

Обоснование и исследование адаптивного метода решения обратных задач геофизики

© 2007 В.А. Кочнев

(Институт вычислительного моделирования СО РАН, Красноярск)

Аннотация

В работе приводится обоснование и исследование адаптивного (итерационно-стохастического) метода, предназначенного для решения систем алгебраических уравнений (линейных и нелинейных, переопределенных и недоопределенных), возникающих в обратных задачах геофизики. Особенностью метода является то, что ставится задача уточнения априорно заданных неизвестных с учетом априорных оценок погрешностей неизвестных, а также погрешностей правых частей уравнений, которые являются результатом измерений и, естественно, содержат ошибки. Уточнение производится последовательно в каждом уравнении системы. Метод использовался для решения обратных задач геофизики с большим числом неизвестных (задачи сейсморазведки, гравиметрии и магнитометрии).

ВВЕДЕНИЕ

Многие обратные задачи геофизики сводятся к решению алгебраических систем (линейных или нелинейных) уравнений, в которых неизвестными являются параметры изучаемой модели среды. Для слоистых сред, которые чаще всего изучают в нефтяной геофизике, параметрами являются мощности слоев и распределение в них физических свойств, таких как скорость пробега упругих волн, плотность, проводимость, магнитная восприимчивость и других. При постановке каждой конкретной задачи выбирается модель среды. При общей постановке вводится вектор неизвестных параметров, обозначаемый через X .

Некоторые измеренные или вычисленные характеристики наблюдаемых полей (времена прихода волн, их амплитуды на определенной частоте и т.д.) обозначаются через вектор U .

Предполагается, что имеется метод решения прямой задачи, т.е. математический оператор перехода $X \rightarrow U$. В общем случае оператор может быть как линейным, так и нелинейным. Запишем его

$$F(X) = U. \quad (1)$$

Тогда обратная задача в общем виде может быть представлена в виде

$$X = F^{-1}(U). \quad (2)$$

Однако, если прямая задача при заданных ограничениях имеет, как правило, единственное решение, то об обратных задачах этого сказать нельзя. Одному и тому же U может удовлетворять различный набор векторов X . Небольшому изменению U могут соответствовать большие изменения в векторе X . Такие обратные задачи относятся к некорректным [Тихонов, 1979]. Их решение становится устойчивым, если наложить ограничение на область решения. Задачи такого типа получили название условно корректных. Разработка методов решения такого класса задач в области геофизики была начата Тихоновым А.Н. (1943) и продолжена Лаврентьевым М.М., Ивановым В.К. и многими их учениками и последователями.

Ограничение области решения заключается в требовании выбора решения в некоторой области, близкой к начальному приближению X_0 , т.е. минимизации функционала

$$\min \theta(X) = \|U - F(X)\|^2 + \alpha \|X - X_0\|^2. \quad (3)$$

где α – параметр регуляризации, зависящий от точности исходных данных. Выбор его является одной из задач метода регуляризации [Тихонов, 1979].

Возможны и другие подходы к постановке обратных задач и методам их решения. Будем рассматривать в дальнейшем статистическую постановку условий задачи.

Предположим $U(u_1, \dots, u_i, \dots, u_m)$

$$u_i = u_i^* + \xi_i, \quad (4)$$

где ξ_i – гауссовская помеха с $N(0, \sigma_i^2)$ или корреляционной матрицей $R_{UU}^{(0)}$;

$$X^{(0)}(x_1^{(0)}, \dots, x_j^{(0)}, \dots, x_n^{(0)}) \quad (5)$$

$$x_j^{(0)} = x_j^* + \Delta x_j^{(0)}$$

где $\Delta x_j^{(0)}$ – неизвестная ошибка априорных данных с $N(0, \sigma_j^2)$ или $R_{xx}^{(0)}$.

Известно несколько постановок при этих условиях задачи. Среди них назовем байесовскую, метода апостериорной вероятности и метода максимума правдоподобия (ММП).

Впервые вероятностная информационная постановка для решения обратных геофизических задач была предложена Халфиным Л.А. (1958). В дальнейшем это направление развивалось Гольцманом Ф.М. (1971) и его учениками, Трояном В.М., Киселевым В.С., Кутьиной О.Г., Калининой Т.Б. и др.

С.В. Гольдин (1970), используя байесовскую постановку, минимизирующую средний риск при квадратичной функции потерь, для решения обратной задачи получил следующую формулу

$$A^T R_{UU}^{-1} A + (R_{xx}^{(0)})^{-1} X = A^T R_{UU}^{-1} U + (R_{xx}^{(0)})^{-1} X_0, \quad (6)$$

где $A_{ij} = \frac{\partial F_i}{\partial x_j}$, которая для частного случая

$$R_{UU} = \sigma_u^2 I, \quad R_{xx}^{(0)} = \sigma_{xx}^2 I \quad \text{и} \quad X_0 = 0 \quad (7)$$

приводит к

$$(A^T A + \alpha I) X = A^T U, \quad (8)$$

где $\alpha = \sigma_u^2 / \sigma_x^2$.

Следует также отметить работы Франклина (1970) и Джексона (1979), в которых задача решена для линейного случая с использованием ММП. Для нелинейного случая решение приводится в работе А. Тараноллы и Б. Валетте (1982). Следуя работе этих авторов, примем:

$$Z_0 = \begin{pmatrix} U \\ X_0 \end{pmatrix}; \quad C_0 = \begin{pmatrix} R_{UU} & R_{UX} \\ R_{XU} & R_{XX} \end{pmatrix}. \quad (9)$$

Обратная задача сводится к решению системы

$$\begin{cases} F(Z) = 0, \\ \max_Z \rho(Z) = \text{const} \cdot \exp\left\{-\frac{1}{2}(Z - Z_0)^T C_0^{-1}(Z - Z_0)\right\}. \end{cases} \quad (10)$$

В такой постановке для нелинейной задачи получено

$$Z_{(l+1)} = Z_0 + C_0 A_l^T \left(A_l \cdot C_0 \cdot A_l^T \right)^{-1}; \quad \{ A_l \cdot (Z_l - Z_0) - f(Z_l) \} \quad (11)$$

или при $R_{XU} = R_{UX}^T = 0$

$$\begin{aligned} X_{(l+1)} &= X_l + \left(A_l^T \cdot R_{UU}^{-1} A_l + R_{XX}^{-1} \right)^{-1} \cdot \\ &\cdot \left\{ A_l^T \cdot R_{UU}^{-1} [U - F(X_l)] - R_{XX}^{-1} (X_l - X_0) \right\}, \end{aligned} \quad (12)$$

а для линейного случая

$$X = X_0 + R_{X_0 X_0} A^T \left[R_{UU} + A \cdot R_{X_0 X_0} A^T \right]^{-1} (U_0 - A X_0), \quad (13)$$

$$R_{XX} = R_{X_0 X_0} - R_{X_0 X_0} A^T \left(R_{UU} + A R_{X_0 X_0} \cdot A^T \right)^{-1} \cdot A \cdot R_{X_0 X_0}, \quad (14)$$

где совместно с уточнением неизвестных изменяется (уменьшается) матрица ковариации.

Это можно проинтерпретировать как проектирование из X_0 в X и получение апостериорного значения R_{XX} . Причем здесь нет ограничений на вид матрицы и ее определитель. Число уравнений может быть меньше, чем число неизвестных.

Используя эту особенность в линейной задаче можно построить рекуррентный метод уточнения параметров, поочередно включая в обработку одно уравнение за другим. В этом случае из (13), (14) получим

$$X^{(k+1)} = X^{(k)} + B_{(k+1)} (U_{(k+1)} - A_{(k+1)} X_{(k)}), \quad (15)$$

$$R_{XX}^{(k+1)} = R_{XX}^{(k)} B_{(k+1)} A_{(k+1)} R_{XX}^{(k)}, \quad (16)$$

где $B_{(k+1)} = R_{XX}^{(k)} A_{(k+1)}^T \left[(\sigma_U^2) + A_{(k+1)} R_{XX}^{(k)} A_{(k+1)}^T \right]^{-1}$, а $A_{(k+1)}$ – коэффициенты $(k+1)$ строки матрицы

A , k меняется от 0 до $m-1$ (m – число уравнений).

Рекуррентные формулы (15), (16) обобщают формулы, полученные Жуковским Е.Л., Липцером Р.Ш. из методов фильтрации Калмана-Бьюси. В формулах Жуковского-Липцера

[Липцер, 1974] вместо (σ_U^2) стоит α – параметр регуляризации. Для решения формальных систем уравнений начальные приближения принимаются равными нулю с $R_{XX}^{(0)} = 1$. Достоинством метода является то, что он за один проход всех уравнений позволяет получить искомое решение и $R_{XX}^{(k+1)}$. Однако, он как и обычный метод обращения (13), (14) позволяет решать лишь системы с небольшим числом неизвестных. Это обусловлено следующими причинами.

1. С ростом числа уравнений растут ошибки, связанные с умножением матриц.
2. Время счета растет пропорционально n^3 .
3. Растет пропорционально n^2 память, необходимая для хранения ковариационных матриц.

В связи с этим встала задача создания метода (в классе итерационных), который бы был лишен указанных недостатков. Обоснование такого итерационного (адаптивного) метода будет дано в последующих разделах.

1. ПОСТАНОВКА ЗАДАЧИ

Как уже говорилось, в задачах геофизики очень часто требуется решать алгебраические системы с большим количеством неизвестных с сильно разреженными матрицами (обратная кинематическая задача сейсморазведки). Для решения таких систем, как правило, используются итерационные методы [Корн, 1973], [Wiggins, 1976], обоснование которых следует не из общих критериев, а из эвристических соображений с последующим исследованием свойств метода.

С учетом сказанного дополним и конкретизируем приведенную выше общую постановку задачи.

Пусть имеется итерационный алгоритм (формулы (1.1), (1.2)), который позволяет по очередному i -ому уравнению уточнять каждый из параметров, входящих в это уравнение, и вместе с тем, неким разумным образом, изменять оценку σ_j^2 , т.е. диагональный элемент, соответствующий неизвестному x_j . В дальнейшем i будет обозначать номер уравнения, а j

– номер неизвестного, k – номер шага, а l – номер итерации. В общем случае $k = i + m(l - 1)$, где m – число уравнений. На первой итерации $k = i$. Итак,

$$x_{j(k+1)} = x_{j(k)} + \alpha_{j(k+1)} \Delta u_i, \quad (1.1)$$

$$\sigma_{j(k+1)}^2 = \gamma_{j(k+1)} \sigma_{j(k)}^2, \quad (1.2)$$

где $\Delta u_i = \tilde{y}_i - f_i(\bar{x}_{1(k)}, \dots, \bar{x}_{j(k)}, \dots, \bar{x}_{n(k)})$, а $\bar{x}_{j(k)}$ – прогнозные значения параметра. При вычислении прогнозных значений могут использоваться методы аппроксимации, учитывающие взаимосвязь неизвестных [Кочнев, 1983].

Для оценки статистических свойств невязки после уточнения параметров на всех уравнениях l -ой итерации будем определять

$$\Delta \bar{u}_{(l)} = \left(\sum_{i=1}^m \Delta u_{i(l)} \cdot P_i \right) / \sum_{i=1}^m P_i, \quad (1.3)$$

$$D_{(l)} = \left(\sum_{i=1}^m \Delta u_{i(l)}^2 \cdot P_i \right) / \sum_{i=1}^m P_i, \quad (1.4)$$

где $P_i = 1/\sigma_i^2$.

В том случае, когда наблюдения равноточны, приходим к общеизвестной оценке невязок: среднему и дисперсии. В конце каждой итерации выдается средняя квадратическая оценка невязки $\sqrt{D_{(l)}}$.

Задача заключается в том, чтобы в (1.1) и (1.2) выбрать такие $\alpha_{j(k+1)}$ и $\gamma_{j(k+1)}$, которые бы обеспечили

- 1) оптимальные на каждом i -ом уравнении распределение невязки и изменение σ_j^2 ;
- 2) уменьшение D при переходе от l -ой итерации к $l+1$.

При выполнении этих требований за решение будем принимать вектор X , полученный на последнем шаге $l+1$ итерации, после которой будет выполнено условие

$$D_{(l)} - D_{(l+1)} \leq |\varepsilon|, \quad (1.5)$$

где ε – малое, наперед заданное число.

2. ОБОСНОВАНИЕ ВЫБОРА КОЭФФИЦИЕНТА α

Наличие невязки Δu_i в общем случае обусловлено следующими причинами:

– отклонением прогнозных значений параметров от истинных x_j^*

$$\Delta x_j = x_j^* - \bar{x}_j; \quad (2.1)$$

– ошибками измерения

$$\xi_i = \tilde{u}_i - u_i^*; \quad (2.2)$$

– неадекватностью математической и физических моделей.

Адекватность математической модели физической обычно исследуется на этапе выбора метода решения прямой задачи. И если прямая задача решается правильно, то есть основания предполагать, что третья причина невязки отсутствует и основная невязка обусловлена первой и второй причинами. Разложим невязку в виде суммы следующих слагаемых

$$\Delta u_i \approx \sum_{j=1}^n \frac{\partial u_i}{\partial x_j} \cdot \Delta x_j + \xi_i. \quad (2.3)$$

В линейном случае производные будут равны коэффициентам уравнения, а в нелинейном – приближенными, если начальное приближение будет находиться вдали от точного вектора неизвестных. По мере уточнения производные будут приближаться к точным. В дальнейшем будем предполагать, что в нелинейном случае начальное приближение находится вблизи точного решения и, следовательно, выражение (2.3) является точным.

Обозначим слагаемые под знаком суммы через z_j , получим

$$\Delta u_i = \sum_{j=0}^n z_j, \quad \text{где } z_0 = \xi_i. \quad (2.4)$$

Предположим, что все слагаемые являются независимыми случайными величинами, распределенными по нормальному закону. Следует заметить, что постулируется независимость не самих неизвестных x_j , а их невязок, точнее $a_{ij}\Delta x_j$. Некоррелируемость Δu_i и параметров Δx будет выполняться при $\Delta u_i \Rightarrow 0$. Если невязки Δu_i не близки к 0, то, как увидим далее, она будет уменьшаться за счет уточнения слагаемых. Показателем уточнения

неизвестных будет уменьшение средних квадратических невязок. Найдем совместную плотность вероятности Δu_i в $n+1$ -мерном пространстве.

$$\rho(z_0, \dots, z_j, \dots, z_n) = \prod_{j=0}^n \frac{1}{(\sqrt{2\pi}\sigma_{z_j})} \cdot \exp\left(-\frac{z_j^2}{2\sigma_{z_j}^2}\right), \quad (2.5)$$

где $\sigma_{z_j}^2 = \left(\frac{\partial u_i}{\partial x_j} \sigma_{x_j}\right)^2$; $\sigma_{z_0}^2 = \sigma_{u_i}^2$; $z_0 = \left(\Delta u_i - \sum_{j=1}^n z_j\right)$.

Значения z_j выберем таким образом, чтобы плотность вероятности была максимальной (метод максимума правдоподобия) [Крамер, 1975].

$$\max(\ln \rho(z_0, \dots, z_j, \dots, z_n)). \quad (2.6)$$

Откуда, поменяв знаки, получим целевую функцию в виде

$$\min(z_1, \dots, z_j, \dots, z_n) = \frac{(\Delta u_i - \sum_1^n z_j)^2}{\sigma_{u_i}^2} + \sum_1^n \frac{z_j^2}{\sigma_{z_j}^2}. \quad (2.7)$$

Если принять

$$\sigma_{u_i}^2 = \text{const} = \sigma_u^2; \sigma_{z_j}^2 = \text{const} = \sigma_z^2; \alpha = \frac{\sigma_{u_i}^2}{\sigma_z^2}, \quad (2.8)$$

то локальная целевая функция будет иметь вид

$$\min Q(z_1, \dots, z_j, \dots, z_n) = (\Delta u_i - \sum_1^n z_j)^2 + \alpha \sum_1^n z_j^2. \quad (2.9)$$

что по форме соответствует регуляризирующему функционалу А.Н. Тихонова [Тихонов, 1979], т.е. в анализируемом методе имеет место локальная регуляризация, и параметры регуляризации для каждого неизвестного и в каждом уравнении определяются величинами $\sigma_{z_j}^2$ и $\sigma_{u_i}^2$.

Дифференцируя (2.7) по каждому из неизвестных, получим систему из n уравнений с n неизвестными [Кочнев, 1983]. Система имеет следующее аналитическое решение

$$z_j = \Delta u_i \cdot \frac{\sigma_{z_j}^2}{\sigma_{u_i}^2 + \sum_{j=1}^n \sigma_{z_j}^2}. \quad (2.10)$$

В знаменатель полученного выражения входит сумма дисперсий неизвестных и дисперсии

правой части. При $\sigma_{u_i}^2 \gg \sigma_{z_j}^2$ неизвестные практически уточняться не будут, т.е. невязка будет полностью отнесена к ошибке наблюдения.

Учитывая обозначения в (2.5) и вводя k – номер очередного шага уточнения, получим окончательно формулы для уточнения значений параметров

$$x_{j(k+1)} = \bar{x}_{j(k)} + \Delta x_{j(k+1)}, \quad (2.11)$$

где $\Delta x_{j(k+1)} = z_j / \frac{\partial u_i}{\partial x_j}$.

Принимая во внимание (2.5), (2.9), (2.10), получим

$$\alpha_{j(k+1)} = \frac{\frac{\partial u_i}{\partial x_j} \cdot \sigma_{x_j(k)}^2}{\sigma_{u_i}^2 + \sum_{j=1}^n \left(\frac{\partial u_i}{\partial x_j} \cdot \sigma_{x_j(k)}^2 \right)^2}. \quad (2.12)$$

Это и будет оптимальная величина коэффициента, следующая из локального критерия максимума правдоподобия. При $\sigma_{u_i}^2 = 0$ и при $\sigma_{x_j}^2 = 1$ получим

$$\alpha_{j(k+1)} = \frac{\frac{\partial u_i}{\partial x_j}}{\sum_{j=1}^n \left(\frac{\partial u_i}{\partial x_j} \right)^2}, \quad (2.13)$$

т.е. приходим к методу Качмажа [Цыпкин, 1968].

3. ВЫБОР КОЭФФИЦИЕНТА $\gamma_{j(k+1)}$

Рассмотрим сначала простую ситуацию, когда значение параметра X уточняется путем непосредственного многократного измерения. Для оценки среднего $\bar{x}_{(k+1)}$ применим адаптивный метод, который на $k+1$ -ом шаге после $k+1$ -го измерения $x_{(k+1)}$ дает

$$\bar{x}_{(k+1)} = \bar{x}_{(k)} + \alpha_{(k+1)}(x_{(k+1)} - \bar{x}_{(k)}) \quad (3.1)$$

Учитывая, что $\frac{\partial u}{\partial x} = 1$ и имеем одно значение, то коэффициент $\alpha_{(k+1)}$ будет равен

$$\alpha_{(k+1)} = \frac{\sigma_{x(k)}^2}{\sigma_{x(k+1)}^2 + \sigma_{x(k)}^2}. \quad (3.2)$$

Из [Крылов, 1954] известно, что при вычислении среднего для неравноточных значений

$z_1, \dots, x_i, \dots, x_m$, имеющих соответствующую дисперсию $\sigma_1^2, \dots, \sigma_i^2, \dots, \sigma_m^2$, по формуле

$$\bar{x}_{(k)} = \frac{\sum x_i \cdot P_i}{\sum P_i} \quad (3.3)$$

наименьшую дисперсию дает оценка при $P_i = 1/\sigma_i^2$.

Следовательно, рекуррентная формула для вычисления $(k+1)$ значения будет иметь вид

$$\bar{x}_{(k+1)} = \frac{\bar{x}_{(k)} \cdot \bar{P}_{(k)} + x_{(k+1)} P_{(k+1)}}{\bar{P}_{(k)} + P_{(k+1)}}. \quad (3.4)$$

Отсюда можем получить формулы (3.1), (3.2) и, кроме этого,

$$\bar{P}_{(k+1)} = \bar{P}_{(k)} + P_{(k+1)} = \frac{1}{\sigma_{(k)}^2} + \frac{1}{\sigma_{(k+1)}^2}, \quad (3.5)$$

где $\bar{P}_{(k)}$ и $\bar{P}_{(k+1)}$ – веса среднего из k и $k+1$ значений, а $\sigma_{(k)}^2$ – дисперсия среднего из k значений.

Преобразуя (3.3), с учетом (3.2), получим

$$\sigma_{(k+1)}^2 = \sigma_{(k)}^2 (1 - \alpha_{(k+1)}), \quad (3.6)$$

т.е. в данном случае

$$\gamma_{(k+1)} = 1 - \alpha_{(k+1)}. \quad (3.7)$$

В разумности полученных формул можно убедиться, приняв все наблюдения равноточными.

Тогда

$$\sigma_{(k+1)}^2 = \sigma_0^2 / (k+1), \quad (3.8)$$

$$\alpha_{(k+1)} = 1/(k+1). \quad (3.9)$$

В результате (3.1) будет иметь вид

$$\bar{x}_{(k+1)} = \bar{x}_{(k)} + \frac{1}{k+1} (x_{(k+1)} - \bar{x}_{(k)}), \quad (3.10)$$

что в точности совпадает с адаптивным алгоритмом оценки среднего [Цыпкин, 1968], дающего несмещенную оценку при $M(\xi) = 0$.

В работе [Фу, 1971] выделены две области уточнения: область, в которой значение уточняемого параметра находится вдали от истинного значения и необходимо сохранить большой шаг, чтобы приблизиться к искомому значению, и область поиска вблизи истинного

значения, когда шаг должен уменьшаться. Приведенные выше алгоритмы (3.1) и (3.4), (3.7) будут работать одинаково в обеих областях, и шаг в них снижается постепенно, независимо от удаления от точного решения. В частности, для ускорения сходимости алгоритмов типа (3.8) в [Фу, 1971] предлагается алгоритм Кестена, суть которого сводится к тому, что $\alpha_{(k)} = 1/n$, где n – число смен знака невязки.

Использование алгоритма Кестена для наших целей нерационально, т.к. он не учитывает неравноточности наблюдений. Поэтому нами предлагается модифицированный вариант алгоритма, который поддерживает $\alpha_{(k)}$ большим – при больших невязках и уменьшает его при малых. В модифицированном алгоритме уточнение параметра остается тем же ((3.1) и (3.2)), а формула (3.4) приобретает вид

$$\overline{\sigma}_{(k+1)}^2 = \overline{\sigma}_{(k)}^2 (1 - \beta_{(k+1)}), \quad (3.11)$$

$$\text{где } \beta_{(k+1)} = \frac{\overline{\sigma}_{(k)}^2}{\psi \cdot (x_{(k+1)} - \bar{x}_{(k)})^2 + \sigma_{(k+1)}^2 + \overline{\sigma}_{(k)}^2} \quad (3.12)$$

при $0 \leq \psi \leq 1$.

В частном случае при $\psi = 0$ получаем первоначальный алгоритм, а при $\psi = 1$ – новый, который учитывает величину невязки. При промежуточных значениях возникают промежуточные модификации алгоритма. В частности, возможно принять $\psi = 1/n$, где n может быть числом смен знака невязки или номером итерации уточнения параметров (для систем уравнений).

Обобщим приведенный подход для уточнения параметров, измеренных не непосредственно, а оцениваемых через функциональные связи. В этом случае оценка $\sigma_{x(k+1)}$ зависит от точности измерения, от слагаемых и коэффициентов, входящих в уравнение. Допустим измеряется \tilde{y} :

$$\tilde{y} = ax + by, \quad (3.13)$$

тогда

$$x = (u - b \cdot y)/a. \quad (3.14)$$

Полагая, что a и b известны точно, для оценки ошибки x запишем:

$$\xi_x = (\xi_u - b \cdot \xi_y)/a. \quad (3.15)$$

Дисперсия оценки x будет равна математическому ожиданию квадрата выражения правой части уравнения (3.15)

$$\sigma_x^2 = (\sigma_u^2 + b^2 \cdot \sigma_y^2 - 2\sigma_u \cdot b \cdot \sigma_y)/a^2. \quad (3.16)$$

Выражение окажется близким к 0, если $\sigma_u^2 \approx b^2 \cdot \sigma_y^2$ и $b > 0$, и будет большим, если $\sigma_u^2 \approx b^2 \cdot \sigma_y^2$ и $b < 0$. Во избежание этих крайних ситуаций третье слагаемое в скобках отбросим, т.е. будем постулировать независимость ошибки измерения и ошибки случайной составляющей второго слагаемого. С учетом сказанного получим

$$\sigma_x^2 = (\sigma_u^2 + b^2 \cdot \sigma_y^2)/a^2. \quad (3.17)$$

Подставив полученную оценку в (3.3), найдем

$$\frac{1}{\sigma_{x(k+1)}^2} = \frac{1}{\sigma_{x(k)}^2} + \frac{a^2}{(\sigma_u^2 + b^2 \cdot \sigma_{y(k)}^2)}. \quad (3.18)$$

Преобразовав которое, получим

$$\sigma_{x(k+1)}^2 = \sigma_{x(k)}^2 \left(1 - \frac{a_{(k+1)}^2 \cdot \sigma_{x(k)}^2}{\sigma_u^2 + a_{(k+1)}^2 \cdot \sigma_{x(k)}^2 + b_{(k+1)}^2 \cdot \sigma_{y(k)}^2} \right). \quad (3.19)$$

Для многомерного случая формула для j -го неизвестного в i -ом уравнении имеет вид

$$\sigma_{j(k+1)}^2 = \sigma_{j(k)}^2 \left(1 - \frac{a_{j(k+1)}^2 \cdot \sigma_{j(k)}^2}{\sigma_{u_i}^2 + \sum_{j=1}^n a_{j(k+1)}^2 \cdot \sigma_{j(k)}^2} \right). \quad (3.20)$$

или, учитывая (2.11), найдем

$$\sigma_{j(k+1)}^2 = \sigma_{j(k)}^2 (1 - a_{j(k+1)} \cdot \alpha_{j(k+1)}).$$

Здесь вывод формулы (3.17) сделан методом индукции из общих подходов метода наименьших квадратов. Ранее эта формула из (1.4) была получена Н.Б.Пивоваровой [Пивоварова, 1984].

Для модифицированного варианта формула имеет вид

$$\sigma_{j(k+1)}^2 = \sigma_{j(k)}^2 \left(1 - \frac{a_{j(k+1)}^2 \cdot \sigma_{j(k)}^2}{\psi \cdot \Delta u_i^2 + \sigma_{u_i}^2 + \sum_{j=1}^n a_{j(k+1)}^2 \cdot \sigma_{j(k)}^2} \right) =$$

$$= \sigma_{j(k)}^2 (1 - \beta_{j(k+1)}) \quad (3.21)$$

при $0 \leq \psi \leq 1$.

Таким образом, обоснован выбор $\gamma_{j(k+1)}$, которое в общем случае будет иметь вид

$$\gamma_{j(k+1)} = 1 - \beta_{j(k+1)}, \quad (3.22)$$

или, в частном случае, при $\psi = 0$

$$\gamma_{j(k+1)} = 1 - a_{j(k+1)} \cdot \alpha_{j(k+1)}. \quad (3.23)$$

Для нелинейного случая следует иметь в виду, что

$$a_{j(k+1)} = \partial u_i / \partial x_j \quad (3.24)$$

при $x_{1(k)}, \dots, x_{j(k)}, \dots, x_{n(k)}$.

Таким образом, в соответствии с поставленной задачей обоснован алгоритм итерационного уточнения параметров в одном уравнении. Он включает формулы: (1.1), (1.2), (2.12), (3.21), (3.22). При переходе к следующему уравнению (от $k+1$ к $k+2$) в качестве априорных будут использоваться уточненные значения параметров и их дисперсии. Так будет продолжаться до тех пор, пока не будут исчерпаны все уравнения системы. В связи с тем, что при выводе формул постулировалась независимость уточненных компонент, возникает необходимость повторять уточнение по всей системе до тех пор, пока они не будут близки к решению, что будет определено выполнением условия (1.5). В предельном случае, когда $x_{(l)} = x^*$, из (1.4) следует

$$D_{(l)} = \left(\sum_{i=1}^n \xi_i^2; P_i \right) \sum_{i=1}^n P_i. \quad (3.25)$$

т.е. $D_{(l)}$ будет характеризовать средневесовую дисперсию помехи, которая от номера итерации не зависит. Следовательно будет выполнено условие (1.5).

Созданный итерационный метод по классификации Фадеева Ф.А. и Фадеевой В.И. является

нестационарным (циклическим), т.к. матрица $H^{(l)}$ в уравнении

$$x^{(l)} = x^{(l-1)} + H^{(l)}(u - Ax^{(l-1)}) \quad (3.26)$$

существенно зависит от номера итерации. Поэтому доказательство общей сходимости построить не удастся. Да это и не имеет особого смысла, т.к. в реальных задачах при огромных матрицах исследовать сходимость метода для конкретной системы задача более трудоемкая, чем ее решение и анализ на сходимость по уменьшению невязки. Поэтому в дальнейшем для анализа сходимости будем использовать подходы, разработанные в методах стохастической аппроксимации [Фу, 1971], [Цыпкин, 1968], [Уайлд, 1967]. В методах стохастической аппроксимации число независимых наблюдений предполагается неограниченным и за счет этого, при определенных ограничениях на свойства помех, возможно получение оценок со сколь угодно высокой точностью. В наших условиях мы вынуждены компенсировать недостаток независимых наблюдений многократным использованием одной и той же выборки с тем, чтобы максимально ослабить влияние априорной информации.

4. ИССЛЕДОВАНИЕ СХОДИМОСТИ В ПРОСТОМ ОДНОМЕРНОМ СЛУЧАЕ

А. Дворецкий предложил рассматривать любую процедуру стохастической аппроксимации как обычный детерминированный метод, но с наложением на него случайной составляющей [Уайлд, 1967]. Следуя этому подходу, исследуем адаптивный алгоритм оценки среднего. Формулу (3.1) для k -го измерения можно представить в следующем виде:

$$\bar{x}_k = x^* - \xi_{k-1} + \alpha_k(x^* - \xi_k - x^* - \bar{\xi}_{k-1}) \quad (4.1)$$

где x^* – точное, но неизвестное нам значение; $\bar{\xi}_{k-1}$ – ошибка среднего значения на $k-1$ шаге; ξ_k – ошибка k -ого измерения.

Из (4.1) и учитывая, что $\xi_{k-1} = \sum_{i=1}^{k-1} \alpha_i \xi_i$, получим

$$\bar{x}_k = x^* + (1 - \alpha_k) \sum_{i=1}^{k-1} \alpha_i \xi_i + \alpha_k \xi_k \quad (4.2)$$

Отсюда видно, что случайная составляющая сосредоточена в последних двух слагаемых.

Обозначим каждое из элементарных слагаемых через r_i . Проиллюстрируем на этом примере условия сходимости, доказанные А. Дворецким [Уайлд, 1967]. Им было показано, что для сходимости процедур стохастической аппроксимации в среднеквадратическом и с вероятностью единица требуется выполнение двух условий:

1) – помеховая составляющая r_k должна быть несмещенной

$$M[r_i] = 0 \quad (4.3)$$

2) – сумма дисперсий случайной составляющей должна быть конечной при любой возможной бесконечной процедуре поиска

$$\sum_{i=1}^{\infty} M[r_i^2] < \infty \quad (4.4)$$

Для выполнения первого условия необходимо, чтобы слагаемые 2 и 3 выражения (4.2) стремились к 0. Третье слагаемое $\alpha_k \cdot \xi_k$ будет стремиться к нулю, если ξ_k – конечно, а α_k – стремится к 0. Второе слагаемое будет стремиться к нулю, если будет выполнено условие $M[\xi_i] = 0$. При невыполнении этого условия оценка среднего окажется смещенной. Для анализа второго условия, следуя Уайлду [Уайлд, 1967], будем считать, что суммарная ошибка, внесенная всеми помехами, равна $\sum_{i=1}^{\infty} r_i$. Если помехи образуют последовательность независимых случайных величин, то дисперсия суммарной ошибки равна математическому ожиданию суммы квадратов отдельных ошибок;

$$D(r_i) = M \left[\sum_{i=1}^{\infty} r_i \right]^2 = \sum_{i=1}^{\infty} r_i^2. \quad (4.5)$$

С учетом сказанного, второе условие сходимости для рассматриваемого случая будет иметь вид:

$$D(r_k) = (1 - \alpha)^2 \cdot \sigma^2 \cdot \sum_{i=1}^{k-1} \alpha_i^2 + \alpha_k^2 \cdot \sigma^2 < \infty \quad (4.6)$$

где σ^2 – оценка дисперсии помехи. Если $k \rightarrow \infty$ и α_k стремится к 0, то второе слагаемое конечно и стремится к 0.

Для анализа первого слагаемого при $0 \leq \alpha_i \leq 1$ можно записать неравенство:

$$(1 - \alpha_i)^2 \cdot \sigma^2 \sum_{i=1}^{k-1} \alpha_i^2 \leq \sigma^2 \sum_{i=1}^{k-1} \alpha_i^2. \quad (4.7)$$

Правая часть неравенства будет конечной в том и только в том случае, если

$$\sum_{i=1}^{\infty} \alpha_i^2 < \infty. \quad (4.8)$$

Таким образом, для стохастической процедуры уточнения параметров необходимо выполнение трех условий;

$$M[\xi_i] = 0, \quad (4.9)$$

$$\alpha_k \rightarrow 0 \quad \text{при} \quad k \rightarrow \infty, \quad (4.10)$$

$$\sum_{i=1}^{\infty} \alpha_i^2 < \infty. \quad (4.11)$$

Из необходимости поиска правильного решения при сколь угодно удаленных начальных значениях возникает еще одно условие сходимости

$$\sum_{i=1}^{\infty} \alpha_i = \infty. \quad (4.12)$$

Как известно [Уайлд, 1967], таким условиям, в частности, удовлетворяет гармонический ряд

$$\alpha_k = 1/k.$$

При $k \rightarrow \infty$, $\alpha_k \rightarrow 0$, $S_1 = \sum_{k=1}^{\infty} \alpha_k \rightarrow \infty$, а $S_2 = \sum_{k=1}^{\infty} \alpha_k^2$ – конечна.

При работе с ЭВМ физически нельзя требовать бесконечности, т.к. количество итераций уточнения всегда конечно. В этом отношении интересно знать, чему же равны суммы S_1 и S_2 при конечных k для гармонического ряда. В табл. 4.1 приведены значения этих сумм при различных k . Как видим, первая сумма растет очень медленно и практически далека от бесконечности, а сумма квадратов стабилизируется уже при первых 100 значениях.

5. ИССЛЕДОВАНИЕ УСЛОВИЙ СХОДИМОСТИ

ДЛЯ ОБЩЕГО ОДНОМЕРНОГО СЛУЧАЯ

Введем уравнение, связывающее точное, но неизвестное значение x^* и его оценку \bar{x}_k на k -ом шаге

$$\bar{x}_k + \Delta x_k = x^* \quad (5.1)$$

функцию наблюдения

$$\tilde{u}_k = a_k x^* + \xi_k \quad (5.2)$$

и прогнозное значение

$$\bar{u}_k = a_k \cdot \bar{x}_k, \quad (5.3)$$

где ξ – случайная ошибка наблюдения в k -ом эксперименте (шаге).

Используя адаптивный метод, получим:

$$\begin{aligned} \bar{x}_k &= \bar{x}_{k-1} + \alpha_k (\tilde{u}_k - \bar{u}_k) = \\ &= x^* - \Delta x_{k-1} + \alpha_k (u^* + \xi_k - (a_k x^* - a_k \Delta x_{k-1})) = \\ &= x^* - \Delta x_{k-1} + \alpha_k (\xi_k + a_k \Delta x_{k-1}); \end{aligned}$$

или $\Delta x_k = (1 - \alpha_k a_k) \Delta x_{k-1} - \alpha_k \xi_k$.

Учитывая, что

$$\Delta x_{k-1} = (1 - \alpha_{k-1} a_{k-1}) \Delta x_{k-2} - \alpha_{k-1} \xi_{k-1}, \quad (5.4)$$

получим

$$\Delta x_k = (1 - \alpha_k a_k)(1 - \alpha_{k-1} a_{k-1}) \Delta x_{k-2} - (1 - \alpha_k a_k) \alpha_{k-1} \xi_{k-1} - \alpha_k \xi_k$$

или

$$\begin{aligned} \Delta x_k &= \Delta x_0 \prod_{i=1}^k (1 - \alpha_i a_i) - \\ &\left[\alpha_1 \xi_1 \prod_{i=2}^k (1 - \alpha_i a_i) + \dots + \alpha_l \xi_l \prod_{i=l+1}^k (1 - \alpha_i a_i) + \dots + \alpha_k \xi_k \right]. \end{aligned} \quad (5.5)$$

где $l < k$, Δx_0 – отклонение априорного значения x_0 от точного. Значение \bar{x}_k будет сходиться к точному значению, если Δx_k будет стремиться к 0, для чего необходимо, чтобы каждое из слагаемых стремилось к 0.

Прежде всего рассмотрим выражение, стоящее под знаком произведения. Из (3.20) для i -го шага это

$$\gamma_i = 1 - \alpha_i a_i. \quad (5.6)$$

Оно изменяется в пределах от 0 до 1,

$$\gamma_i = \frac{\sigma_{u_i}^2}{\sigma_{u_i}^2 + a_i^2 \sigma_{(i-1)}^2}. \quad (5.7)$$

Из (5.5) видим, что чем γ_i ближе к нулю, тем меньше Δx_k и, следовательно, больше сходимость на i -ом шаге. Приняв в качестве эталона γ_i , у которого $a_i = 1$, и сравнивая его с $\gamma_{(a>1)}$ и $\gamma_{(a<1)}$, можем утверждать, что при одинаковых $\sigma_{u_i}^2$ и $\sigma_{u_{i-1}}^2$

$$\gamma_{(a>1)} < \gamma_{(a=1)} < \gamma_{(a<1)}. \quad (5.8)$$

Взяв за основу условия сходимости (4.9)–(4.12), доказанные для $\gamma_{(a=1)}$, можно высказать суждения о влиянии коэффициента a на скорость сходимости. Чем больше a_i по сравнению с 1, тем больше сходимость на i -ом шаге и наоборот.

Поведение γ_i и их произведений для частного случая $\sigma_{u_i}^2 = \sigma_1^2 = \text{const}$ для различных a приведены в таблице 5.1.

6. ИССЛЕДОВАНИЕ СХОДИМОСТИ ДЛЯ МНОГОМЕРНОГО СЛУЧАЯ

НА ОДНОМ ШАГЕ

При анализе сходимости будем рассматривать линейный случай, но это не сужает общности. Дело в том, что вместо линейных коэффициентов $a_{j(k)}$ в нелинейном случае будут $\frac{\partial u_{(k)}}{\partial x_j}$ (при $x_{1(k-1)} \dots x_{j(k-1)} \dots x_{n(k-1)}$). При повторной итерации при выходе на то же нелинейное уравнение будем получать другие коэффициенты, но сути это не меняет, т.к. ниже не оговаривается постоянство коэффициентов $a_{j(k)}$. По мере приближения к решению коэффициенты становятся более стабильными. Есть еще один нюанс, связанный с отбрасыванием последующих членов разложения остатка Δu в ряд Тейлора. При вводе членов второго порядка мы вынуждены были бы разбрасывать невязку на все слагаемые ряда, но т.к. этого не делается, то большая доля невязки идет на члены с первыми производными, т.е. это можно представить себе как некоторую релаксацию, которая, как известно, [Ортега, 1975],

увеличивает скорость сходимости.

В каждом конкретном нелинейном уравнении может быть верхняя или нижняя релаксация со своим коэффициентом релаксации.

Запишем уравнение, взятое из системы на k -ом шаге

$$a_k x + b_k y = u_k. \quad (6.1)$$

Прогнозное значение \bar{u}_k будет зависеть от неизвестных, вычисленных на $k-1$ шаге, т.е.

$$a_k \bar{x}_{k-1} + b_k \bar{y}_{k-1} = \bar{u}_k. \quad (6.2)$$

Пусть

$$\tilde{u}_k = u_k^* + \xi_k; \quad \bar{x}_k + \Delta x_k = x^*; \quad \bar{y}_k + \Delta y_k = y^*, \quad (6.3)$$

где ξ_k – ошибка правой части k -го уравнения; x^*, y^*, u^* – точные, но неизвестные значения параметров; \bar{x}_k и \bar{y}_k – оценки значения параметров x_k и y_k на k -ом шаге; $\Delta x_k, \Delta y_k$ – отклонения оценок от точных значений.

Отсюда получим следующее выражение для невязки Δu_k

$$\Delta u_k = \tilde{u}_k - \bar{u}_k = \xi_k + a_k \Delta x_{k-1} + b_k \Delta y_{k-1}. \quad (6.4)$$

Следуя адаптивному методу, получим следующие выражения для оценок \bar{x}_k и \bar{y}_k на k -ом шаге

$$\bar{x}_k = \bar{x}_{k-1} + \alpha_k \Delta u_k, \quad (6.5)$$

$$\bar{y}_k = \bar{y}_{k-1} + \beta_k \Delta u_k, \quad (6.6)$$

или

$$\Delta x_k = \Delta x_{k-1} (1 - \alpha_k a_k) - \alpha_k b_k \Delta y_{k-1} - \alpha_k \xi_k, \quad (6.7)$$

$$\Delta y_k = \Delta y_{k-1} (1 - \beta_k b_k) - \beta_k a_k \Delta x_{k-1} - \beta_k \xi_k, \quad (6.8)$$

где

$$\alpha_k = \frac{a_k \sigma_{x(k-1)}^2}{\sigma_{u(k)}^2 + a_k^2 \sigma_{x(k-1)}^2 + b_k^2 \sigma_{y(k-1)}^2}, \quad (6.9)$$

$$\beta_k = \frac{b_k^2 \sigma_{y(k-1)}^2}{\sigma_{u(k)}^2 + a_k^2 \sigma_{x(k-1)}^2 + b_k^2 \sigma_{y(k-1)}^2}, \quad (6.10)$$

где $\sigma_{u(k)}^2$ – оценка дисперсии ошибки ξ_k ; $\sigma_{x(k-1)}^2, \sigma_{y(k-1)}^2$ – условные дисперсии оценок \bar{x}_{k-1} и \bar{y}_{k-1} , полученные на $k-1$ шаге.

В дальнейшем будем анализировать условия сходимости одного неизвестного, например, x .

Оценка на k -ом шаге x_k будет стремиться к точному значению x^* , если Δx_k , уменьшаясь по абсолютной величине, стремится к 0.

$$\Delta x_k = (1 - \alpha_k a_k) \Delta x_{k-1} - \alpha_k (b_k \Delta y_{k-1} - \xi_k). \quad (6.11)$$

В многомерном случае

$$\Delta x_{1(k)} = (1 - \alpha_{1(k)} a_{1(k)}) \Delta x_{1(k-1)} - \alpha_{1(k)} \left(\sum_{j=2}^2 a_{j(k)} \Delta x_{j(k-1)} + \xi_k \right). \quad (6.12)$$

Если остатки неизвестных Δx_j – случайны, то выражение, стоящее во второй скобке так же является случайным. Обозначив его φ_k , получим

$$\varphi_k = \sum_{j=2}^2 a_{j(k)} \Delta x_{j(k-1)} + \xi_k. \quad (6.13)$$

Условная дисперсия значения φ_k может быть оценена, т.к. нам известны условные оценки дисперсий каждого из слагаемых. Полагая, что случайные слагаемые, входящие в (13), независимы, получим

$$\sigma_{\varphi(k)}^2 = \sum_{j=2}^2 a_j^2 \sigma_{x_{j(k-1)}}^2 + \sigma_{u(k)}^2. \quad (6.14)$$

Переписав выражение (12) с учетом (13), получим

$$\Delta x_{1(k)} = (1 - \alpha_{1(k)} a_{1(k)}) \Delta x_{1(k-1)} - \alpha_{1(k)} \varphi_k. \quad (6.15)$$

С учетом (14) формула для $\alpha_{1(k)}$ в многомерном варианте примет вид

$$\alpha_{1(k)} = \frac{a_{1(k)} \sigma_{x1(k-1)}^2}{\sigma_{y(k)}^2 + a_{1(k)}^2 \sigma_{x1(k-1)}^2}. \quad (6.16)$$

При переходе от $(k-1)$ к k -му шагу остаток $\Delta x_{1(k)}$ будет уменьшаться по абсолютной величине, если

$$0 \leq (1 - \alpha_{1(k)} a_{1(k)}) < 1, \quad (6.17)$$

$$\text{и } \alpha_{1(k)} \varphi_k \rightarrow 0 \text{ при } k \rightarrow \infty. \quad (6.18)$$

Условие (17) выполняется, т.к. $0 < \alpha_{1(k)} a_{1(k)} \leq 1$. Причем Δx_k будет приближаться к 0 тем быстрее, чем больше $a_{1(k)}$ и $\sigma_{1(k-1)}^2$. При малых $a_{1(k)}$, по сравнению с другими коэффициентами и при малых $\sigma_{1(k-1)}^2, \alpha_k \rightarrow 0$, остаток будет оставаться практически одним и тем же. Это свойство можно использовать тогда, когда какой-то параметр известен точно. Задав в качестве априорного точное значение и соответствующую ему малую σ_{x_1} , можно достичь того, что это значение параметра практически не будет изменяться. Наличие этой ситуации будет легко распознавать по результатам счета, т.к. $\sigma_{j(k)}^2$ будет в этом случае равно $\sigma_{j(0)}^2$, ибо

$$\sigma_{j(k)}^2 = \sigma_{j(k-1)}^2 (1 - a_k \alpha_k). \quad (6.19)$$

Условие (18) будет выполняться, если $M[\varphi_k] = 0$ и $\alpha \rightarrow 0$ при $k \rightarrow \infty$, т.е. имеют место условия, следующие из теоремы Дворецкого.

7. ИССЛЕДОВАНИЕ УСЛОВИЙ СХОДИМОСТИ НА ПРИМЕРЕ

СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ С ДВУМЯ НЕИЗВЕСТНЫМИ

Для исследования сходимости оценок неизвестных не отдельного параметра, а всего вектора в соответствии с [Уайлд, 1967] введем расстояние r от точки с координатами \bar{x}_k, \bar{y}_k до точки x^*, y^* и запишем

$$r_k^2 = (\bar{x}_k - x^*)^2 + (\bar{y}_k - y^*)^2. \quad (7.1)$$

При переходе от $k-1$ к k -ому шагу условие сходимости будет иметь вид

$$r_k^2 - r_{k-1}^2 < 0. \quad (7.2)$$

Пусть для двух неизвестных в линейной системе

$$\bar{x}_k = x^* - \Delta x_k; \bar{y}_k = y^* - \Delta y_k; \quad (7.3)$$

$$u_k^* = a_k x^* + b_k y^*; \tilde{u}_k = u_k^* + \xi_k.$$

В соответствии с адаптивным методом вычислим значение \bar{u}_k используя оценки неизвестных, полученные на предыдущем шаге

$$\bar{u}_k = a_k \bar{x}_{k-1} + b_k \bar{y}_{k-1}. \quad (7.4)$$

Найдем невязку

$$\Delta u_k = \tilde{u} - \bar{u}_k \quad (7.5)$$

и оценку каждого из неизвестных

$$\begin{aligned} \bar{x}_k &= \bar{x}_{k-1} + \alpha_k \Delta u_k, \\ \bar{y}_k &= \bar{y}_{k-1} + \beta_k \Delta u_k, \end{aligned} \quad (7.6)$$

Выражения для α_k и β_k даны в (6.9) и (6.10). Развернем выражения для r_k^2 и r_{k-1}^2

$$r_k^2 = (x^* - \bar{x}_{k-1} - \alpha_k \Delta u_k)^2 + (y^* - \bar{y}_{k-1} - \beta_k \Delta u_k)^2, \quad (7.7)$$

$$r_{k-1}^2 = (x^* - \bar{x}_{k-1})^2 + (y^* - \bar{y}_{k-1})^2. \quad (7.8)$$

Найдем разность между (7.7) и (7.8) и, преобразовав ее, получим:

$$\Delta r_k^2 = -(x^* - \bar{x}_{k-1})\alpha_k \Delta u_k - (y^* - \bar{y}_{k-1})\beta_k \Delta u_k + \frac{1}{2}(\alpha_k^2 + \beta_k^2)\Delta u_k^2. \quad (7.9)$$

Поделив на Δu_k^2 , поменяв знак и преобразовав, получим новую безразмерную величину, которую обозначим W'_k

$$W'_k = \frac{\Delta r_k^2}{\Delta u_k^2} = \frac{\alpha_k(x^* - \bar{x}_{k-1}) + \beta_k(y^* - \bar{y}_{k-1})}{\xi_k + a_k(x^* - \bar{x}_{k-1}) + b_k(y^* - \bar{y}_{k-1})} - \frac{1}{2}(\alpha_k^2 + \beta_k^2). \quad (7.10)$$

С учетом смены знака условие сходимости будет иметь вид:

$$W'_k > 0. \quad (7.11)$$

Подставив в (7.10) значения α_k и β_k и преобразовав, получим

$$\begin{aligned} W_k &= W'_k \cdot (\sigma_{u(k)}^2 + a_k^2 \sigma_{x(k-1)}^2 + b_k^2 \sigma_{y(k-1)}^2) = \\ &= \frac{a_k \sigma_{x(k-1)}^2 \Delta x_{(k-1)} + b_k \sigma_{y(k-1)}^2 \Delta y_{(k-1)}}{\xi_k + a_k \Delta x_{(k-1)} + b_k \Delta y_{(k-1)}} - \end{aligned}$$

$$-\frac{1}{2} \frac{a_k^2 \sigma_{x(k-1)}^4 + b_k^2 \sigma_{y(k-1)}^4}{\sigma_{u(k)}^2 + a_k^2 \sigma_{x(k-1)}^2 + b_k^2 \sigma_{y(k-1)}^2}. \quad (7.12)$$

Условие сходимости имеет вид

$$W_k > 0. \quad (7.13)$$

Сходимость будет зависеть от соотношения параметров, связанных с x и y . Для удобства анализа введем следующие соотношения:

$$b = \gamma \cdot a; \quad \sigma_y^2 = \eta^2 \cdot \sigma_x^2; \quad (y^* - \bar{y}_k) = \theta(x^* - \bar{x}_k). \quad (7.14)$$

Подставив их в (12), получим

$$W_k = \frac{a_k \Delta x_k (1 + \gamma_k \eta_k^2 \theta_k)}{\xi_k + a_k \Delta x_k (1 + \gamma_k \theta_k)} - \frac{1}{2} \frac{(1 + \gamma_k^2 \eta_k^4)}{1 + \gamma_k^2 \eta^2} > 0. \quad (7.15)$$

Выражение (7.15) определяет условия сходимости к точному решению на k -ом шаге.

Рассмотрим частный случай:

Если $\xi_k = 0$, то будем иметь

$$W = \frac{1 + \gamma_k \eta_k^2 \theta_k}{1 + \gamma_k \theta_k} - \frac{1}{2} \frac{1 + \gamma_k^2 \eta_k^4}{1 + \gamma_k^2 \eta^2} > 0. \quad (7.16)$$

В этом случае условия сходимости будут зависеть от соотношений параметров. При $\eta = 1$ имеем $W = 1/2$, т.е. имеет место сходимость независимо от величины других параметров. Этот частный случай соответствует методу Качмажа.

Выводы по теоретическому исследованию условий сходимости метода.

1. Помеховая составляющая наблюдений должна быть несмещенной (условие 4.3). Если в реальной задаче присутствует неслучайная компонента, искажающая наблюдение, она должна быть введена в модель наблюдения, как это делается при решении обратной кинематической задачи [Кочнев, 1989] или поправки вводятся в данные наблюдений, как это делается в гравиметрии или магнитометрии.

2. Помеховая составляющая должна быть конечной (условие 4.4). В реальной ситуации это означает отсутствие «ураганных» помех. Конечно, такие наблюдения необходимо выявлять и

исключать.

3. Адаптивный метод удовлетворяет условиям сходимости по Дворецкому (4.10, 4.11, 4.12).
4. В адаптивном методе в большей степени уточняются те неизвестные, у которых наибольшие коэффициенты и наоборот.
5. Задание погрешностей позволяет влиять на уточнение отдельных параметров.
6. Показателем сходимости является средняя квадратическая невязка по итерациям.

Перейдем к рассмотрению возможностей метода при решении модельных и реальных задач магнитометрии.

8. ПРИМЕРЫ, ИЛЛЮСТРИРУЮЩИЕ ВОЗМОЖНОСТИ МЕТОДА

Приведем простые задачи, допускающие проверку и наглядное представление результата.

Пример 1. От начального приближения $x_0 = 0.5$; $y_0 = 3$, используя различные итерационные методы, найти решение следующей системы

$$\begin{aligned}x - y &= -1, \\ -0.5x + 2y &= 2\end{aligned}$$

Точное решение системы $x^* = 0$; $y^* = 1$.

Отклонение $\Delta x = |x_0 - x^*| = 0.5$; $\Delta y = |y_0 - y^*| = 2$, т.е. из (7.14) $\theta = 4$. Рассмотрим сначала наиболее благоприятные условия, положив $\sigma_{x_0} = 0.5$; $\sigma_{y_0} = 2$; (т.е. $\eta = 4$ и $\theta/\eta = 1$).

Используя адаптивный метод, за 4 итерации приближаемся к точному решению (табл. 8.1 и рис. 2). Этот же пример решим, положив $\sigma_{x_0} = \sigma_{y_0} = 1$ ($\eta = 1$ и $\theta/\eta = 4$). Как видно из рис. 1 такое же приближение удастся получить за 5 итераций.

И, наконец, предположим, что априорная информация о достоверности неверна, приняв $\sigma_{x_0} = 2$; $\sigma_{y_0} = 0.5$ ($\eta = 0.25$; $\theta/\eta = 16$). На рис. 2 видим, что для получения того же результата необходимо на две итерации больше, чем в первом варианте и на одну по сравнению со вторым.

Для сопоставления на рис. 2 приведена траектория результатов, полученных с помощью метода Качмажа. В данном примере большая скорость сходимости оказалась у адаптивного

метода (при первом и втором варианте АСКО), а медленная – у метода Качмажа.

Пример 2. Начиная от точки $x_0 = y_0 = 2$, найдем решение системы численными итерационными методами

$$\begin{aligned}x - 5y &= 0, \\ x + 5y &= 0.\end{aligned}$$

Следует заметить, что решить эту систему методами покоординатного спуска без поворота системы координат нельзя. Они не дают решения, т.к. не сходятся. Метод Качмажа (рис. 3) сходится на каждом шаге, но по мере приближения к решению скорость сходимости постепенно уменьшается. Что касается адаптивного метода (рис. 3 и табл. 8.2), то его траектория становится более “целеустремленной” по мере приближения к решению. Это объясняется тем, что из-за больших коэффициентов перед y , σ_y^2 сильно уменьшается и невязка в большей степени идет на уточнение координаты x .

Пример 3. Начиная от точки $x_0 = 0.5$; $y_0 = 3$ с использованием адаптивного метода найдем решение несовместной системы трех уравнений.

$$1) \quad x - y = -1 \quad 2) \quad -0.5 + 2y = 2 \quad 3) \quad 0.333x + y = 2.167,$$

имеющих различные погрешности правых частей

$$\sigma_{u_1} = \sigma_{u_2} = 0.001; \quad \sigma_{u_3} = 0.5.$$

Из рис. 4 и табл. 8.3 видно, что траектория решения приближается к точке пересечения 1 и 2 линий, а уравнение 3 оказывает влияние только при первом (с его участием) уточнении (отрезок от 2 к 3 шагу), пока и являются большими.

Пример 4. Используя адаптивный метод, найдем решение системы из примера 3, приняв

$$\sigma_{u_1} = \sigma_{u_2} = \sigma_{u_3} = 0.5.$$

Как видно из рис. 5 и табл. 8.4, решение постепенно приближается к центру треугольника.

Пример 5. Используя адаптивный метод, решим нелинейную систему

$$\sqrt{t^2 + \frac{0.01}{v^2}} = 1.005,$$

$$\sqrt{t^2 + \frac{1}{v^2}} = 1.118,$$

при следующих начальных данных:

$$t_0 = 0.95; \quad v_0 = 2.2;$$

$$\sigma_{t_0} = 0.050; \quad \sigma_{v_0} = 0.2; \quad \sigma_u = 0.001.$$

Точное решение: $t_0 = 1, \quad v = 2$.

Результаты уточнения для $\psi = 0$ приведены в табл. 8.5.

Пример 6. (Взято из [Уайлд, 1967, с.263]). От начального приближения $x = 10$ найти корень для функции

$$y = x^3 + x + \xi_k,$$

где ξ_k – случайное число, которое предлагается выбирать $\pm C$, бросая монету. Не меняя сути задания, в качестве помехи используем гармоническую функцию $\xi_k = C \cdot \sin(0.5\pi k)$.

На этом примере сопоставим три метода: Ньютона, стохастической аппроксимации и адаптивный. При отсутствии помех, т.е. при $C = 0$, метод Ньютона и адаптивный при решении этой задачи практически совпадают. Отличие результатов имеет место при наличии помех, что видим из таблицы 8.6, в которой сведены результаты поиска корня при $C = 2$.

Из таблицы видно, что методы Ньютона и адаптивный (при двух вариантах дисперсии $\sigma_{x_0}^2 = 100$ и $\sigma_{x_0}^2 = 10$) первоначально одинаково быстро приближаются к решению, но, начиная с 7-го шага, видны отличия, которые отчетливо проявляются на 11 и 15 шагах. В результате адаптивного метода флуктуации оценок корня, вызванные помехой, постепенно уменьшаются из-за постепенного уменьшения σ_x^2 , а в результатах метода Ньютона они остаются на одинаковом уровне (± 2).

Сходимость метода стохастической аппроксимации сильно зависит от выбора шага.

Если в этом примере взять $\alpha_k = 1/k$, то процесс уточнения будет очень быстро расходиться. Уже на первом шаге будет получено $x_1 \approx -500$, а на втором $x_2 \approx 40 \cdot 10^6$. Поэтому, по рекомендации [Уайлд, 1967], возьмем $\alpha_k = 0.002/k$ и дополнительно $\alpha_k = 0.01/k$. Результаты счета сведем в таблицу 8.7. Из таблицы 8.7 видно, что результаты метода стохастической аппроксимации практически сильно зависят от первого или второго шага, а дальше они уточняются очень мало.

Приведенные выше примеры показывают, что адаптивный метод, являющийся статистическим обобщением метода Качмажа (2.12) позволяет:

- а) ускорять сходимость решения задачи не только за счет учета достоверности начальных приближений, но и за счет собственных свойств алгоритма (примеры 1, 2);
- б) решать несовместные неравноточные системы (примеры 3, 4), т.е. системы, не имеющие единственное решение;
- в) решать нелинейные системы (примеры 5 и 6).

Метод используется для решения как линейных, так и нелинейных обратных геофизических задач с большим числом ($10^2 - 10^4$) неизвестных [Кочнев, 1985], [Кочнев, 1987], [Кочнев, 1988], [Кочнев, 1993], [Кочнев, 1996], [Кочнев, 2001], [Кочнев, 2006].

Выводы

Для решения обратных геофизических задач, характеризующихся неустойчивостью и большим числом неизвестных, предложен адаптивный итерационно-стохастический метод. Приведена постановка задачи, обоснован выбор критерия и метод уточнения параметров по невязке с одновременным изменением условных дисперсий. Исследованы условия сходимости и приведены модельные и реальные примеры, иллюстрирующие некоторые его возможности, сходимость, устойчивость к случайным помехам и способность решать задачи с большим количеством неизвестных.

Благодарности

Автор благодарит сотрудников, принимавших участие в разработке конкретных методов и

технологий решения обратных задач геофизики: Г.А.Устюжанина, А.В.Антоненко, Н.Б.Пивоварову, Т.А.Тушко, В.С.Полякова, И.В.Гоза, Д.В.Васильева, В.Ю.Сидорова, О.В.Кашафутдинова.

ЛИТЕРАТУРА

1. Гольдин С.В., Бембель Р.М. О некоторых способах регуляризации метода наименьших квадратов при обработке геофизических наблюдений. Математические методы в нефтяной геологии и геофизике / Тр. ЗапСибНИГНИ. – Вып. 36. – Тюмень, 1970. – С. 86–102.
2. Гольцман Ф.М. Статистические модели интерпретации. – М.: Наука, 1971. – 328 с.
3. Жуковский Е.А. Статистическая регуляризация алгебраических систем уравнений // Журнал вычислительной математики и матем. физики. – Т. 12. – N 1. – 1972. – С. 185–191.
4. Корн Г., Корн Т. Справочник математика для научных работников и инженеров. – М.: Наука, 1973. – 831 с.
5. Кочнев В.А. Адаптивное прослеживание отраженных волн и оценка их параметров по данным многократных систем наблюдений. // Геология и геофизика. – N 2.–1983.– С. 95–103.
6. Кочнев В.А. Адаптивный метод решения систем линейных уравнений в задачах геофизики. Применение ЭВМ в задачах управления. / Красноярск, 1985. – С. 62–71.
7. Кочнев В.А. Адаптивный метод решения некоторых обратных задач сейсморазведки. Численные модели геофизики. / Новосибирск, 1987. – С. 55–61.
8. Кочнев В.А. Адаптивные методы интерпретации сейсмических данных (монография) Наука. Сиб.отд. Новосибирск., 1988 – 152 с.
9. Кочнев В.А., Хвостенко В.И. Адаптивный метод решения обратных задач гравиметрии. // Геология и геофизика, №7, 1996, с.120-129.
10. Кочнев В.А. Путь осознания возможностей математических моделей и алгебраических уравнений в геофизике. Геофизика, №5, 2001 г
11. Кочнев В.А. Адаптивные методы решения обратных задач геофизики. ВЦ СО РАН, Красноярск, учебное пособие, 1993.
12. Кочнев В.А. Модель среды – основа для интерпретации и решения прямых и обратных

задач. Мат. 33-й сессии Межд. сем. им. Д.Г.Успенского «Вопросы теории и практики геологической интерпретации гравитационных, магнитных и электрических полей.», Екатеринбург, 2006

13. Кочнев В.А., Пивоварова Н.Б., Тушко Т.А. Исследование адаптивного метода решения обратной задачи сейсмологии. Препринт, ВЦ СО АН СССР, №19, г.Красноярск, 1984.

14. Кочнев В.А., Антоненко А.В. Рекурсивное уточнение параметров двухслойной модели среды по годографам преломленных волн // Проблемы нефти и газа Тюмени. – Вып. 60. – Тюмень, 1983. – С. 59–62.

15. Кочнев В.А., Гоз И.В. Нераскрытые возможности магнитометрии. – Геофизика, №6, 2006, с.51-55

16. Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 648 с.

17. Крылов А.Н. Лекции о приближенных вычислениях. – М.: Госиздат. технико-теоретической литературы, 1954. – 400 с.

18. Липцер Р.Ш., Ширяев А.Н. Статистика случайных процессов. – М.: Наука, 1974. – 696 с.

19. Ортега Дж., Рейнболдт В. Итерационные методы решения нелинейных систем уравнений со многими неизвестными. – М.: Мир, 1975. – 558 с.

20. Пивоварова Н.Б. Об оптимальном выборе параметров регуляризации // Модели и численные методы. – Новосибирск, 1984. – С. 23–31.

21. Растригин Л.А. Системы экстремального управления. – М.: Наука, 1974. – 632 с.

22. Страхов В.Н. Что делать? (О развитии гравиметрии и магнитометрии в России в начале XXI века). РАН, объединенный институт физики земли им. В.Ю. Шлихта. - М., 1998. - 24с.

23. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. – М.: Наука, 1979. – 288 с.

24. Федотов А.М. Линейные некорректные задачи со случайными ошибками в данных. – Новосибирск: Наука, 1982. – 189 с.

25. Фу К. Последовательные методы в распознавании образов и обучении машин. – М.: Наука, 1971. – 256 с.

26. Цыпкин Я.З. Адаптация и обучение в автоматических системах. – М.: Наука, 1968. – 400 с.

27. Уайлд Дж. Методы поиска экстремума. – М.: Мир, 1967. – 267 с.
28. Jackson D.D. The use of a priori data to resolve non-uniqueness in linear inversion // Geophys. J.H. Astron. Soc., 57, p. 137–157, 1979.
29. Tarantola A., Valette B. Generalized nonlinear inverse problems solved using the least squares criterion // Reviews of Geophysics and space Physics, vol. 20., N 2, p. 219–232, 1982.
30. Wiggins R.A. The general linear inverse problem.– Rev. Geophys.Space Phys.,10(1),251-285, 1972.
31. Wiggins R.A., Lerner R.I. Residual statics estimation as a general linear inverse problem. – Geophysics, vol. 41, N 5.

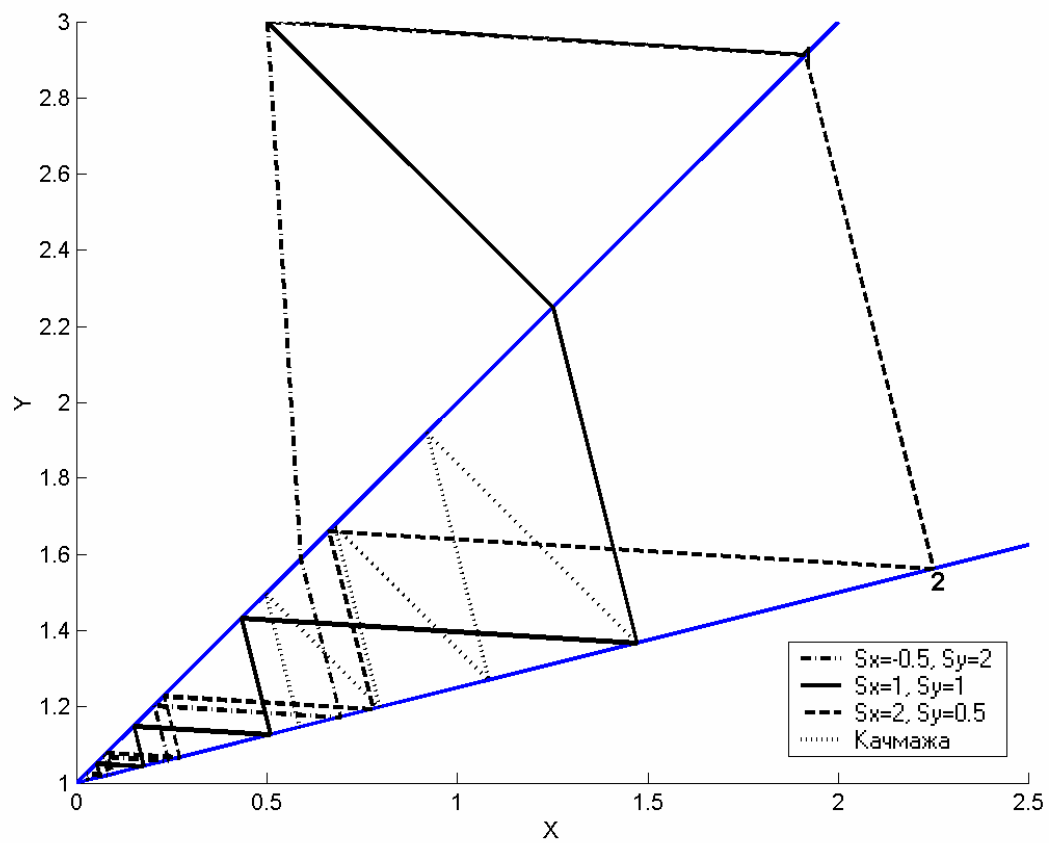


Рис. 1. Сопоставление итеративных методов решения системы линейных уравнений

$$x - y = 1; \quad -0.5x + 2y = 2$$

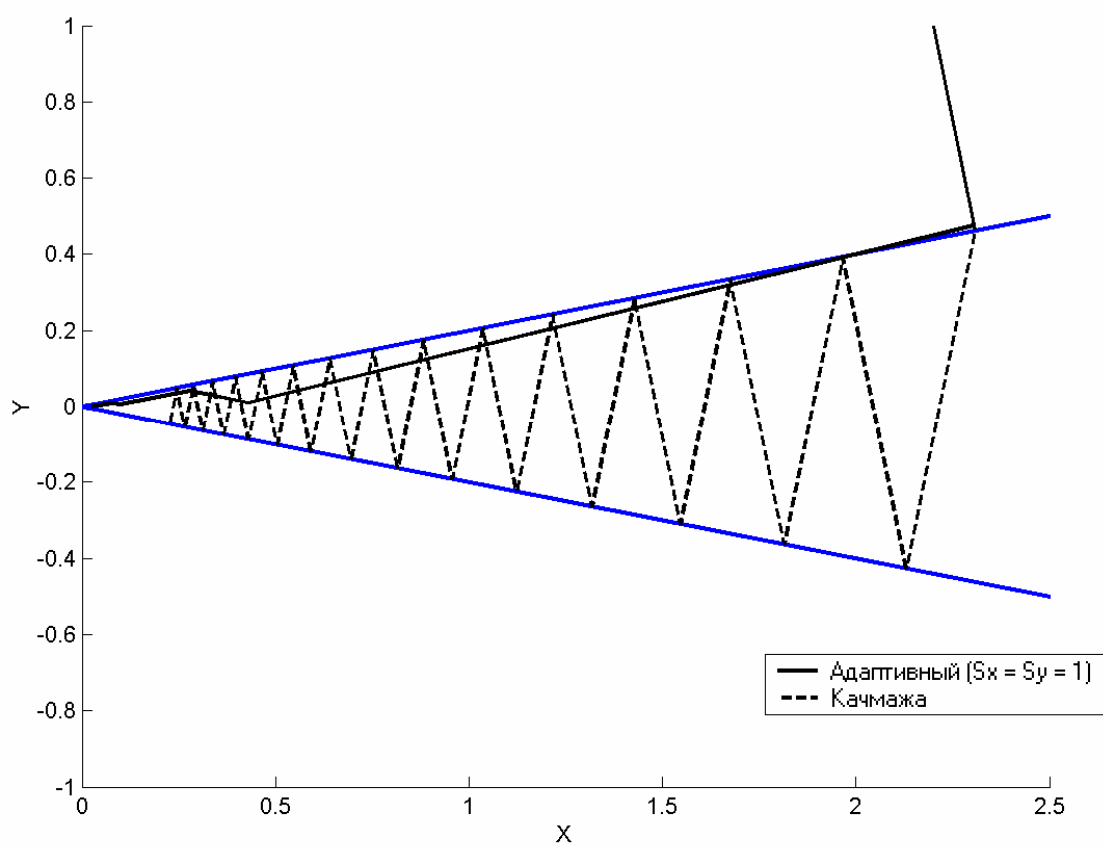


Рис. 2. Сопоставление методов при решении системы уравнений

$$x-5y=0; \quad x+5y=0$$

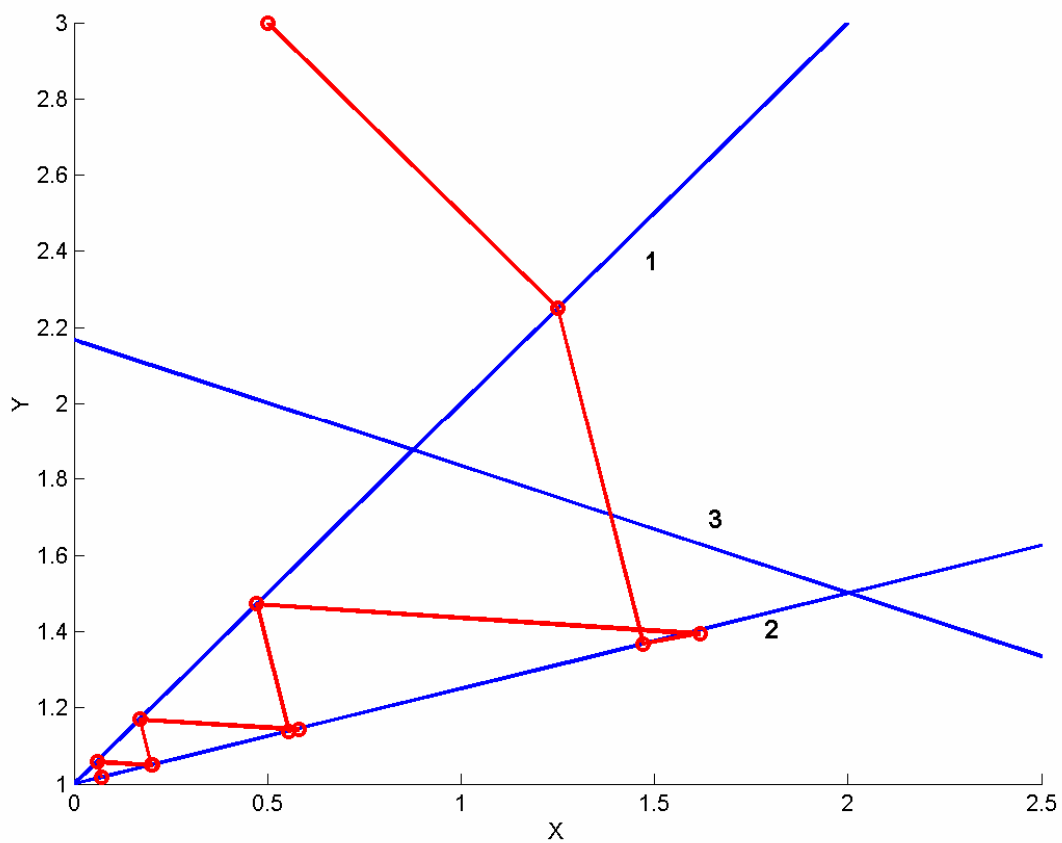


Рис. 3. Траектория поиска адаптивным методом решения неравноточной системы

$$x-y=1; -0.5x+2y=2; 0.333x+y=2.167$$

$$\sigma_{u1}=\sigma_{u2}=0.001; \sigma_{u3}=0.5$$

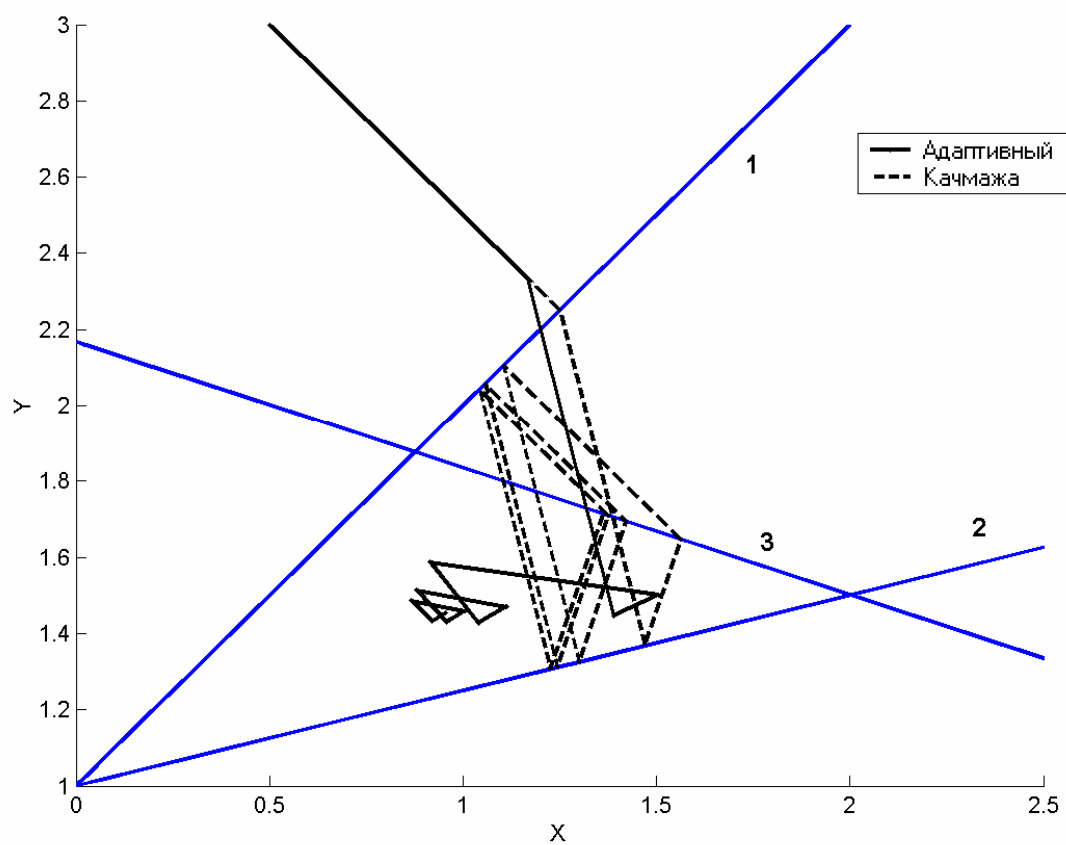


Рис. 4. Траектория решения адаптивным методом системы трех уравнений с двумя неизвестными. Уравнения те же, что и на рис.3

$$\sigma_{u1}=\sigma_{u2}=\sigma_{u3}=0.5$$

Таблица 4.1

k	10	100	1000	10000	20000	30000
S_1	2.93	5.19	7.48	9.78	10.48	10.88
S_2	1.55	1.63	1.64	1.64	1.64	1.64

Таблица 5.1

номер измерения	$a = 0.5$		$a = 1$		$a = 2$	
	γ_i	$\prod \gamma_i$	γ_i	$\prod \gamma_i$	γ_i	$\prod \gamma_i$
1	4/5	4/5	1/2	1/2	1/5	1/5
2	5/6	4/6	2/3	1/3	5/9	1/9
3	6/7	4/7	3/4	1/4	9/13	1/13
4	7/8	4/8	4/5	1/5	13/17	1/17
5	8/9	4/9	5/6	1/6	17/21	1/21

Таблица 8.1

Номер итерации	Номер уравнения	Δu_k	x_k	y_k	$\sigma_{x(k)}$	$\sigma_{y(k)}$
0		—	0.5	3	0.5	2
1	1	1.5	0.58	1.59	0.48	0.48
	2	-0.88	0.69	1.17	0.47	0.12
2	1	-0.52	0.2	1.2	0.114	0.114
	2	-0.3	0.24	1.06	0.11	0.03
3	1	-0.18	0.07	1.07	0.03	0.03
	2	-0.1	0.8	1.02	0.030	0.006
4	1	-0.06	0.02	1.02	0.006	0.006
	2	0.03	0.02	1.01	0.006	0.002

Таблица 8.2

Номер итерации	Номер уравнения	Δu_k	x_k	y_k	$\sigma_{x(k)}$	$\sigma_{y(k)}$
0			2	2	1	1
1	1	8	2.3	0.46	0.99	0.72
	2	-4.6	2.06	-0.41	0.96	0.40
2	1	-4.1	1.7	0.34	0.93	0.15
	2	-3.4	1.2	-0.25	0.88	0.17
3	1	-2.5	0.83	0.16	0.81	0.10
	2	-1.6	0.44	-0.09	0.71	0.06
4	1	-0.89	0.16	0.03	0.54	0.03
	2	-0.32	0.03	-0.05	0.33	0.014
5	1	-0.05	0.006	0.0001	0.17	0.007
	2	0.003	0.0002	0.00003	0.086	0.003

Таблица 8.3

Номер итерации	Номер уравнения	Δu_k	x_k	y_k	σ_{x_k}	σ_{y_k}
0	0	—	0.5	3	1	1
1	1	-1.5	1.25	2.25	0.7	0.7
	2	-1.87	1.47	1.36	0.68	0.17
	3	0.3	1.62	1.4	0.63	0.17
2	1	-1.22	0.47	1.47	0.16	0.16
	2	-0.71	0.56	1.14	0.15	0.04
	3	0.84	0.58	1.14	0.15	0.04
3	1	-0.43	0.17	1.17	0.037	0.037
	2	-0.25	0.2	1.05	0.036	0.009
	3	1.05	0.20	1.05	0.036	0.009
4	1	-0.15	0.07	1.06	0.008	0.009
	2	-0.17	0.08	1.02	0.008	0.002
	3	1.1	0.08	1.02	0.008	0.002

Таблица 8.4

Номер итерации	Номер уравнения	Δu_k	x_k	y_k	σ_{x_k}	σ_{y_k}
0	0	0	0.5	3	1	1
1	1	-1.5	1.167	2.33	0.74	0.74
	2	-2.0	1.38	1.45	0.72	0.28
	3	0.25	1.5	1.5	0.67	0.25
2	1	-1	0.91	1.58	0.43	0.24
	2	-0.72	1.03	1.42	0.41	0.18
	3	0.40	1.11	1.47	0.4	0.17
3	1	-0.64	0.88	1.51	0.32	0.16
	2	-0.58	0.95	1.43	0.31	0.14
	3	0.42	1.00	1.46	0.30	0.13
4	1	-0.54	0.87	1.48	0.26	0.13
	2	-0.54	0.92	1.43	0.25	0.11
	3	0.42	0.95	1.45	0.25	0.11

Таблица 8.5

Номер итерации	Номер уравнения	Δu_k	t_k	v_k	σ_{t_k}	σ_{v_k}
0	0		0.95	2.200	0.050	0.2
1	1	0.051	1.001	2.196	0.001	0.199
	2	0.016	1.008	1.983	0.001	0.018
2	1	-0.001	1.000	1.983	0.0008	0.018
	2	-0.002	1.000	1.997	0.0007	0.009

Таблица 8.6

k	Метод Ньютона		Адаптивный $\psi = 1$			Адаптивный $\psi = 1$		
	—	10	—	10	100	—	10	100
0	—	10	—	10	100	—	10	100
1	1012	6.64	1012	6.64	97	1012	6.64	9.97
2	299	4.4	299	4.39	84	299	4.39	9.82
3	87	2.91	87	2.91	51	87	2.91	9.13
4	27.6	1.87	27.6	1.87	1.87	27.8	1.87	6.9
5	10.4	0.96	10.4	0.97	6.2	10.5	0.98	4.0
6	1.85	0.47	1.87	0.49	0.99	1.9	0.51	0.92
7	-1.4	1.3	-1.38	1.00	0.62	-1.3	0.98	0.58
8	3.6	0.74	2.00	0.64	0.42	1.9	0.63	0.39
9	3.15	-0.44	2.9	0.016	0.38	2.9	0.03	0.36
10	-0.53	-0.11	1.6	0.011	0.27	0.03	0.02	0.26
11	-2.1	1.93	-1.98	0.44	0.26	-1.98	0.43	0.25
12	9.07	1.18	0.52	0.34	0.20	0.52	0.34	0.20
13	4.8	0.24	2.38	-0.03	0.19	2.38	-0.02	0.18
14	0.26	0.025	-0.03	-0.02	0.16	-0.02	-0.017	0.15
15	-1.9	2.0	-2.0	0.25	0.15	-2.0	0.25	0.15

Таблица 8.7

k	MCA $\alpha_k = 0.01/k$		MCA $\alpha_k = 0.02/k$	
	$y_k + \xi_k$	x_k	$y_k + \xi_k$	x_k
1	1012	-0.12	1012	-1012
2	-0.12	-0.12	-1084	0.6
3	-2.12	-0.11	-1.18	0.6
4	-0.11	-0.11	0.83	0.6
5	1.89	-0.11	2.8	0.59
6	-0.11	-0.11	0.8	0.59
7	-2.1	-0.11	-1.19	0.59