УДК 004.852

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

ПСЕВДОПРОЕКЦИОННЫЕ АЛГОРИТМЫ ОЦЕНИВАНИЯ, ОСНОВАННЫЕ НА АППРОКСИМАЦИИ ОПЕРАЦИИ ОРТОГОНАЛЬНОГО ПРОЕЦИРОВАНИЯ

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

Введение

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

$$y_n = \mathbf{c}^{*\mathrm{T}} \mathbf{x}_n + \xi_n, \tag{1}$$

где y_n — наблюдаемый выходной сигнал; $\mathbf{x}_n = (x_{1n}, x_{2n}, \ldots, x_{Nn})^{\mathrm{T}}$ — вектор входных сигналов $N \times 1$; $\mathbf{c}^* = (c_1{}^*, c_2{}^*, \ldots, c_N{}^*)^{\mathrm{T}}$ — вектор искомых (оцениваемых) параметров $N \times 1$; ξ_n — помеха; Т — символ транспонирования; n — дискретное время, сводится к минимизации некоторого наперед выбранного функционала качества (критерия идентификации) от ошибки оценивания. Использование квадратичного функционала приводит к различным алгоритмам идентификации, позволяющим получить оценки с искомого вектора c^* при нормальных распределениях помехи ξ_n .

Одним из наиболее эффективных и наиболее простых в вычислительном отношении одношаговых алгоритмов оценивания является алгоритм Качмажа, предложенный в работе [1] и имеющий вид

$$\mathbf{c}_n = \mathbf{c}_{n-1} + \frac{y_n - \mathbf{c}_{n-1}^{\mathrm{T}} \mathbf{x}_n}{\|\mathbf{x}_n\|^2} \mathbf{x}_n, \tag{2}$$

где $\|\bullet\|$ — евклидова норма.

Оценки скорости сходимости данного алгоритма при идентификации стационарных объектов впервые получены в [2–6]. Если в [2–4] рассматривался регулярный случай, то в [5, 6] были получены оценки, учитывающие статистические свойства сигналов и помех.

Для повышения вычислительной устойчивости алгоритма (2) В.М. Чадеев [2, 3] предложил его модификацию — регуляризованный алгоритм, использующий в знаменателе помимо $\|\mathbf{x}_n\|^2$ некоторое положительное число, параметр регуляризации. В работах [5, 6] получены неасимптотические и асимптотические оценки скорости сходимости регуляризованного алгоритма Качмажа, которые показали, что введение регуляризующей добавки, улучшая устойчивость алгоритма, приводит к замедлению скорости его сходимости.

Возможность ускорения алгоритма Качмажа путем использования не одного, а ряда измерений рассмотрена в [7]. В настоящей публикации исследовалась модификация алгоритма Качмажа, достаточно эффективна при оценивании нестационарных параметров.

Публикации Ш.Г. Лелашвили [8, 9] послужили толчком к разработке нового класса алгоритмов — многошаговых проекционных алгоритмов [10–14], в которых при построении оценки на n-м шаге используется не только вновь поступающая информация, как это происходит в алгоритме Качмажа, но и информация о ряде предыдущих шагов $n - 1$, $n - 2$… Количество таких шагов определяет память алгоритма. При этом благодаря лучшей экстраполяции и фильтрации в ряде случаев удается добиться существенного сокращения времени идентификации. Преимущества такого подхода отмечались в [15], а в [16] проведен сравнительный анализ одно- и двухшаговых алгоритмов. Общая форма многошаговых алгоритмов изучалась в [8–14]. В этих работах рассматривался проекционный алгоритм вида

$$\mathbf{c}_n = X_n^{s\mathrm{T}}(X_n^s X_n^{s\mathrm{T}})^{-\mathbf{1}} Y_n^S, \tag{3}$$

где $Y_n^S = (y_n, y_{n-1}, \ldots, y_{n-s+1})^{\mathrm{T}}$ — вектор $S \times 1$; $X_n^s = (\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-S+1})^{\mathrm{T}}$ — матрица $S \times N$.

В [11–14] установлены свойства случайных псевдообратных матриц и матриц проецирования, которые позволили определить скорость сходимости данных алгоритмов и сделать вывод, что учет в данных алгоритмах информации об S предыдущих шагах равносилен в смысле скорости сходимости уменьшению размерности исходного пространства N на S. Таким образом, применение многошаговых проекционных алгоритмов позволяет существенно ускорить процесс идентификации и является достаточно эффективным при оценивании нестационарных параметров [12, 14].

Как видно из (3), реализация алгоритма требует вычисления обратной матрицы наблюдений $(X_n^s X_n^{s\mathrm{T}})^{-\mathbf{1}}$ размерности $S \times S$.

Радикальным способом уменьшения требуемого объема памяти и упрощения многошаговых алгоритмов является использование рекуррентных соотношений для вычисления оценок (рекуррентная форма данных алгоритмов предложена в [10]). Однако, как показывает практика, хотя рекуррентные алгоритмы устраняют необходимость операции обращения матриц, что значительно упрощает вычисления и снижает чувствительность к ошибкам округления, для их реализации требуется значительное количество арифметических операций. В связи с этим представляет интерес разработка алгоритмов, обладающих близкими к проекционным алгоритмам свойствами, но более простых в реализации. Так как наиболее простым, но достаточно эффективным одношаговым проекционным алгоритмом является алгоритм Качмажа (2), целесообразно рассмотреть процедуры, использующие многократное применение этого алгоритма.

Рассматриваемые в данной статье методы осуществляют аппроксимацию операции ортогонального проецирования вектора ошибки $\theta_{n-1} = \mathbf{c}^* - \mathbf{c}_{n-1}$ на $L_n^{(S)}$ — линейную оболочку S входных векторов $\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-S+1}$.

Процедуры последовательного проецирования

При построении процедур оценивания используются известные свойства матриц проецирования, сформулированные в леммах 1 и 2.

Лемма 1. Матрицы проецирования симметричны $(P = P^{\mathrm{T}})$ и идемпотентны $(P^2 = P)$.

Лемма 2. Пусть P_Ω — оператор проецирования на линейное подпространство Ω. Тогда, если подпространство $\omega \subset \Omega$, то

$$P_\Omega P_\omega = P_\omega,\ \ P_\Omega = P_\omega + P_{\Omega\perp\omega}, \tag{4}$$

где $\Omega \perp \omega$ — ортогональное дополнение к ω в Ω.

Доказательства лемм приводятся в [17].

Для удобства применения используемых свойств матриц, входящих в алгоритмы оценивания, будем сразу формулировать их для случая, когда эти матрицы строятся по S последовательным векторам входного сигнала $\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-S+1}$, хотя все формулы справедливы и в общем случае. Так, лемма 2 в применении к подпространствам $\Omega = L_{n|s}$ (линейная оболочка s векторов $\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-S+1}$) и $\omega = L_{n-i|s-i}$ (линейная оболочка *s*–*i* векторов $\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-S+1}$) дает следующее свойство операторов проецирования на вложенные линейные оболочки $L_{n-i|s-i} \subset L_{n|s}$:

$$\begin{aligned} P_n^S P_{n-i}^{S-i} = P_{n-i}^{S-i} P_n^S &= P_{n-i}^{S-i}\ , \\ P_n^S = P_{n-1}^{S-1} &+ P_{\mathbf{x}_n \perp L_{n-1}^{S-1}}, \end{aligned} \tag{5}$$

где $\mathbf{x}_n \perp L_{n-1}^{S-1}$ — компонента вектора $\mathbf{x}_n$, ортогональная подпространству L_{n-1}^{S-1}).

В этих методах проецирование θ_{n-1} на $L_n^{(S)}$ производится в несколько этапов, шаг алгоритма разбивается на несколько подшагов. На *n*-й итерации после поступления нового вектора $\mathbf{x}_n$ производится первый подшаг по стандартному алгоритму Качмажа (2), т.е. θ_{n-1} ортогонально проецируется на $\mathbf{x}_n$ (полученную проекцию обозначим θ'_n):

$$\theta'_n = (I - P_n^{(1)})\theta_{n-1}. \tag{6}$$

Матрица проецирования на вектор $\mathbf{x}_n$ имеет вид

$$P_n^{(1)} = \frac{\mathbf{x}_n \mathbf{x}_n^{\mathrm{T}}}{\|\mathbf{x}_n\|^2}.$$

Обозначим $\theta_{n-1\|}(L_n^{(S)})$ ортогональную проекцию вектора θ_{n-1} на $L_n^{(S)}$.

При точном проецировании — $\theta_{n\|}(L_n^{(S)}) = 0$, при псевдопроецировании — $\theta_{n\|}(L_n^{(S)}) \neq 0.$

Задача построения метода состоит в поиске алгоритма, уменьшающего величину проекции θ'_n на $L_n^{(S)}$. В этом случае скорость сходимости алгоритма будет приближаться к скорости сходимости алгоритма с точным проецированием. Как показано ниже, для получения скорости сходимости, близкой к скорости сходимости алгоритма с точным проецированием, достаточно добиться уменьшения $\theta'_{n\|}(L_n^{(S)})$ в λ раз. Точному проецированию отвечает $\lambda = \infty$, но уже при $\lambda \sim 2-3$ для $N-S>1$ достигается скорость сходимости, близкая к алгоритму с точным проецированием. Для получения таких λ предлагаются следующие процедуры.

Процедура 1. Осуществляется последовательное проецирование θ'_n на предшествующие векторы: $\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-S+1}$.

Процедура 2. Осуществляется проецирование θ'_n на некоторый вектор z_n, направление которого коррелирует с направлением проекции θ'_n на $L_n^{(S)}$. В качестве такого вектора может быть взят, например, вектор вида

$$z_n = \sum_{i=1}^{S-1} P_{n-i}^{(1)} \theta_n.$$

Процедура 3. При $S < 0{,}5N$ производится повторное проецирование на вектор $\mathbf{x}_{n-S+1}$.

Исследование вопросов сходимости процедур

Проведем сначала оценку скорости сходимости алгоритмов как функцию параметра λ, характеризующего убывание проекции вектора θ'_n на $L_n^{(S)}$. Для вычисления скорости сходимости рассмотрим

$$\Psi_n = M_{x_n / x_{n-1}, \ldots} \{ \|\theta_{n-1}\|^2 - \|\theta_n\|^2 \}. \tag{7}$$

Утверждение 1. Для рассматриваемых псевдопроекционных процедур величина $\psi_n = \alpha_{n-1} - \alpha_n$ определяется соотношением

$$\psi_n = \frac{1}{N-s+1} \left\{ 1 - \frac{1}{N-s+1} \left(\frac{S-1}{N} \right)^2 \frac{1}{1-\lambda^{-1}} \right\} \alpha_{n-1}, \tag{8}$$

причем для процедуры 1 $\lambda \approx \mathrm{e}$, а для процедуры 2 $\lambda = 2$.

Здесь $\alpha_i = M\{\| \Theta_i \|^2\}$; $\theta_i = c^* - c_i$ — ошибка оценивания.

Доказательство. Учитывая подшаг $\theta_{n-1} \to \theta'_n$ по алгоритму Качмажа, запишем

$$\|\theta_{n-1}\|^2 - \|\theta_n\|^2 = \|\theta_{n-1}\|^2 - \|\theta'_n\|^2 + \|\theta'_n\|^2 - \|\theta_n\|^2.$$

Известно, что [2, 3, 6]

$$M_{x_n / x_{n-1}, \ldots} \{ \|\theta_{n-1}\|^2 - \|\theta'_n\|^2 \} = \frac{1}{N} \|\theta_{n-1}\|^2. \tag{9}$$

Векторы θ'_n и θ_n имеют разные проекции на $L_n^{(S)}$, но их проекции на $L_n^{(S)\perp}$ — ортогональное дополнение к $L_n^{(S)}$ — одинаковы, так как проецирование на направление, лежащее в $L_n^{(S)}$, не может изменить величину проекции на $L_n^{(S)\perp}$. Это утверждение справедливо и в случае, когда θ_n получается из θ'_n проецированием на $L_{n-1}^{(S-1)}$. Поэтому можем записать $\|\theta'_n\|^2 - \|\theta_n\|^2 = \left\|\theta'_n (L_{n-1}^{(S-1)})\right\|^2 - \left\|\theta_{n\|} (L_{n-1}^{(S-1)})\right\|^2$.

Введем параметр λ соотношением

$$M \left\{ \left\| \theta_{n\|} \left(L_{n-1}^{(S-1)} \right) \right\|^2 \right\} = \frac{1}{\lambda} M \left\{ \left\| \theta'_{n\|} \left(L_{n-1}^{(S-1)} \right) \right\|^2 \right\}, \tag{10}$$

отражающим тот факт, что рассматриваемые модифицированные алгоритмы уменьшают проекцию вектора θ'_n на $L_{n-1}^{(S-1)}$ в среднем в λ раз. Перепишем (9)

$$\Psi_n = M_{x_n / x_{n-1}, \ldots}\{\|\theta_{n-1}\|^2 - \|\theta_n\|^2\} =$$

$$= \frac{1}{N} M_{x_n / x_{n-1}, \ldots}\{\|\theta_{n-1}\|^2\} + (1 - \frac{1}{\lambda}) M_{x_n / x_{n-1}, \ldots} M\left\{ \left\|\theta'_{n\|}\left(L_{n-1}^{(S-1)}\right)\right\|^2 \right\}.$$

Для анализа сходимости процедур потребуются некоторые результаты, устанавливаемые следующей леммой.

Лемма 3. Пусть в пространстве E_N будет ортонормированный базис $e_1, e_2, \ldots, e_N$, первые $S-1$ векторов которого образуют линейное подпространство, натянутое на $S-1$ вектор $\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-S+2}$. Тогда при принятых предположениях относительно статистических свойств сигналов имеет место формула

$$M_e\{P_e P_{n-1}^{(S-1)} P_e\} = \frac{S-1}{N(N+2)} I + \frac{2}{N(N+2)} (I - P_{n-1}^{(S-1)}), \tag{11}$$

где

$$M_{\mathrm{x}_n / \mathrm{x}_{n-1}, \ldots}\{\bullet\} = M_e\{\bullet\}; \;\; \mathrm{e} = \mathrm{x}_n \|\mathrm{x}_n\|^{-1}.$$

Доказательство. Рассмотрим операторы проецирования на e и $L_{n-1}^{(S-1)}$:

$$P_e = P_n^{(1)}, \;\; P_e = ee^{\mathrm{T}}, \;\; P_{n-1}^{(S-1)} = e_1 e_1^{\mathrm{T}} + \ldots + e_{S-1} e_{S-1}^{\mathrm{T}}.$$

В базисе $e_1, e_2, \ldots, e_{S-1}$ оператор $P_e P_{n-1}^{(S-1)} P_e$ запишем следующим образом:

$$P_e P_{n-1}^{(S-1)} P_e = ee^{\mathrm{T}} (e_1 e_1^{\mathrm{T}} + \ldots + e_{S-1} e_{S-1}^{\mathrm{T}}) ee^{\mathrm{T}},$$

где $e_1^{\mathrm{T}} = (1, 0, \ldots, 0)$, $e_2^{\mathrm{T}} = (0, 1, \ldots, 0), \ldots, e_{S-1}^{\mathrm{T}} = (0, 0, \ldots, 1)$.

Учитывая, что $e^{\mathrm{T}} e_i = e_i^{\mathrm{T}} e = e^{(i)}$, $e^{(i)}$ — i-я компонента вектора e, имеем

$$e^{\mathrm{T}} e_i e_i^{\mathrm{T}} e = (e^{(i)})^2,$$

$$P_e P_{n-1}^{(S-1)} P_e = e(e^{(1)2} + e^{(2)2} + \ldots + e^{(S-1)2}) e^{\mathrm{T}} =$$
$$= (e^{(1)2} + e^{(2)2} + \ldots + e^{(S-1)2}) ee^{\mathrm{T}}.$$

Выпишем компоненты этой матрицы в выбранном базисе

$$P_e P_{n-1}^{(S-1)} P_e = \begin{pmatrix} (e^{(1)2} + \ldots + e^{(S-1)2}) e^{(1)2} \ldots (e^{(1)2} + \ldots + e^{(S-1)2}) e^{(1)} e^{(N)} \\ \ldots \\ (e^{(1)2} + \ldots + e^{(S-1)2}) e^{(N)} e^{(1)} \ldots (e^{(1)2} + \ldots + e^{(S-1)2}) e^{(N)2} \end{pmatrix}.$$

Вычислим математическое ожидание обеих частей данной матрицы, которое сводится к усреднению правой части по вектору e. Функция распределения этого случайного вектора обладает угловой симметрией вследствие того, что компоненты $\mathbf{x}_n$ имеют нулевое математическое ожидание и одинаковую дисперсию. Поэтому недиагональные члены матрицы, имеющие вид $(e^{(1)2} + \ldots + e^{(S-1)2}) e^{(i)} e^{(j)}$, $i \neq j$, при усреднении дают 0. При усреднении диагональных элементов встречаются выражения двух типов:

$$M_{22} = M_e\{e^{(i)^2} e^{(j)^2}\} \text{ и } M_4 = M_e\{e^{(i)^4}\},$$

 ISSN 0572-2691

причем при $i = j < S-1$ присутствуют члены первого и второго вида, а при $i = j > S-1$ — только первого.

Вычислим M_{22}. С учетом явного вида функции распределения вектора $\mathbf{x}$, которая в силу независимости компонент — просто произведение N гауссовских функций с одинаковыми параметрами, имеем

$$f(x) = \sqrt{\frac{\omega}{\pi}} e^{-\omega \|\mathbf{x}\|^2}, \quad \omega = (2\sigma_x^2)^{-1},$$

для M_{22} получаем

$$M_e\{e^{(i)^2} e^{(j)^2}\} = M_x \left\{ \frac{x^{(i)^2} x^{(j)^2}}{\|\mathbf{x}\|^4} \right\} = \left(\sqrt{\frac{\omega}{\pi}} \right)^N \int\limits_{-\infty} \dots \int\limits_{-\infty} e^{-\omega \|\mathbf{x}\|^2} \frac{x^{(i)^2} x^{(j)^2}}{\|\mathbf{x}\|^4} dx^{(1)} \dots dx^{(N)}.$$

Введем функцию от параметра

$$J(\omega) = \int\limits_{-\infty} e^{-\omega \|\mathbf{x}\|^2} \frac{x^{(i)^2} x^{(j)^2}}{\|\mathbf{x}\|^4} d\mathbf{x}.$$

Воспользовавшись известной формулой дифференцирования по параметру, продифференцируем $J(\omega)$ дважды по ω:

$$\frac{\partial^2 J}{\partial \omega^2} = \int\limits_{-\infty} \dots \int\limits_{-\infty} e^{-\omega (x^{(1)2} + \dots + x^{(N)2})} x^{(i)^2} x^{(j)^2} dx^{(1)} \dots dx^{(N)}.$$

N-мерный интеграл распадается на произведение N однократных интегралов. Интегрирование по $N-2$ переменным (кроме i- и j-й) дает просто нормировочный множитель $\sqrt{\pi\omega^{-1}}$. Интегрирование по i- и j-й переменным дает дисперсию, умноженную на нормировочный множитель:

$$\int\limits_{-\infty} e^{-\omega x^{(i)2}} x^{(i)2} dx^{(i)} = \frac{1}{2\omega} \sqrt{\frac{\pi}{\omega}}.$$

С учетом этого получаем

$$\frac{\partial^2 J}{\partial \omega^2} = \frac{1}{4\omega^2} \left(\sqrt{\frac{\pi}{\omega}} \right)^N.$$

Интегрируя это дифференциальное уравнение дважды по ω, с учетом очевидного граничного условия $J \to 0$ при $\omega \to \infty$

$$J(\omega) = \left(\frac{\pi}{\omega} \right)^N \frac{1}{N(N+2)}.$$

Таким образом, для M_{22} имеем

$$M_e \left\{ \frac{x^{(i)^2} x^{(j)^2}}{\|\mathbf{x}\|^4} \right\} = \frac{1}{N(N+2)}.$$

M_4 вычисляем аналогичным способом:

$$M_{\mathbf{x}} \left\{ \frac{x^{(i)^4}}{\|\mathbf{x}\|^4} \right\} = \left(\sqrt{\frac{\omega}{\pi}} \right)^N \int e^{-\omega \|\mathbf{x}\|^2} \frac{x^{(i)^4}}{\|\mathbf{x}\|^4} d\mathbf{x};$$

$$J(\omega)=\int e^{-\omega\|\mathbf{x}\|^2}\frac{x^{(i)4}}{\|\mathbf{x}\|^4}d\mathbf{x};$$

$$\frac{\partial^2 J}{\partial\omega^2}=\int e^{-\omega\|\mathbf{x}\|^2}x^{(i)4}d\mathbf{x}=\left(\sqrt{\frac{\pi}{\omega}}\right)^{N-1}\int x^{(i)4}e^{-\omega x^{(i)}2}dx^{(i)}.$$

Обозначая

$$J_0=\int e^{-\omega x^{(i)2}}dx=\sqrt{\frac{\pi}{\omega}},$$

получаем

$$\int x^{(i)4}e^{-\omega x^{(i)}2}dx=\frac{\partial^2 J_0}{\partial\omega^2}=\frac{3}{4\omega^2}\omega^{-0.5};\frac{\partial^2 J_0}{\partial\omega^2}=\frac{3}{4\omega^2}\left(\frac{\pi}{\omega}\right)^N.$$

Это отличается от полученного на предыдущей странице выражения для $\frac{\partial^2 J}{\partial\omega^2}$ только множителем 3. Окончательно выражение для M_4 принимает вид

$$M_4=M_{\mathbf{x}}\left\{\frac{x^{(i)4}}{\|\mathbf{x}\|^4}\right\}=\frac{3}{N(N+2)}.$$

Для искомого математического ожидания оператора $P_eP_{n-1}^{(S-1)}P_e$ в выбранном базисе имеем

$$M_e\{P_eP_{n-1}^{(S-1)}P_e\}=$$

$$=\left(\begin{array}{ccc|ccc}
\frac{S-1}{N(N-2)}+\frac{2}{N(N-2)} & \cdots & 0 & & & \\
\vdots & & & & 0 & \\
0 & \cdots & \frac{S-1}{N(N-2)}+\frac{2}{N(N-2)} & & & \\
\hline
 & & & \frac{S-1}{N(N-2)} & \cdots & 0 \\
 & 0 & & \vdots & & \vdots \\
 & & & 0 & \cdots & \frac{S-1}{N(N-2)}
\end{array}\right).$$

Учитывая вид матрицы $P_{n-1}^{(S-1)}$ в этом базисе, получаем окончательно выражение (11). Так как по определению

$$\theta'_{n\|}(L_{n-1}^{(S-1)})=P_{n-1}^{(S-1)}(I-P_e)\theta_{n-1},$$

то имеем

$$\left\|\theta'_{n\|}(L_{n-1}^{(S-1)})\right\|^2=\theta_{n-1}^T(I-P_e)P_{n-1}^{(S-1)}P_{n-1}^{(S-1)}(I-P_e)\theta_{n-1}$$

и с учетом (4),(5), т.е. того, что $(P_{n-1}^{(S-1)})^2=P_{n-1}^{(S-1)}$, можем записать

$$\left\|\theta'_{n\|}(L_{n-1}^{(S-1)})\right\|^2=\theta_{n-1}^{\mathrm{T}}P_{n-1}^{(S-1)}\theta_{n-1}-2\theta_{n-1}^{\mathrm{T}}P_eP_{n-1}^{(S-1)}\theta_{n-1}+\theta_{n-1}^{\mathrm{T}}P_eP_{n-1}^{(S-1)}P_e\theta_{n-1}. \quad (12)$$

 ISSN 0572-2691

Тогда выражение для $M\left\{\left\|\theta'_{n\|}\left(L_{n-1}^{(S-1)}\right)\right\|^2\right\}$ примет вид

$$M\left\{\left\|\theta'_{n\|}\left(L_{n-1}^{(S-1)}\right)\right\|^2\right\}=M\{\|\theta_{n-1}\|^2\}\left(1-\frac{2}{N}+\frac{2}{N(N+2)}\right)+\frac{S-1}{N(N+2)}M\{\|\theta_{n-1}\|^2\}. \quad (13)$$

Введение параметра τ соотношением

$$M\left\{\left\|\theta'_{n\|}\left(L_{n-1}^{(S-1)}\right)\right\|^2\right\}=\tau M\{\|\theta_{n-1}\|^2\} \quad (14)$$

позволяет записать формулу (13) в виде

$$M_e\left\{\left\|\theta'_{n\|}\right\|^2\right\}=\left[\tau\left(1-\frac{2}{N}+\frac{2}{N(N+2)}\right)-\frac{S-1}{N(N+2)}\right]M\left\{\|\theta_{n-1}\|^2\right\}.$$

А для интересующей нас величины Ψ с точностью до $o(N^{-2})$ получаем

$$\Psi=\left\{\frac{1}{N}+\left(1-\frac{1}{\lambda}\right)\left[\tau\left(1-\frac{2}{N}+\frac{2}{N^2}\right)+\frac{s-1}{N^2}\right]\right\}M\{\|\theta_{n-1}\|^2\}. \quad (15)$$

Для определения входящего в (15) параметра τ поступим следующим образом: вычислим величину проекции

$$\left\|\theta'_{n\|}(L_{n-1}^{(S-1)})\right\|^2=\theta_n^T P_n^{(S)}\theta_n.$$

Воспользуемся снова свойством оператора проецирования на вложенные линейные оболочки, которыми в нашем случае будут $P_n^{(S)}$ и $P_{n-1}^{(S)}$: $P_n^{(S)}=P_{n-1}^{(S-1)}+P_{e_\perp}$. Здесь введено обозначение $e_\perp$ для вектора $(I-P_{n-1}^{(S-1)})e$. Будем считать, что

$$\theta_{n\|}(L_{n-1}^{(S-1)})\approx\theta_{n\|}(L_n^{(S)}).$$

Это предположение оправдывается анализом последующих результатов и выполняется с точностью до величин порядка $o(S^{-1})$. Тогда можно записать

$$\begin{aligned}M\left\{\left\|\theta'_{n\|}\left(L_{n-1}^{(S-1)}\right)\right\|^2\right\}&=M\left\{\left\|P_{n-1}^{(S-1)}\theta_n\right\|^2\right\}-M\left\{\left\|P_{e_\perp P_{n-1}^{(S-1)}}\theta_n\right\|^2\right\}=\\&=\frac{1}{\lambda}\left\{\left\|\theta'_{n\|}\left(L_{n-1}^{(S-1)}\right)\right\|^2\right\}+M\left\{\left\|P_{e_\perp}\theta'_{n\|}\right\|^2\right\}\end{aligned} \quad (16)$$

по определению параметра λ, введенного как показатель уменьшения квадрата длины проекции вектора θ'_n на $L_{n-1}^{(S-1)}$.

Учитывая, что

$$\theta'_{n|}=(I-P_e)\theta_{n-1},\ \ e_\perp=(I-P_{n-1}^{(S-1)})e,$$

получаем

$$\left\|P_{e_\perp L_{n-1}^{(S-1)}}\theta'_n\right\|^2=\theta_{n-1}^T(I-P_e)P_{e_\perp L_{n-1}^{(S-1)}}(I-P_e)\theta_{n-1}.$$

Проведем усреднение по вектору e, выбирая базис, в котором, как и ранее, оператор проецирования $P_{n-1}^{(S-1)}$ имеет канонический вид.

Для $e_\perp$ в этом базисе имеем

$$e_\perp = (\underbrace{o,\ldots,o,}_{S-1} e_S,\ldots,e_N).$$

Поскольку оператор проецирования на вектор $\mathbf{x}$ есть $\mathbf{x}\mathbf{x}^{\mathrm{T}}\|\mathbf{x}\|^{-2}$, можем записать соотношение

$$P_e P_{e_\perp} = (P_{e_\perp} P_e)^{\mathrm{T}} = \frac{e(e^{\mathrm{T}} e_\perp) e_\perp^{\mathrm{T}}}{\|e_\perp\|^2} = e e_\perp^{\mathrm{T}}.$$

Для усреднения произведений операторов проецирования, входящих в формулу (15), воспользуемся леммой 3. При этом

$$M\{P_e P_{e_\perp} P_e\} = \frac{S-1}{N(N+2)} I + \frac{2}{N(N+2)}(I - P_{n-1}^{(S-1)});$$

$$M\{P_{e_\perp}\} = \frac{1}{S-1}(I - P_{n-1}^{(S-1)});$$

$$M\{P_e P_{e_\perp}\} = M\{e e_\perp^{\mathrm{T}}\} = \frac{1}{N}(I - P_{n-1}^{(S-1)}).$$

Окончательно, с учетом того, что введение параметра τ позволяет записать

$$M\{\theta_{n-1}^{\mathrm{T}} P_{n-1}^{(S-1)} \theta_{n-1}\} = \tau M\{\|\theta_{n-1}\|^2\},$$

для интересующей нас величины $M_e\{\|P_{e_\perp}\theta'_n\|^2\}$ получаем

$$M_e\{\|P_{e_\perp}\theta'_n\|^2\} = (1-\tau)\left(\frac{1}{S-1} - \frac{2}{N} + \frac{2}{N^2}\right)\|\theta_{n-1}\|^2 + \frac{S-1}{N^2}.$$

Подставляя полученный результат в соотношение (15), выпишем уравнение для τ, дающее вместе с (14) систему, при решении которого получаем оценку скорости сходимости рассматриваемого алгоритма:

$$\tau(1-\psi)(1+\delta) = \frac{1}{\lambda}\left[\tau\left(1 - \frac{2}{N} + \frac{2}{N^2}\right) - \frac{S-1}{2}\right] + (1-\tau)\left(\frac{1}{S-1} - \frac{2}{N} + \frac{2}{N^2}\right);$$

$$\psi = \left\{\frac{1}{N} + \left(1 - \frac{1}{\lambda}\right)\left[\tau\left(1 - \frac{2}{N} + \frac{2}{N(N+2)}\right)\right]\right\} M\{\|\theta_{n-1}\|^2\}.$$

Решение этой системы с точностью до величин более высокого порядка малости дает для ψ выражение

$$\psi = \frac{1}{S-1}\left[1 - \frac{2}{S-1}\left(\frac{S-1}{2}\right)^2 \frac{1}{1-\lambda^{-1}}\right] M\{\|\theta_{n-1}\|^2\}. \tag{17}$$

Анализ этой формулы показывает, что уже при $\lambda = 2$ и $N - S > 1$ скорость сходимости упрощенного проекционного алгоритма близка к скорости сходимости алгоритма с точным проецированием, что подтверждается численным моделированием. При $\lambda \to \infty$ из (17) видно, что при $N - S > 1$ скорость сходимости,

 ISSN 0572-2691

вычисленная по рассмотренной методике, очень близка к скорости сходимости S-шагового проекционного алгоритма [11, 13], что оправдывает предположения, поскольку $\lambda \to \infty$ соответствует полному аннулированию проекции θ'_n на $L_{n-1}^{(S-1)}$, и мы приходим к алгоритму с точным проецированием.

Отметим, что хотя формально условие $N - S > 1$ не учитывалось, сделанные в ходе вычислений предположения при $N - S \sim 1$ будут несправедливыми, эффективное значение λ приближается к единице и формула (17) теряет смысл.

Отметим теперь, какие значения λ достигаются в предложенных вариантах алгоритмов. Шаг процедуры 1 состоит из S подшагов: нулевого по алгоритму Качмажа и последующих $(S-1)$-подшагов повторного проецирования на векторы $\mathbf{x}_{n-1}, \mathbf{x}_{n-2}, \ldots, \mathbf{x}_{n-S+1}$. Можно показать, что при этом убывание $\left\|\theta_{n\|}\right\|^2$ в $(S-1)$-мерном подпространстве $L_{n-1}^{(S-1)}$ происходит быстрее, чем убывает величина $\left\|\theta_n\right\|^2$ в пространстве той же размерности за S шагов по алгоритму Качмажа, потому что величина проекции $\left\|\theta_{n\|}\right\|^2$ на текущий вектор $\mathbf{x}_n$ оказывается в среднем больше, чем на другие векторы, на которые последнее проецирование производилось позднее.

Таким образом, для процедуры 1 получаем оценку

$$\lambda \approx \left(1 - \frac{1}{S-1}\right)^{S-1}, \ \lambda \sim e.$$

В процедуре 2 после нулевого подшага по алгоритму Качмажа по $\mathbf{x}_n$ осуществляется подшаг в направлении вектора

$$\mathbf{z}_n = \sum_{i=1}^{S} (y_{n-i} - c_{n-1}^{\mathrm{T}} \mathbf{x}_{n-i}) \mathbf{x}_{n-i}$$

в соответствии с

$$\theta_n = \theta'_n - \frac{{\theta'_n}^{\mathrm{T}} \mathbf{z}_n}{\left\|\mathbf{z}_n\right\|^2} \mathbf{z}_n.$$

Оценим, как будет уменьшаться величина проекции θ'_n на $L_{n-1}^{(S-1)}$, что и даст нам оценку параметра λ, поскольку он определяется из (10) как степень этого уменьшения. Вектор $\mathbf{z}_n$ лежит в $L_{n-2}^{(S-1)}$, его направление оказывается коррелированным с направлением проекции θ'_{n-1} на $L_{n-2}^{(S-1)}$, обозначаемой $\theta'_{n-1\|}$. По определению $P_{n-2}^{(S-1)}$ и $\theta_{n-1\|}$ с учетом подшага по алгоритму Качмажа имеем

$$\theta_{n-1\|} = P_{n-1}^{(S-1)} \theta_{n-1}; \ \theta'_{n\|} = P_{n-1}^{(S-1)} \theta_{n-1\|}; \ \theta_{n\|} = P_{n-1}^{(S-1)} (I - P_{\mathbf{z}_n}) \theta'_n.$$

Поскольку $\mathbf{z}_n \in L_{n-1}^{(S-1)}$, по свойству проекционных операторов (4)

$$P_{n-1}^{(S-1)} P_{\mathbf{z}_n} = P_{\mathbf{z}_n}; \ \theta_{n\|} = \theta'_{n\|} - P_{\mathbf{z}_n} \theta'_{n\|}.$$

Отсюда

$$\left\|\theta_{n\|}\right\|^2 = \left\|\theta'_{n\|}\right\|^2 - {\theta'_n}^{\mathrm{T}} P_{\mathbf{z}_n} \theta'_{n\|}$$

с учетом явного вида оператора проецирования на вектор $\mathbf{z}_n$

$$M\{\|\theta_n\|^2\} = M\{\|\theta'_{n\|}\|^2\} - M\left\{\frac{(\theta'^{\mathrm{T}}_{n\|}\mathbf{z}_n)^2}{\|\mathbf{z}_n\|^2}\right\}.$$

Будем характеризовать уменьшение $M\{\|\theta'_{n\|}\|^2\}$ величиной $M\{\psi_n\}$, $\psi_n = \|\theta'_{n\|}\|^2 - \|\theta_{n\|}\|^2$. Оценим эту величину, считая, что $\theta'_{n\|}$ независимы от $\mathbf{x}_{n-1}, \mathbf{x}_{n-2}, \ldots, \mathbf{x}_{n-S+1}$. Это соответствует тому, что проекции θ'_n на векторы $\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{n-S+1}$ имеют в среднем одинаковую величину, потому что основной вклад в эти проекции дает подшаг $\theta_n \to \theta'_n$ $\|\bullet\|$ по вектору $\mathbf{x}_n$, независимому от векторов $\mathbf{x}_{n-1}, \mathbf{x}_{n-2}, \ldots, \mathbf{x}_{n-S+1}$.

При вычислении $M\{\psi\}$ для упрощения записи примем обозначения

$$\theta = \theta'_{n\|}(L^{(S-1)}_{n-1});\ \ \mathbf{x}_{n-i} = \mathrm{d}_i;\ \ L^{(S-1)}_{n-1} = L_d.$$

Введем в L_d базис $\{e_i\}$ так, чтобы $e_1 \| \theta$, $\theta = \theta e_1$. Размерность L обозначим l. В данном случае $l = S-1$, но для общности вычисления проведем при произвольном l.

Разложим θ'_n на составляющие $\theta'_{n\|}$ и $\theta'_{n\perp}$, $\theta'_n = \theta'_{n\|} + \theta'_{n\perp}$.

Имеем далее

$$\mathbf{z}_n^T\theta'_n = \mathbf{z}_n^T\theta'_{n\|},$$

поскольку $\theta'_{n\perp}$ ортогонально $L^{(S-1)}_{n-1}$, а значит, и всем векторам из $L^{(S-1)}_{n-1}$. В упрощенных обозначениях имеем

$$\mathbf{z}_n = \sum_{i=1}^{S-1}(\theta^{\mathrm{T}}d_i)d_i,\ \ \theta'^{\mathrm{T}}{}_{n\|}\mathbf{z}_n = \sum_{i=1}^{S-1}(\theta^{\mathrm{T}}d_i)^2$$

и в выбранном базисе $\theta^{\mathrm{T}}d_i = \|\theta\| d_{1i}$ (d_{1i} — первая компонента d_i).

Для ψ получаем

$$M\{\psi\} = M\left\{\frac{(\mathbf{z}_n^{\mathrm{T}}\theta'_{n\|})^2}{\|\mathbf{z}_n\|^2}\right\} = \|\theta\|^2 M_{\{d_i\}}\left\{\frac{\mathrm{d}_{11}^2 + \ldots + \mathrm{d}_{1s-1}^2}{\sum_{i,\,j=1}^{S-1} d_{1i}d_{1j}(d_i^{\mathrm{T}}d_j)}\right\}. \tag{18}$$

Для вычисления этой величины рассмотрим матрицу Δ размерности $l\times(S-1)$, столбцами которой являются координаты векторов d_i, $(i = 1,\ldots,S-1)$ в выбранном базисе l-мерного пространства L. Элементы этой матрицы — независимые гауссовы случайные величины с нулевым математическим ожиданием и одинаковой дисперсией, и $M(\psi)$ получается усреднением выражения в (18) по $l\times(S-1)$-мерному фазовому пространству, являющимся прямым произведением $(S-1)$ l- мерных пространств, отвечающих векторам d_i. Представим это пространство как прямое произведение l $(S-1)$-мерных пространств, отвечающих векторам D_K, $(k = 0, S-1)$, компонентами которых являются строки матрицы Δ, т.е. $D_{iK} = d_{Ki}$.

 ISSN 0572-2691

При введенных векторах D_K величина $M\{\psi\}$ выражается следующим образом:

$$M\{\psi\} = M_{\{D\}}\left\{ \frac{\|D_1\|^4}{\sum_{m=1}^{l}(D_1^{\mathrm{T}} D_m)^2} \right\}, \tag{19}$$

поскольку из определения D_K следует

$$\sum_{i,\ j,\ k=1}^{i,\ j=s-1,\ k=l} d_{1i} d_{1j} d_{k_i} d_{kj} = \sum_{i,\ j,\ k=1}^{i,\ j=s-1,\ k=l} D_{i1} D_{j1} D_{ik} d_{jk} = \sum_{k=1}^{l} (D_1^{\mathrm{T}} D_k)^2.$$

Плотность распределения вероятности величин D_i в $l \times (S-1)$-мерном пространстве компонент набора $\{D_i\}$ имеет вид

$$\left(\sqrt{\frac{\pi}{\omega}}\right)^{l(S-1)} \exp\{-\omega(\|D_1\|^2 + \ldots + \|D_{s-1}\|^2)\}.$$

Воспользуемся формулой

$$\int_0^{\infty} \mathrm{e}^{-tz} \mathrm{d}z = \frac{1}{z},\ \ z > 0.$$

Подставив в нее знаменатель дроби для ψ из (19)

$$z = \|D_1\|^2 + \sum_{m=1}^{l} \left(\frac{D_1^T D_m}{\|D_1\|^2} \right)^2.$$

Получим для $M\{\psi\}$ следующее выражение:

$$M\{\psi\} = \left(\sqrt{\frac{\omega}{\pi}}\right)^{l(S-1)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} exp\{-\alpha(\|D_1\|^2 + \ldots + \|D_{s-1}\|^2)\} \times$$

$$\times \exp\left\{ t\left(\|D_1\|^2 + \sum_{m=1}^{l} \left(\frac{D_1^T D_m}{\|D_1\|^2} \right)^2 \right) \right\} \|D_1\|^2 dD_1 \cdots dD_l dt. \tag{20}$$

Изменив в (20) порядок интегрирования (интегрируя сначала по $D_2, \ldots, D_l$ при фиксированном D_1), получим произведение $l-1(S-1)$-мерных интегралов, каждый из которых (обозначим их J_1) имеет вид

$$J_1 = \left(\sqrt{\frac{\omega}{\pi}}\right)^{S-1} \underbrace{\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty}}_{S-1} \mathrm{e}^{-\omega(\|D\|^2 + t(D_{\bar{n}}^{\mathrm{T}}))} dD.$$

Здесь $\bar{n}$ — единичный вектор $D_1\|D_1\|^{-1}$. Но J_1 не зависит от $\bar{n}$, поскольку подынтегральная функция инвариантна относительно поворота. Поэтому вычислим J_1 в базисе, в котором $e_1 \,|\, \bar{n}$.

В этом случае выражение, стоящее в экспоненте, равно

$$(\alpha + t)D_1^2 + \ldots + \alpha(D_2^2 + \ldots + D_{S-1}^2)$$

и $(S-1)$-мерный интеграл распадается на произведение $S-2$ интегралов, дающих с множителем нормировки $\sqrt{\frac{\omega}{\pi}}$ единицу, и интеграла

$$J=\sqrt{\frac{\omega}{\pi}}\int\limits_{-\infty}^{\infty}\mathrm{e}^{-(\alpha+t)D_1^2}\,\mathrm{d}D_1,$$

который легко вычисляется и равен $\sqrt{\frac{\omega}{\omega+t}}$.

Остается вычислить интеграл по D_1 и t. Сначала вычислим интеграл по D_1:

$$J_1=\int\limits_{-\infty}^{\infty}\mathrm{e}^{-(\omega+t)\|D_1\|^2}\|D_1\|^2\,dD_1.$$

Поскольку

$$J(\gamma)=\int\limits_{-\infty}^{\infty}\mathrm{e}^{-\gamma\|D_1\|^2}\,\mathrm{d}D_1=\left(\sqrt{\frac{\pi}{\gamma}}\right)^{S-1},$$

имеем

$$J_1=-\frac{\partial J(\gamma)}{\partial\gamma}\bigg|_{\gamma=\omega+t}=\frac{S-1}{2}\left(\sqrt{\frac{\pi}{\omega+t}}\right)^{S-1}\frac{1}{\omega+t}.$$

И окончательно, используя формулу

$$\int\limits_{0}^{\infty}(a+x)^{-\nu}\,dx=\frac{a^{-\nu+1}}{\nu-1},$$

получаем для $M\{\psi_n\}$ выражение

$$M\{\psi_n\}=\frac{s-1}{2}\int\limits_{0}^{\infty}(1+t)^{-\frac{(l+S)}{2}}\,dt=\frac{S-1}{l+S-1}.$$

При $l=S-1$ $M\{\psi_n\}=0{,}5$, что отвечает значению параметра $\lambda=2$. Таким образом, доказано утверждение 1.

Рассмотрим теперь, с какой скоростью будет сходиться алгоритм, в котором используется процедура 3. Вычислить эту скорость можно, не используя приближение, полученное выше для алгоритмов с $N-S>1$.

Утверждение 2. Процедура 3 сходится со скоростью геометрической прогрессии со знаменателем

$$\varepsilon=1-\frac{1+\beta}{N}, \tag{21}$$

где β зависит от соотношения S и N.

Доказательство. Проецирование на вектор $\mathbf{x}_{n-s}$ при $S>1$ может уменьшить величину $\|\theta_n\|^2$, потому что предыдущие $S-1$ шагов алгоритма изменяют направление вектора θ_n по сравнению с его направлением на $(n-S)$-м шаге, когда он перпендикулярен $\mathbf{x}_{n-s}$ и $\theta_{n-S}^{\mathrm{T}}\,\mathbf{x}_{n-s}=0$.

 ISSN 0572-2691

Поэтому сначала рассмотрим вспомогательную задачу, формулируемую так: насколько в среднем увеличивается квадрат нормы проекции вектора θ на вектор e после итерации по алгоритму Качмажа вдоль случайного вектора b. Пусть векторы e и b независимы (это справедливо для проецирования на вновь приходящие векторы $\mathbf{x}_n$). Введя обозначения P_e и P_b для операторов проецирования на векторы e и b, с учетом явного вида P_e и P_b вычислим величину $M_b\left\{\left\|\theta_e'\right\|^2\right\}$, $\theta' = (I - P_b)\theta$, где, как и ранее, $\theta' = P_e\theta'$, $\theta_e = P_e\theta$.

Имеем

$$M_b\left\{\left\|\theta_e'\right\|^2\right\} = M\{\theta^{\mathrm{T}}(I - P_b)P_e^2(I - P_b)\theta\} \tag{22}$$

и задача сводится к усреднению по случайному вектору b операторов P_bP_e и $P_bP_eP_b$. Так как b не зависит от e,

$$M\{P_bP_e\} = [M\{P_b\}]P_e = \frac{1}{N}P_e, \tag{23}$$

поскольку

$$M_b\{P_b\} = \frac{1}{N}I. \tag{24}$$

Математическое ожидание оператора $P_bP_eP_b$ вычислялось выше даже в более общей ситуации. В нашем случае имеем

$$M_b\{P_bP_eP_b\} = \frac{1}{N(N+2)}I + \frac{2}{N(N+2)}P_e. \tag{25}$$

Окончательно с точностью до членов $o(N^{-2})$ получаем

$$M\left\{\left\|\theta_e'\right\|^2\right\} = \left\|\theta_e\right\|^2\left(1 - \frac{2}{N}\right) + \frac{1}{N^2}\left\|\theta\right\|^2.$$

Ясно, что если рассматривается проецирование на вектор $\mathbf{x}_{n-S+1}$, на который уже производилось проецирование на одном из предыдущих шагов, то эта ситуация моделируется задачей, аналогичной рассмотренной выше, но с корреляцией между векторами θ и e, поскольку проекция θ_n на $\mathbf{x}_{n-S+1}$ в среднем в β раз меньше, чем проекция θ_n на случайно ориентированный вектор e. Введенный здесь параметр β как раз и будет характеризовать, насколько дополнительно уменьшится величина квадрата нормы после проецирования на вектор $\mathbf{x}_{n-S+1}$, на который уже производилось проецирование $S-1$ шагов назад (тогда вектор θ_{n-S+1} не зависел от этого вектора $\mathbf{x}_{n-S+1}$). После $(S-1)$-го шага θ_{n-S+1} и $\mathbf{x}_{n-S+1}$ ортогональны, затем проекция θ_{n-i} на $\mathbf{x}_{n-S+1}$ возрастает по величине, но если величина проекции θ на случайным образом ориентированный вектор e дается формулой $M\{\theta^{\mathrm{T}}P_e\theta\} = N^{-1}[\theta]^2$, то сохраняющаяся «антикорреляция» векторов θ_{n-i} и $\mathbf{x}_{n-S+1}$ приводит к тому, что величина проекции θ_n на $\mathbf{x}_{n-S+1}$ будет меньше. Этот факт описывается соотношением

$$M_e\{\theta^T P_e\theta\} = \beta\frac{1}{N}\left\|\theta\right\|^2, \text{ где } \beta < 1.$$

Формулы (22) и (23) не изменятся и подстановка β в (22) дает формулу

$$M_b\left\{\left\|\theta_e'\right\|^2\right\}=\left\|\theta_e\right\|^2\left(1-\frac{2\beta}{N}\right)+\frac{1}{N^2}\left\|\theta\right\|^2. \tag{26}$$

Работу алгоритма теперь можно описать следующим образом. Шаг по алгоритму Качмажа дает убывание величины $M\{\left\|\theta_n\right\|^2\}$ в $1-N^{-1}$ раз, а производимое затем проецирование на вектор $\mathbf{x}_{n-S+1}$ уменьшает $\left\|\theta\right\|^2$ еще в $1-\beta N^{-1}$ раз. Уравнение для определения параметра β получается итерированием соотношений (22) и (26) с учетом изменения $\left\|\theta\right\|^2$ на каждом шаге. Обозначим t_i математическое ожидание проекции θ_n на вектор x_n после i шагов алгоритма $M\{\left\|\theta_{n+i}\right\|^2\}=\nu_i$. С учетом этого имеем следующую систему рекуррентных уравнений:

$$\begin{cases} t_{i+1}=t_i(1-2\beta_i N^{-1})+N^{-2}\nu_i, \\ \nu_{i+1}=\nu_i(1-\beta_i N^{-1}), \;\; i=\overline{0,S-1}, \end{cases} \tag{27}$$

причем $\beta_{2i+1}=I$, $\beta_{2i}=\beta$. Начальные условия при $i=0$ после шага по алгоритму Качмажа дают $t_o=0$, $\nu_o=\nu_o$.

Решение системы (26) в общем виде запишем так:

$$\nu_{i+1}=\nu_o\prod_{j=1}(1-\beta_j N^{-1});$$

$$t_{i+1}=\sum_{l=1}^{i}\prod_{j=l}\left(1-\frac{2\beta_i}{N}\right)\prod_{i=1}^{l-1}\left(1-\frac{\beta_i}{N}\right). \tag{28}$$

Необходимо определить, какое значение примет t_i на $2(S-1)$-м шаге, поскольку на этом шаге происходит повторное проецирование на вектор $\mathbf{x}_n$, величина $\left\|\theta_n\right\|^2$ уменьшается как раз на βN^{-1}, и скорость сходимости процедуры 3 будет выражаться формулой

$$\varepsilon=\left(1-\frac{1}{N}\right)\left(1-\frac{1}{\beta}\right).$$

Поскольку по определению параметра β имеем

$$M\left\{\left\|\theta'_{n+2(S-1)}\right\|^2-\left\|\theta_{n+2(S-1)}\right\|^2\right\}=t_{2(S-1)}=\frac{\beta}{N}\nu_{2(S-1)}, \tag{29}$$

то в принципе подстановка $t_{2(S-1)}$ и $\nu_{2(S-1)}$ из общего решения дает уравнение для β. Получим это уравнение, выписанное с точностью до членов $o(N^{-1})$, поскольку реально всегда $N>1$. При этом

$$\varepsilon=\left(1-\frac{\beta_i}{N}\right)\left(1-\frac{\beta_{i+1}}{N}\right).$$

Выразив t_{i+2} и ν_{i+2} через t_i и ν_i, с учетом того, что $\varepsilon^2=(1-2\beta_i N^{-1})(1-2\beta_{i+1}N^{-1})$, можно записать с точностью до членов

$$t_{i+2} = \varepsilon^2 t_i + 2N^{-2}\nu_i,\ \ \nu_{i+2} = \varepsilon\nu_i.$$

Отсюда находим

$$\nu_{2(S-1)} = \varepsilon^{S-1}\nu_o,\ \ t_{2(S-1)} = 2N^{-2}\nu_o(\varepsilon^{2(S-2)} + \ldots + \varepsilon^{(S-2)}).$$

Этот вывод более нагляден, чем формальное разложение общего решения системы (26), даваемого формулами (27), по малому параметру N^{-1}. Подставляя $t_{2(S-1)}$ и $\nu_{2(S-1)}$ в (28), получаем для ε алгебраическое уравнение

$$1-\varepsilon = N^{-1} + 2N^{-1}(1+\varepsilon+\ldots+\varepsilon^{S-2}). \tag{30}$$

При $s \to \infty$ из (29) имеем уравнение для $1-\varepsilon = (1+\beta)N^{-1}$:

$$1-\varepsilon = \frac{1}{N} + \frac{2}{N^2(1-\varepsilon)},$$

имеющее положительный корень $\varepsilon = 1-2N^{-1}$, что согласуется с «геометрической моделью» уравнения, даваемой проекционным алгоритмом, так как условие $s \to \infty$ означает, что «память» о проецировании на вектор $\mathbf{x}_n$ уже стерта и дополнительное проецирование на этот вектор увеличивает скорость сходимости ровно вдвое, если эту скорость характеризовать величиной $1-\varepsilon$. При $SN^{-1} < 1$ из (30) получаем

$$\varepsilon \approx 1 - \frac{1}{N}\left(1 + \frac{2(S-1)}{N}\right).$$

При $s \sim N$ с учетом того, что $\varepsilon = 1-(1+\beta)N^{-1}$, получаем для β уравнение

$$\beta = \frac{2}{1+\beta}\left[1-\left(1-\frac{1+\beta}{2}\right)^{s-1}\right],$$

которое при $SN^{-1} \sim 1$ можно преобразовать к виду

$$\beta = \frac{2}{1+\beta}\left(1-e^{-(1+\beta)\frac{S-1}{N}}\right).$$

Например, при $SN^{-1} = 0{,}5$ имеем корень $\beta = 0{,}7$.

Скорость сходимости алгоритма с такой же глубиной S будет равной скорости сходимости многошагового проекционного алгоритма (с точным проецированием) при $S \approx 0{,}4N$.

Заключение

Как показали результаты исследований, использование аппроксимации оператора ортогонального проецирования в алгоритмах оценивания позволяет при незначительном снижении скорости сходимости алгоритмов существенно упростить их реализацию и повысить их вычислительную устойчивость вследствие устранения операции обращения матрицы наблюдений. При этом для гарантированного улучшения их вычислительных свойств целесообразно применить достаточно простую в данном случае процедуру регуляризации алгоритма Качмажа с использованием существующих рекомендаций по выбору параметра регуляризации [6].

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

ПСЕВДОПРОЕКЦІЙНІ АЛГОРИТМИ ОЦІНЮВАННЯ, ЩО БАЗУЮТЬСЯ НА АПРОКСИМАЦІЇ ОПЕРАЦІЇ ОРТОГОНАЛЬНОГО ПРОЕЦІЮВАННЯ

Одним з найбільш ефективних і найбільш простих в обчислювальному відношенні однокрокових алгоритмів оцінювання є алгоритм Качмажа, запропонований в [1]. У багатьох подальших роботах було розглянуто можливість прискорення алгоритму Качмажа шляхом використання не одного, а ряду вимірювань. Роботи [8, 9] послужили поштовхом до розробки нового класу багатокрокових проекційних алгоритмів [10–14], в яких при побудові оцінки на *n*-му кроці використовується не тільки нова інформація, як це відбувається в алгоритмі Качмажа, а й інформація про ряд попередніх кроків $n-1$, $n-2$ .... . Кількість таких кроків визначає пам'ять алгоритму. При цьому завдяки кращій екстраполяції і фільтрації у ряді випадків вдається домогтися істотного скорочення часу ідентифікації. Реалізація багатокрокового (*S*-крокового) проекційного алгоритму вимагає обчислення зворотної матриці спостережень розмірності $S \times S$. В [11–14] встановлено властивості випадкових псевдообернених матриць і матриць проеційованих, які дозволили визначити швидкість збіжності даних алгоритмів і зробити висновок про те, що урахування в даних алгоритмах інформації про *S* попередніх кроків рівносильне в сенсі швидкості збіжності зменшенню розмірності вихідного простору *N* на *S*. У даній роботі пропонуються і досліджуються псевдопроекційні алгоритми, що володіють близькими до багатокрокових проекційних алгоритмів властивостями, але більш прості в реалізації. Дані алгоритми використовують апроксимацію операції точного проеціювання і будуються на основі однокрокового адаптивного алгоритму Качмажа. Отримано оцінки швидкості збіжності запропонованих процедур і показано, що використання такої апроксимації дозволяє при незначному зниженні швидкості збіжності алгоритмів істотно спростити їх реалізацію та підвищити обчислювальну стійкість внаслідок усунення операції обертання матриці спостережень.

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

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

PSEUDOPROJECTION ESTIMATION ALGORITHMS BASED ON APPROXIMATION OF ORTHOGONAL PROJECTION OPERATION

Kaczmarz algorithm proposed in [1] is one of the most effective and most computationally simple one-step estimation algorithms. In a number of subsequent studies the possibility of acceleration algorithm Kaczmarz by the use of not one but a series of measurements was examined. Research made in [8, 9] was a push for the development of a new class of algorithms — multistage projection algorithms [10–14], in which during the construction of estimates on the *n*-th step not only new information is used, as it comes in Kaczmarz algorithm, but also information about number of previous steps $n–1$, $n–2$ .... . The number of these steps determines the algorithm's memory. In this case, thanks to a better extrapolation and filtering, in some cases it is possible to reduce significantly the time of identification. The implementation of multi-step (S -step) projection algorithm requires the computation of the inverse matrix of observations with $S \times S$ dimension. In [11–14] properties of random pseudoinverse matrix and the projection matrix are set. It helped to determine the rate of convergence of these algorithms and conclude that taking into account information about previous *S* steps is equivalent in terms of convergence speed to reduction of dimension *N* in the original space to *S*. In this paper we propose and investigate pseudoprojective algorithms that have close to the multi-stage projection algorithm properties, but are simpler to implement. These algorithms use an approximation of the exact projection operation and are based on one-step Kaczmarz adaptive algorithm. Estimates of the convergence rate of the proposed procedures are obtained. It is shown that the use of such approximation allows, with a slight decrease in the convergence rate of the algorithms, to simplify significantly their implementation and increase computational stability due to eliminating the rotation operation of the observation matrix.

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

1. Kaczmarz, S. Approximate solution of systems of linear equations. *Int. J. of Control*. 1993. № 57. P. 1269–1271. https://doi.org/10.1080/00207179308934446.
2. Чадеев В.М. Определение динамических характеристик объектов в процессе их нормальной эксплуатации для целей самонастройки. *Автоматика и телемеханика*. 1964. 25. № 9. С. 1302–1306.
3. Райбман Н.С., Чадеев В.М. Адаптивные модели в системах управления. М.: Сов. радио, 1966. 156 с.
4. Aved'jan A.D. Bestimmung der Parameter linearer Modelle stationarer und instationarer Strecken. *Messen, steuern, regeln*. 1971. N 9. P. 348–350.
5. 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.
6. 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.
7. Либероль, Б.Д., Руденко, О.Г., Тимофеев, В.А. Модифицированный алгоритм Качмажа для оценивания параметров нестационарных объектов. *Проблемы управления и информатики.* 1995. № 4. С. 81–89
8. Лелашвили, Ш. Г. Применение одного итерационного метода для анализа многомерных автоматических систем. *Схемы автоматического управления*. 1965. С.19–33.
9. Лелашвили, Ш. Г. Некоторые вопросы построения статистической модели многомерных объектов. *Автоматическое управление*. 1967. С. 59–96.
10. Аведьян, Э.Д. Модифицированные алгоритмы Качмажа для оценки параметров линейных объектов. *Автоматика и телемеханика.* 1978. № 5. С. 64–72.
11. Ищенко Л.А., Либероль Б.Д., Руденко О.Г. Проекционные алгоритмы идентификации линейных объектов. *Докл. АН УССР. Сер. А.* 1985. № 7. С. 62–64.
12. Ищенко Л.А., Либероль Б.Д., Руденко О.Г. Адаптивное оценивание параметров нестационарных объектов. *Докл. АН УССР. Сер. А.* 1985. № 12. С. 70–72.
13. Ищенко Л.А., Либероль Б.Д., Руденко О.Г. О свойствах одного класса многошаговых адаптивных алгоритмов идентификации *Кибернетика.* 1986. №1. С.92–96.
14. Либероль Б.Д., Руденко О.Г. О свойствах проекционных алгоритмов оценивания параметров нестационарных объектов. *Докл. АН УССР. Сер. А.* 1990. № 4. С. 71–74.
15. Цыпкин Я.З. Адаптация и обучение в автоматических системах. М.: Наука, 1968. 399 с.
16. Поляк Б.Т Сравнение скорости сходимости одношаговых и многошаговых алгоритмов оптимизации при налички помех измерений. *Изв. АН СССР. Техническая кибернетика*.1974. № 5. С. 9–12.
17. Себер Дж. Линейный регрессионный анализ. М.: Мир, 1980. 456 с.

Получено 02.12.2019