

АНАЛИЗ И СИНТЕЗ СИГНАЛОВ И ИЗОБРАЖЕНИЙ

УДК 621.391 : 53.08

Ю. Е. Воскобойников, Л. А. Литвинов

*(Новосибирск)*ВЫБОР МОМЕНТА ОСТАНОВА
В ИТЕРАЦИОННЫХ АЛГОРИТМАХ
ВОССТАНОВЛЕНИЯ СИГНАЛОВ И ИЗОБРАЖЕНИЙ

Для восстановления сигналов и изображений большой размерности часто применяют итерационные алгоритмы, использующие в качестве параметра регуляризации номер итерации, на которой прекращается построение минимизирующей последовательности. Выбор момента останова является важной проблемой, и для ее успешного решения требуется существенная априорная информация, отсутствующая на практике. В представленной работе на основе метода перекрестной значимости (cross-validation method) предлагается алгоритм оценивания номера итерации, на которой решение имеет минимальную ошибку.

Введение. Многие задачи обработки и интерпретации экспериментальных данных после дискретизации оператора сводятся к решению системы линейных алгебраических уравнений вида

$$K\varphi = f, \quad (1)$$

где K – матрица размера $N \times M$; φ, f – векторы соответствующей размерности. При большой размерности системы для нахождения устойчивого решения применяют итерационные алгоритмы [1–3]. При этом в качестве параметра регуляризации [4–6] используется номер итерации, на которой прекращается построение решения, т. е. момент останова итерационного алгоритма. Используемый на практике критерий невязки (или более общий – критерий обобщенной невязки) требует задания числовых характеристик, определяющих уровень погрешности задания матрицы K и вектора правой части f [1, 2]. К сожалению, во многих случаях эти характеристики бывают неизвестными, и выбор момента останова будет существенно затруднен при использовании итерационных алгоритмов на практике.

В данной работе на основе метода перекрестной значимости (cross-validation method) предлагается алгоритм оценивания номера итерации, на которой решение имеет минимальную ошибку.

Итерационные алгоритмы восстановления. Относительно системы (1) сделаем следующие предположения: а) вместо точной правой части f задан вектор $\tilde{f} = \bar{f} + \eta$, где η – случайный вектор; б) система (1) несовместна (из-за погрешности задания правой части), а матрица K плохо обусловлена или, возможно, вырождена (следствие некорректности исходной задачи). Тогда в качестве решения системы (1) принимают нормальное псевдорешение $\tilde{\varphi}^+$, т. е. вектор, имеющий минимальную норму среди всех векторов, доставляющих минимум функционалу $\Psi[\varphi] = \frac{1}{2} \|\tilde{f} - K\varphi\|^2$ (функционалу метода наименьших квадратов).

Для вычисления нормального псевдорешения (при больших размерностях векторов φ, \tilde{f}) используют итерационные алгоритмы минимизации. Приведем два итерационных алгоритма, часто используемые на практике.

Алгоритм метода наискорейшего спуска. Минимизирующая последовательность $\{\varphi^{(n)}\}$ определяется соотношением

$$\varphi^{(n+1)} = \varphi^{(n)} - \beta_n g^{(n)}, \quad (2)$$

где

$$g^{(n)} = -K^T (\tilde{f} - K\varphi^{(n)}), \quad \beta_n = \frac{\|g^{(n)}\|^2}{\|Kg^{(n)}\|^2}, \quad n = 0, 1, 2, \dots$$

Заметим, что вектор $g^{(n)}$ является градиентом функционала $\Psi[\varphi]$.

Алгоритм метода сопряжения градиентов. Минимизирующая последовательность $\{\varphi^{(n)}\}$ определяется соотношением

$$\varphi^{(n+1)} = \varphi^{(n)} - \beta_n p^{(n)}, \quad (3)$$

где

$$p^{(n)} = g^{(n)} + \gamma_n p^{(n-1)}, \quad g^{(n)} = -K^T (\tilde{f} - K\varphi^{(n)}),$$

$$\gamma_n = \frac{\|g^{(n)}\|^2}{\|g^{(n-1)}\|^2}; \quad \beta_n = \frac{\langle g^{(n)}, p^{(n)} \rangle}{\|Kp^{(n)}\|^2}; \quad n = 0, 1, 2, \dots; \quad p^{(-1)} = \bar{0}; \quad p^{(0)} = g^{(0)}$$

(здесь $\bar{0}$ – нулевой вектор). Заметим, что данный алгоритм имеет следующее свойство: минимум функционала $\Psi[\varphi]$ достигается не более чем за M итераций (при условии, что все арифметические операции на компьютере выполняются без погрешностей).

Очевидно, что для этих двух алгоритмов необходимо задать стартовую точку – вектор $\varphi^{(0)}$. Можно использовать один из следующих вариантов: 1) $\varphi^{(0)} = \bar{0}$; 2) $\varphi^{(0)} = K^T \tilde{f}$. Оба эти варианта являются гарантией того, что любой элемент $\varphi^{(n)}$ будет перпендикулярен нуль-пространству матрицы K , т. е. векторам, для которых $K\varphi = \bar{0}$. Это свойство перпендикулярности является необходимым требованием нормального псевдорешения.

Если вектор правой части f задан точно и матрица K хорошо обусловлена, то в качестве нормального псевдорешения принимается предельная точка минимизирующей последовательности $\{\varphi^{(n)}\}$. Плохая обусловленность матрицы K и погрешности задания правой части не позволяют принять предельную точку в качестве искомого решения из-за больших ошибок. В этих условиях минимизирующие процедуры (2), (3) должны обладать еще дополнительным свойством – свойством регуляризации, в качестве параметра которой выступает номер итерации останова итерационной процедуры. Почему номер итерации выполняет эту функцию? Чтобы ответить на этот вопрос, рассмотрим функционал

$$\Delta(n) = \|\varphi^{(n)} - \bar{\varphi}^+\|^2, \quad (4)$$

где $\bar{\varphi}^+$ – нормальное псевдорешение, построенное при «точной» правой части \bar{f} . После старта итерационной процедуры величина $\Delta(n)$ монотонно уменьшается (за счет уменьшения систематической ошибки), а затем при дальнейшем росте числа итераций начинает увеличиваться (за счет роста случайной ошибки, обусловленной передачей в решении $\varphi^{(n)}$ погрешностей задания правой части). Очевидно, что нужно остановить итерационный алгоритм до увеличения $\Delta(n)$ на некоторой итерации (итерации останова, номер которой обозначим n_s). При этом алгоритм выбора n_s должен гарантировать сходимость решений $\varphi^{(n_s)}$:

$$\|\varphi^{(n_s)} - \bar{\varphi}^+\| \rightarrow 0 \text{ при } \|\tilde{f} - \bar{f}\| \rightarrow 0. \quad (5)$$

Это условие является необходимым для того, чтобы итерационный алгоритм являлся регуляризирующим.

Для вычисления n_s можно использовать критерий невязки (или его обобщение) [1, 2], но необходимая для его применения априорная информация о числовых характеристиках, определяющих уровень погрешностей задания матрицы K и вектора правой части f , во многих практических случаях бывает неизвестной.

Алгоритмы оценивания номера оптимальной итерации. Итерацию, в которой достигается единственный минимум функционала $\Delta(n)$, назовем оптимальной, а номер этой итерации обозначим n_{opt} . Очевидно, что в качестве момента останова итерационной процедуры восстановления следует брать оптимальную итерацию (т. е. $n_s = n_{\text{opt}}$). Для точного вычисления n_{opt} требуется существенная априорная информация о векторах η , $\bar{\varphi}^+$, которая, к сожалению, отсутствует на практике. Поэтому попытаемся оценить n_{opt} при помощи метода перекрестной значимости [4, 7].

Этот метод достаточно давно использовался для выбора параметра регуляризации α при решении плохо обусловленных систем алгебраических уравнений. Регуляризованное решение φ_α находится из системы уравнений

$$(K^T K + \alpha I)\varphi_\alpha = K^T \tilde{f}.$$

В качестве параметра регуляризации принималось значение α_U , доставляющее минимум функционалу

$$U(\alpha) = \frac{\frac{1}{N} \|\tilde{f} - K\varphi_\alpha\|^2}{\left\{ \frac{1}{N} \text{Sp}[I - D(\alpha)] \right\}^2}, \quad (6)$$

где I – единичная матрица; $D(\alpha) = K(K^T K + \alpha I)^{-1} K^T$ – матрица размера $N \times N$; $\text{Sp}[A]$ – след матрицы A . Так как процедура одновременной минимизации требует многократного вычисления $U(\alpha)$, то присутствие в $D(\alpha)$ обратной матрицы обуславливает существенные вычислительные затраты (порядка N^3) при реализации этого метода. Изложим варианты метода перекрестной значимости для итерационных алгоритмов (2), (3), свободные от этого недостатка.

Обозначим через $e^{(n)} = \tilde{f} - K\varphi^{(n)}$ вектор невязки и введем в рассмотрение матрицу невязки $E^{(n)}$ размера $N \times N$, позволяющую выразить вектор невязки в виде $e^{(n)} = E^{(n)} \tilde{f}$.

Справедливы следующие утверждения.

Утверждение 1. Матрица невязки $E^{(n)}$ для алгоритма метода наискорейшего спуска определяется рекуррентным соотношением

$$E^{(n+1)} = (I - \beta_n K K^T) E^{(n)}, \quad n = 0, 1, 2, \dots,$$

а матрица $E^{(0)}$ задается выражением

$$E^{(0)} = \begin{cases} I, & \text{если } \varphi^{(0)} = \bar{0}; \\ I - K K^T, & \text{если } \varphi^{(0)} = K^T \tilde{f}. \end{cases}$$

Утверждение 2. Матрица невязки $E^{(n)}$ для алгоритма метода сопряженных градиентов определяется рекуррентным соотношением

$$E^{(n+1)} = (I - \beta_n K K^T) E^{(n)} - \frac{\beta_n \gamma_n}{\beta_{n-1}} [E^{(n)} - E^{(n-1)}],$$

$$E^{(-1)} = I, \quad n = 0, 1, 2, \dots, \quad \gamma_0 = 0,$$

а матрица $E^{(0)}$ задается выражением

$$E^{(0)} = \begin{cases} I, & \text{если } \varphi^{(0)} = \bar{0}; \\ I - K K^T, & \text{если } \varphi^{(0)} = K^T \tilde{f}. \end{cases}$$

Доказательство утверждений непосредственно следует из определения матрицы невязки и соотношений (2), (3).

С использованием введенной матрицы невязки функционал (6) для итерационных алгоритмов можно переписать как

$$U(n) = \frac{\frac{1}{N} \|E^{(n)} \tilde{f}\|^2}{\left\{ \frac{1}{N} \text{Sp}[E^{(n)}] \right\}^2}, \quad n=0,1,2,\dots$$

Исследование эффективности оценки номера оптимальной итерации. Можно предположить, что n_U отличается от n_{opt} , и поэтому для определения проигрыша решения $\varphi^{(n_U)}$ по сравнению с $\varphi^{(n_{\text{opt}})}$ были проведены многочисленные вычислительные эксперименты. Результаты некоторых экспериментов обсудим далее.

В качестве матрицы K была взята матрица размера 500×400 . Число обусловленности матрицы определялось выражением $\text{cond}(K) = \lambda_{\max} / \lambda_{\min}$, где λ_{\max} , λ_{\min} – максимальное и минимальное сингулярные числа матрицы K . В экспериментах число обусловленности менялось от 300 до $3 \cdot 10^9$. Правая часть системы (1) искажалась нормально распределенным шумом с относительным уровнем $\delta = \|\eta\| / \|\tilde{f}\|$.

На рис. 1 приведен график функционала относительной ошибки $\varepsilon(n) = \|\varphi^{(n)} - \bar{\varphi}^+\| / \|\bar{\varphi}^+\|$ (кривая 1) и график функционала $U(n)$ (кривая 2) для итерационной процедуры (2). Число обусловленности матрицы равно $3 \cdot 10^9$, а $\delta = 0,02$. Видно, что номера n_{opt} и n_U совпадают, и, следовательно, постро-

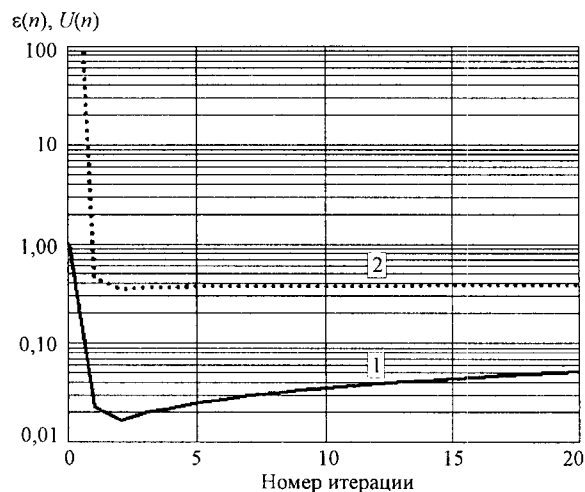


Рис. 1

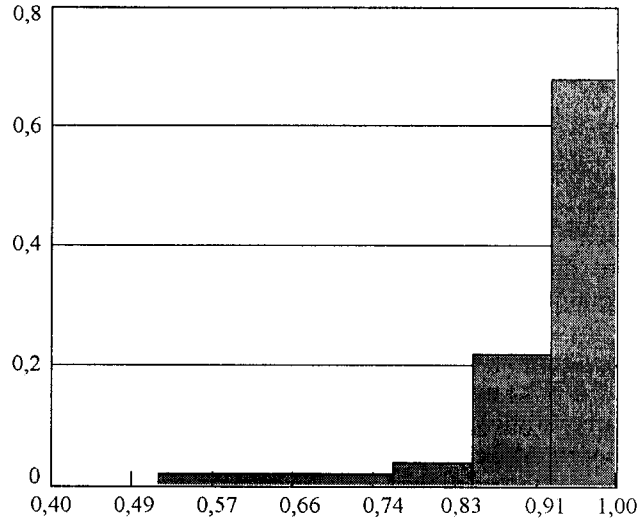


Рис. 2

енный критерий останова позволяет получать решения с наименьшими ошибками. Аналогичным образом ведут себя кривые и для метода сопряженных градиентов.

Для более «тонкого» исследования эффективности оценки n_U введем величину $E_U = \frac{\|\varphi^{(n_{opt})} - \bar{\varphi}^+\|}{\|\varphi^{(n_U)} - \bar{\varphi}^+\|}$, которую назовем коэффициентом эффективности оценки n_U . Очевидно, что значения E_U изменяются в интервале $[0, 1]$, и чем ближе значение E_U к единице, тем меньше «проигрыш» в точности решения $\varphi^{(n_U)}$ по сравнению с $\varphi^{(n_{opt})}$. Так как E_U является случайной величиной, то по выборке объемом 50 вычислялись следующие числовые ха-

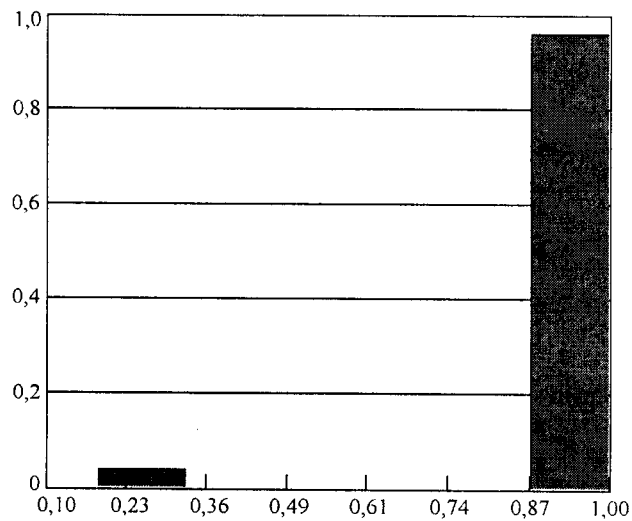


Рис. 3

Таблица 1

Уровень шума δ	\min_E	m_E	\max_E
0,00015	1,0	1,0	1,0
0,005	0,913	0,987	1,0
0,01	0,738	0,949	1,0
0,05	0,515	0,939	1,0
0,10	0,358	0,911	1,0
0,15	0,213	0,919	1,0
0,25	0,46	0,981	1,0
0,50	1,0	1,0	1,0

рактические характеристики: минимальное значение \min_E , максимальное значение \max_E , выборочное среднее m_E и строились гистограммы относительных частот. На рис. 2 представлена гистограмма для метода наискорейшего спуска, при этом $\text{cond}(K) = 3 \cdot 10^9$, $\delta = 0,05$, а на рис. 3 – гистограмма для уровня шума $\delta = 0,20$. Видно, что большая часть значений коэффициента эффективности находится в области 0,8–1,0, но с увеличением уровня шума появляются меньшие значения коэффициента, хотя и с незначительной вероятностью (0,02–0,04).

В табл. 1 для метода наискорейшего спуска, разных уровней шума и $\text{cond}(K) = 3 \cdot 10^9$ приведены числовые характеристики случайной величины E_U , а в табл. 2 – для $\text{cond}(K) = 3 \cdot 10^4$. Видно, что при уменьшении числа обусловленности эффективность оценки увеличивается. В табл. 3 приведены числовые характеристики для метода сопряженных градиентов и для $\text{cond}(K) = 3 \cdot 10^9$. Из сравнения табл. 1 и 3 видно сохранение закономерностей изменения числовых характеристик величины E_U .

Таблица 2

Уровень шума δ	\min_E	m_E	\max_E
0,00015	1,0	1,0	1,0
0,005	0,843	0,968	1,0
0,01	0,761	0,963	1,0
0,05	0,768	0,975	1,0
0,10	0,527	0,982	1,0
0,15	1,0	1,0	1,0
0,25	1,0	1,0	1,0
0,50	1,0	1,0	1,0

Т а б л и ц а 3

Уровень шума δ	\min_E	m_E	\max_E
0,00015	1,0	1,0	1,0
0,005	0,661	0,941	1,0
0,01	0,601	0,960	1,0
0,05	0,326	0,930	1,0
0,10	0,849	0,993	1,0
0,15	0,850	0,976	1,0
0,25	0,601	0,992	1,0
0,50	1,0	1,0	1,0

Заключение. Предложенный в работе алгоритм позволяет достаточно точно оценить номер итерации, на которой решение имеет минимальную ошибку. Аналогичные алгоритмы могут быть построены и для других итерационных процессов: метода простой итерации и метода наименьших ошибок. Алгоритм также допускает обобщение на случай бесконечномерного операторного уравнения 1-го рода.

СПИСОК ЛИТЕРАТУРЫ

1. Алифанов О. М., Артюхин Е. А., Румянцев С. В. Экстремальные методы решения некорректных задач. М.: Наука, 1988.
2. Гилязов С. Ф. Методы решения линейных некорректных задач. М.: Изд-во МГУ, 1987.
3. Вайникко Г. М., Веретенников А. Ю. Итерационные процедуры в некорректных задачах. М.: Наука, 1986.
4. Воскобойников Ю. Е., Преображенский Н. Г., Седельников А. И. Математическая обработка эксперимента в молекулярной газодинамике. Новосибирск: Наука, 1984.
5. Морозов В. А., Гребенников А. И. Методы решения некорректно поставленных задач: алгоритмический аспект. М.: Изд-во МГУ, 1992.
6. Воскобойников Ю. Е. Оценивание оптимального параметра регуляризирующих алгоритмов восстановления изображений // Автометрия. 1995. № 3. С. 68.
7. Golub G. H., Heath M., Wahba G. Generalized cross-validation as a method for choosing a good ridge parameter // Technometrics. 1979. 21, N 2. P. 215.

Новосибирский государственный
архитектурно-строительный университет,
voscob@ngasu.nsk.su

Поступила в редакцию
6 ноября 2003 г.