УДК 516.6

И. Н. Карабан

Некоторые аспекты регуляризации плохо обусловленных систем линейных уравнений

Многие задачи обработки данных экспериментов сводятся к решению систем плохо обусловленных алгебраических уравнений. Из-за широты области решений таких систем возникает задача выбора единственного решения на основе дополнительной информации. Для этого необходимо располагать эффективным численным методом получения множества решений. Ниже эта задача рассматривается применительно к некоторым методам устойчивого решения некорректных задач [1].

В соответствии с методом замены уравнения $\tilde{A}z = \tilde{u}$ близким ему уравнением имеем

$$(\tilde{A} + \alpha E)\, z_\alpha = \tilde{u}, \tag{1}$$

где $\alpha > 0$ — числовой параметр; E — единичная матрица; $\tilde{A}$ $[n \times n]$ — симметричная положительно определенная матрица; z_α — решение уравнения.

Остановимся на определении диапазона значений α, где каждая пара компонентов z_α связана линейной зависимостью. Это позволяет установить однозначную связь между всеми компонентами решения, что необходимо при выборе единственного решения. По формулам Крамера имеем

$$z_{\alpha i} = [|\tilde{A}_i| + \alpha y_i + \varphi_1(\alpha)]/[|\tilde{A}| + \alpha s + \varphi_2(\alpha)], \tag{2}$$

где $z_{\alpha i}$ — i-я компонента z_α; $|\tilde{A}_i|$ — определитель матрицы, полученной заменой i-го столбца $\tilde{A}$ на $\tilde{u}$; s, y_i — величины, зависящие от $\tilde{A}$ и $\tilde{u}$; $\varphi_1(\alpha)$, $\varphi_2(\alpha)$ — алгебраические полиномы соответственно $(n-1)$-й и n-й степени.

Для получения разложения знаменателя (2) удобно ввести в рассмотрение n-вектор a_j, в j-й строке которого находится α, а в остальных — нули.

Теорема 1. *Знаменатель выражения (2) содержит следующие члены: определитель $|\tilde{A}|$; определители матриц, полученных из $\tilde{A}$ заменой j-го столбца на a_j ($j = \overline{1, n}$); определители матриц, полученных из $\tilde{A}$ заменой всех возможных пар столбцов с номерами p и q соответственно на a_p и a_q столбцы; определители матриц, полученных из $\tilde{A}$ заменой всех возможных троек столбцов с номерами p, q, r соответственно на a_q, a_p, a_r столбцы; определители матриц, полученных из $\tilde{A}$ заменой всех возможных сочетаний четырех, пяти..., $n-1$ столбцов на соответствующие столбцы a_j; определители матрицы, все диагональные элементы которой равны α, а все недиагональные — нули.*

Доказательство. Разложим определитель $|\tilde{A} + \alpha E|$ в ряд Тейлора в точке $\alpha = 0$. Для вычисления необходимых производных применим известное правило дифференцирования определителя по параметру, которое и дает требуемый результат.

Следует отметить, что для разложения знаменателя в ряд в принципе можно использовать известное представление определителя $|\tilde{A} - \lambda E|$ при $\lambda = -\alpha$ [2] Однако примененное выше разложение отличается более наглядными зависимостями между искомыми коэффициентами разложения и элементами $\tilde{A}$, $\tilde{u}$.

Полученное разложение позволяет определить числовые значения коэффициентов при α^k, $k = \overline{1, n}$, в числителе и знаменателе выражения (2). Имеют место два следствия доказанной теоремы.

Следствие 1. *Коэффициент знаменателя (2) при α^k равен алгебраической сумме миноров матрицы $\tilde{A}$, полученных вычеркиванием всех возможных сочетаний из k столбцов и k строк, при этом в каждом таком сочетании номера строки и столбца совпадают между собой.*

Следствие 2. *Коэффициент числителя (2) при α^k равен алгебраической сумме миноров матриц, полученных вычеркиванием (из $\tilde{A}$ с i-м столбцом, замененным на $\tilde{u}$) всех возможных сочетаний k столбцов и k строк. В каждом сочетании номера строки и столбца совпадают между собой, но не совпадают с номером i столбца $\tilde{A}$, замененного на $\tilde{u}$.*

Имея числовые значения коэффициентов при α^k в числителе и знаменателе (2), можно оценить диапазон значений α, в котором связь любых двух компонентов z_α близка к линейной. Для этого достаточно путем перебора найти область значений α, в которой одновременно выполняется $n+1$ неравенство

$$\| A_i | + \alpha y_i | \gg | \varphi_1(\alpha) |, \; \| \tilde{A} | + \alpha s | \gg | \varphi_2(\alpha) |. \tag{3}$$

Использование (3) для идентификации области линейной связи между компонентами регуляризованного решения позволяет заметно сократить машинное время по сравнению с многократным решением уравнения (1) для той же цели. При этом экономия машинного времени увеличивается по мере роста n и числа рассматриваемых значений α (т. е. шага по α).

Теорема 2. *Любая пара компонентов регуляризованного решения связана линейной зависимостью*

$$z_{\alpha i} \approx (z_i y_j - z_j y_i)/(y_j - z_j s) - (y_i - z_i s)\, z_{\alpha j}/(y_j - z_j s), \tag{4}$$

где z_i и z_j — соответственно i-я и j-я компоненты решения уравнения $\tilde{A}z = \tilde{u}$.

Доказательство. Пренебрегая нелинейными членами в (2) при условии соблюдения (3), определим значения параметра α из выражений (2), соответствующих i-й и j-й компонентам решения: $\alpha_i = (z_{\alpha i} | \tilde{A} | -$

$-|A_i|)/(y_i - z_{\alpha i}s)$; $\alpha_j = (z_{\alpha j}|\tilde{A}| - |A_j|)/(y_j - z_{\alpha i}s)$. Учитывая, что $\alpha_i = \alpha_j$ и опуская промежуточные выкладки, получаем требуемый результат.

Полагая $\varphi_1(\alpha) \approx 0$, $\varphi_2(\alpha) \approx 0$, формулу (2) перепишем в виде

$$z_{\alpha i} = y_i/s + D/(\alpha + |A|/s), \tag{5}$$

где $D = (|A_i|s - |\tilde{A}|y_i)/s^2$.

Видно, что в силу симметрии и положительной определенности $\tilde{A}$ вертикальная асимптота (5) характеризуется отрицательным значением параметра $\alpha = -|\tilde{A}|/s$. При различных сочетаниях знаков y_i и D могут иметь место четыре типа зависимостей.

Выражение (2) может быть использовано для получения семейства решений на основе метода регуляризации А. Н. Тихонова. Известно, что точка минимума сглаживающего функционала

$$M^\alpha[z, \tilde{u}, \tilde{A}] = \|\tilde{A}z - \tilde{u}\|^2 + \alpha\|z\|^2 \tag{6}$$

(здесь $\|\cdot\|$ — евклидова норма) находится из уравнения Эйлера

$$(\tilde{A}'\tilde{A} + \alpha E)z_\alpha = \tilde{A}'\tilde{u}, \tag{7}$$

сходного с (1). По аналогии с (6) может быть получена формула

$$z_{\alpha i} = y_i/s + D/(\alpha + |\tilde{A}|^2/s), \tag{8}$$

где y_i — алгебраическая сумма главных миноров матрицы, полученной заменой i-го столбца $\tilde{A}'\tilde{A}$ на $\tilde{A}'\tilde{u}$; $s > 0$ — алгебраическая сумма главных миноров матрицы $\tilde{A}'\tilde{A}$; $D = (|\tilde{A}_i||\tilde{A}| - y_i|\tilde{A}|^2)/s^2$, $|\tilde{A}_i|$—определитель матрицы, полученной из $\tilde{A}$ заменой i-го столбца на $\tilde{u}$.

Использование линейных соотношений (4) при решении неустойчивых (плохо обусловленных) задач двумя рассмотренными методами позволяет заметно сократить объем вычислений, необходимых для получения семейства решений z_α с различными числовыми значениями α. При этом экономия машинного времени возрастает по мере уменьшения шага при переборе α, а также при увеличении размерности задачи. Соотношения (3) отражают специфику неустойчивых задач, где область возможных решений вырождается в гиперлинию. При этом фиксация любой компоненты решения однозначно определяет все остальные его составляющие. Поэтому любая компонента решения z_α формально может играть роль параметра α, варьированием которого подбирается подходящее решение.

Зависимости (2) при $\varphi_1(\alpha) \to 0$, $\varphi_2(\alpha) \to 0$ позволяют получить удобную формулу для расчета параметра регуляризации по методу невязки.

Теорема 3. *Значение параметра регуляризации, найденного по методу невязки, задается формулой*

$$\alpha_{1,2} \approx (\delta^2|A|^2 s)\Big/\left(\sum_{i=1}^n B_i^2 - \delta^2 s\right) \pm$$
$$\pm\left[(\delta^2(|A|^2 s)^2\Big/\left(\sum_{i=1}^n B_i^2 - \delta^2 s\right)^2 + (\delta^2|A|^4)\Big/\left(\sum_{i=1}^n B_i^2 - \delta^2 s\right)\right]^{1/2}, \tag{9}$$

где δ — параметр, характеризующий близость приближенной и точной правой части уравнения $Az = \tilde{u}$ (A — точно заданная матрица).

Доказательство. По аналогии с (2) запишем выражения для $z_{\alpha i}$ применительно к (7):

$$z_{\alpha i} \approx [|\tilde{A}||\tilde{A}_i| + \alpha(\delta)y_i]/[|\tilde{A}|^2 + \alpha(\delta)s]. \tag{10}$$

Нелинейные члены в числителе и знаменателе (10) опущены, т. е. предполагается, что имеют место неравенства $\| \tilde{A} \| \tilde{A}_i | + \alpha(\delta) y_i | \gg | \varphi_1(\alpha(\delta)) |$, $\| \tilde{A} |^2 + \alpha(\delta) s | \gg | \varphi_2(\alpha(\delta)) |$, где $\alpha(\delta)$ находится по формуле (9). Выбирая δ с учетом этих неравенств и подставляя (10) в соотношение $\| Az_\alpha - \tilde{u} \| = \delta^2$ ($\| \cdot \|$ — евклидова норма), получаем

$$\sum_{i=1}^{n} \left[\sum_{k=1}^{n} a_{ik} (| A_k \| A | + \alpha y_k) - u_i (| A |^2 + \alpha s) \right] \approx \delta^2 (| A |^2 + \alpha s)^2.$$

Последнее соотношение можно переписать в виде

$$\sum_{i=1}^{n} \left[\sum_{k=1}^{n} a_{ik} | A_k \| A | - \tilde{u}_i | A |^2 + \alpha \left(\sum_{k=1}^{n} a_{ik} y_k - \tilde{u}_i s \right) \right]^2 \approx \delta^2 (| A |^2 + \alpha s)^2.$$

Учитывая, что $\sum_{k=1}^{n} a_{ik} | A_k \| A | - \tilde{u}_i | A |^2 = 0$, получаем квадратное уравнение относительно α: $\alpha^2 \sum_{i=1}^{n} B_i^2 \approx \delta^2 (| A |^2 + \alpha s)^2$, где $B_i = \sum_{k=1}^{n} a_{ik} y_k - \tilde{u}_i s$. Решая это уравнение, получаем (9), что и требовалось доказать.

Следствие 1. *При* $\delta^2 > \left(\sum_{i=1}^{n} B_i^2 \right) / s^2$ *имеем* $\alpha_1 < 0$, $\alpha_2 < 0$.

Следствие 2. *При* $\delta^2 < \left(\sum_{i=1}^{n} B_i^2 \right) / s^2$ *имеем* $\alpha_1 > 0$, $\alpha_2 < 0$. *Таким образом положительное значение параметра* α *существует только при достаточно малом* δ, *определяемом из последнего неравенства.*

Формулы, аналогичные (9), в принципе могут быть получены применительно к другим методам выбора параметра регуляризации. Использование аналитических формул в вычислительном отношении нагляднее, чем применение методов ньютоновского типа, рассмотренных, например, в [3]. Поэтому приближенные аналитические решения следует иметь в виду как дополнение к численным методам определения параметра регуляризации.

1. *Тихонов А. Н., Арсенин В. Я.* Методы решения некорректных задач.— М.: Наука, 1979.— 284 с.
2. *Гантмахер Ф. Р.* Теория матриц.— М.: Наука, 1967.— 574 с.
3. *Гордонова В. И., Морозова В. А.* Численные алгоритмы выбора параметра в методе регуляризации.— Журн. вычислит. матем. и мат. физики 1973, 13, № 3, с. 539—545.

ВНИИВО

Поступила в редакцию 21.12.81
после переработки — 09.09.82