTING

The Kaczmarz algorithm, which is now widely used due to its simplicity and efficiency in solving various technical problems, was proposed in [3] for solving systems of linear algebraic equations. Analytical studies carried out in [8–10] allowed to obtain non-asymptotic and asymptotic convergence rate estimates of this algorithm in the presence of measurement interference. In [11–14], the problem of acceleration of the Kaczmarz algorithm by using not one but a number of previous measurements was considered. In [12–16], multi-step projection algorithms were proposed, which used information about L of the previous steps when constructing the estimate at the next n -th step. In [15, 16], the speed of convergence of these algorithms was determined and it was shown that the use of information in the multi-step projection algorithms of information about the previous steps is equivalent in the sense of the speed of convergence to reducing the dimension of the output space N by L . A significant positive effect of using information about the previous steps from one side and the need of the matrix inversion on the other side, indicate the feasibility of developing algorithms that have properties similar to multi-step projection algorithms, but which are simpler to implement. One such approach is approximation of the orthogonal projection operation of the estimation error vector onto the input vectors linear shell, carried out on the basis of a one-step Kaczmarz projection algorithm, was studied in [17]. This article discusses a modification of the Kaczmarz algorithm that uses the weighting of the estimates obtained in a number of previous steps. The properties of pseudo-projection estimation algorithms that use information weighing and are based on the one-step adaptive Kaczmarz algorithm are investigated. Estimates of the algorithms convergence rate were obtained and it is shown that the use of the information weighting operation in the estimation algorithms allows to reduce the size of their convergence region (in comparison with the Kaczmarz algorithm with $\gamma = \text{const} = 1$), which is determined by the ratio of signal and interference dispersions, with a slight decrease in the convergence rate.

Keywords: Kaczmarz algorithm, projection algorithm, recursive procedure, asymptotic estimate, identification accuracy.

1. Вазан М. Стохастическая аппроксимация. М.: Мир, 1972. 289 с.
2. Невельсон М.Б., Хасьминский П.З. Стохастическая аппроксимация и рекуррентное оценивание. М.: Наука, 1972. 304 с.
3. Kaczmarz S. Angenäherle Auflösung von Systemen linearer Gleichungen Bull. Int. Acad. Polon. Sci. Lett., C 1, Sci. Math. Nat., Ser. A. 1937. P. 355–357.

4. Nagumo I., Noda A. A learning method for system identification. *IEEE Trans. Autom. Control*. 1967. AC-12. N 3. P. 282–287.
5. Чадеев В.М. Определение динамических характеристик объектов в процессе их нормальной эксплуатации для целей самонастройки. *Автоматика и телемеханика*. 1964 **25**, № 9. С. 1302–1306.
6. Райбман Н.С., Чадеев В.М. Адаптивные модели в системах управления. М.: Сов. радио, 1966. 156 с.
7. Aved'jan A. D. Bestimmung der Parameter linearer Modelle stationärer und instationärer Strecken. *Messen, steuern, regeln*. 1971. N 9, P. 348–350.
8. 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.
9. Либероль Б.Д., Руденко О.Г., Бессонов А.А. Исследование сходимости одношаговых адаптивных алгоритмов идентификации. *Международный научно-технический журнал «Проблемы управления и информатики»*. 2018. **5**. С. 19–32.
10. Rudenko O., Bezsonov O., Romanyk O., Lebediev V. Analysis of convergence of adaptive single-step algorithms for the identification of non-stationary objects. *Eastern-European Journal of Enterprise Technologies*. 2019. N 1/4 (97). P. 6–14.
11. Либероль Б.Д., Руденко О.Г., Тимофеев В.А. Модифицированный алгоритм Качмажа для оценивания параметров нестационарных объектов. *Международный научно-технический журнал «Проблемы управления и информатики»*. 1995. № 4. С. 81–89.
12. Лелашвили Ш. Г. Применение одного итерационного метода для анализа многомерных автоматических систем. *Схемы автоматического управления*. 1965. С. 19–33.
13. Лелашвили Ш. Г. Некоторые вопросы построения статистической модели многомерных объектов. *Автоматическое управление*. 1967. С. 59–96.
14. Аведьян Э.Д. Модифицированные алгоритмы Качмажа для оценки параметров линейных объектов. *Автоматика и телемеханика*. 1978. **5**. С. 64–72
15. Ищенко Л.А., Либероль Б.Д., Руденко О.Г. О свойствах одного класса многошаговых адаптивных алгоритмов идентификации. *Кибернетика*, 1986. № 1. С. 92–96.
16. Либероль Б.Д., Руденко О.Г. О свойствах проекционных алгоритмов оценивания параметров нестационарных объектов. *Доклады АН УССР*, сер. А. 1990. № 4. С. 71–74.
17. Либероль Б.Д., Руденко О.Г., Бессонов А.А. Псевдопроеционные алгоритмы оценивания, основанные на аппроксимации операции ортогонального проецирования. *Международный научно-технический журнал «Проблемы управления и информатики»*. 2020. № 2. С. 16–33.
18. Ланкастер П. Теория матриц. М.: Наука, 1978. 280 с.
19. Поляк Б.Т Сравнение скорости сходимости одношаговых и многошаговых алгоритмов оптимизации при наличии помех измерений. *Техническая кибернетика*.1974. № 5. С. 9–12.

Получено 20.02.2020