

Член-корреспондент АН СССР Л. А. ВАЙНШТЕЙН

О ЧИСЛЕННОМ РЕШЕНИИ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ ПЕРВОГО РОДА С ИСПОЛЬЗОВАНИЕМ АПРИОРНЫХ СВЕДЕНИЙ О ВОССТАНАВЛИВАЕМОЙ ФУНКЦИИ

Оптимальная обработка данных при численном решении интегральных уравнений первого рода (в случае, когда какие-либо количественные сведения об искомой функции отсутствуют — известно лишь, что она существует и изменяется не слишком быстро) сводится ⁽¹⁾ к одновременному сглаживанию погрешностей в заданной и искомой функциях с помощью двух систем ортонормальных функций, связанных с ядром интегрального уравнения. Оптимальной обработкой называется такая, для которой средняя квадратичная ошибка минимальна; фактически следует пользоваться подоптимальным приближением с весовой функцией $w(t)$, оставляющей в разложениях сравнительно небольшое число членов. Это приближение сохраняет свое значение при более общей постановке задачи ⁽²⁾, позволяя уточнить результаты при одновременном упрощении вычислений; оно приемлемо и к вычислительно устойчивым задачам.

Если об искомой функции имеются какие-то априорные сведения, то часто они позволяют сразу представить ее в виде линейной комбинации сравнительно небольшого числа ортонормальных функций, коэффициенты при которых и нужно определить. В этом случае обработку данных следует производить иначе. Для простоты и конкретности предположим, что с помощью какой-то формулы квадратур интегральное уравнение сведено к системе линейных уравнений

$$\sum_s K_{rs} z_s = u_r, \quad r, s = 0, 1, \dots, n, \quad (1)$$

с плохо обусловленной матрицей $\|K_{rs}\|$; здесь z_s — значения искомой функции в опорных точках, u_r — такие же значения известной функции. Пусть априорные сведения позволяют записать z_s в виде

$$z_s = \sum_{j=0}^p b_j P_j(s), \quad (2)$$

где P_j — некоторая система ортонормальных функций, $j = 0, 1, \dots, p$; $p < n$ или даже $p \ll n$,

$$\sum_s P_j(s) P_{j'}(s) = \delta_{jj'}. \quad (3)$$

Так будет, если, например, искомая функция является полиномом степени p или периодической функцией, содержащей известное число обертонов.

Наряду с ортонормальным базисом P_j можно ввести другой ортонормальный базис

$$P_k^*(s) = \sum_j \alpha_{kj} P_j(s), \quad P_j(s) = \sum_k \alpha_{jk} P_k^*(s), \quad j, k = 0, 1, \dots, p, \quad (4)$$

где коэффициенты α_{kj} удовлетворяют соотношениям

$$\sum_j \alpha_{kj} \alpha_{k'j} = \delta_{kk'}. \quad (5)$$

Оказывается ⁽³⁾, что для каждой матрицы $\|K_{rs}\|$ существует базис P_k^* , позволяющий представить решение в наиболее простом виде.

Подставляя выражения (2) в систему (1), мы приходим к системе

$$\sum_j b_j Q_j(r) = u_r, \quad Q_j(r) = \sum_s K_{rs} P_j(s), \quad (6)$$

не имеющей точного решения (поскольку $p < n$). Это связано с тем, что в значения u_r и как следствие — в значения b_j и z_s входят погрешности (помехи); в частности, u_r можно представить в виде ⁽¹⁾

$$u_r = \bar{u}_r + \delta u_r, \quad (7)$$

где δu_r — погрешность. Поэтому вместо системы (6) надо решать систему

$$\sum_{j'} C_{jj'} b_{j'} = a_j, \quad j, j' = 0, 1, \dots, p, \quad (8)$$

$$C_{jj'} = \sum_r Q_j(r) Q_{j'}(r), \quad a_j = \sum_r u_r Q_j(r); \quad (9)$$

значения b_j , удовлетворяющие системе (8), реализуют минимум суммарной квадратичной ошибки

$$E' = \sum_r \left[\sum_j b_j Q_j(r) - u_r \right]^2, \quad (10)$$

т. е. дают решение исходной системы (1) по методу наименьших квадратов.

Элементы симметричной матрицы $\|C_{jj'}\|$ можно представить в виде (ср. ⁽¹⁾)

$$C_{jj'} = \sum_k \chi_k \alpha_{kj} \alpha_{kj'}, \quad j, j', k = 0, 1, \dots, p, \quad (11)$$

где числа α_{kj} , удовлетворяющие соотношениям (5), получаются из уравнений

$$\sum_{j'} C_{jj'} \alpha_{kj'} = \chi_k \alpha_{kj}, \quad \chi_k > 0. \quad (12)$$

Если с помощью этих чисел α_{kj} , ввести функции $P_k^*(s)$ по формулам (4), а также функции $Q_k^*(r)$ и величины a_k^* и b_k^* по формулам

$$Q_k^*(r) = \sum_s K_{rs} P_k^*(s) = \sum_j \alpha_{kj} Q_j(r), \quad (13)$$

$$a_k^* = \sum_j \alpha_{kj} a_j = \sum_r u_r Q_k^*(r),$$

$$b_k^* = \sum_j \alpha_{kj} b_j, \quad z_s = \sum_k b_k^* P_k^*(s),$$

то система (8) примет простейший вид

$$\chi_k b_k^* = a_k^*. \quad (14)$$

Оптимальное решение, соответствующее минимальной средней квадратичной ошибке

$$E_k = \overline{(\hat{b}_k^* - b_k^*)^2} = \frac{\overline{(\delta a_k^*)^2}}{\chi_k (1 + \tau_k)}, \quad \tau_k = \frac{\overline{(\delta a_k^*)^2}}{(\bar{a}_k^*)^2}, \quad (15)$$

дается выражением (ср. ⁽¹⁾)

$$\hat{b}_k^* = \hat{a}_k^* / \chi_k, \quad \hat{a}_k^* = a_k^* / (1 + \tau_k), \quad (16)$$

в то время как «очевидное» решение $b_k^* = a_k^* / \chi_k$ каждого уравнения (14) может привести к значению средней квадратичной ошибки, сильно превышающему минимальное значение (15).

Таким образом, оптимальное решение системы (1) при условиях (2) имеет вид

$$\hat{z}_s = \sum_{k=0}^p \frac{a_k^*}{\chi_k (1 + \tau_k)} P_k^*(s); \quad (17)$$

для него суммарная средняя квадратичная ошибка

$$E = \sum_{j=0}^p \overline{(\hat{b}_j - \bar{b}_j)^2} = \sum_s \overline{(\hat{z}_s - \bar{z}_s)^2} \quad (18)$$

принимает согласно формуле (15) минимальное значение

$$E_{\min} = \sum_k E_k. \quad (19)$$

Величина τ_k характеризует относительную погрешность, с которой мы знаем величину a_k^* ; при этом числитель τ_k , равный

$$(\delta a_k^*)^2 = \sum_{r, r'} \overline{\delta u_r \delta u_{r'}} Q_k^*(r) Q_k^*(r'), \quad (20)$$

предполагается известным, а знаменатель τ_k , равный $(\bar{a}_k^*)^2$, неизвестен; ближе всего к $\bar{a}_k^*$ стоит $\hat{a}_k^*$, поскольку величина $(\hat{a}_k^* - \bar{a}_k^*)^2$ минимальна. Заменяя $\bar{a}_k^*$ на $\hat{a}_k^*$, преобразуем формулу (17) к виду

$$\hat{z}_s = \sum_{k=0}^p \frac{a_k^*}{\chi_k} w(t_k) P_k^*(s), \quad t_k = \frac{(\delta a_k^*)^2}{(a_k^*)^2}, \quad (21)$$

а $w(t_k)$ — весовая функция, введенная ранее ⁽¹⁾ и равная нулю при $t_k > 1/4$.

В ходе построения подоптимального решения (21) помехи (погрешности) отфильтровывались дважды: сначала в функции u , при замене (6) на (8), а затем в функции z , при написании формул (16) и (17). Первая фильтрация минимизирует величину E' , определяемую формулой (10); вторая фильтрация минимизирует величину E , определяемую формулой (18). Что мы выигрываем, используя представление (2) для функции z ? Если считать, что функции $P_j(s)$ уже известны, то вычисления существенно упрощаются по сравнению со случаем, когда априорных сведений нет; теперь надо построить представление (11) для симметричной матрицы $(p+1)$ -го порядка, в то время как ранее ⁽¹⁾ приходилось строить аналогичное представление для несимметричной матрицы более высокого $((n+1)$ -го) порядка. Ошибка E при уменьшении числа каналов обработки от $n+1$ до $p+1$ также должна уменьшиться.

В работе ⁽⁴⁾ система (1) приближенно решалась с помощью разложений

$$\hat{z}_s = \sum_{k=0}^m b_k P_k(s), \quad \hat{u}_r = \sum_{j=0}^m c_j P_j(r) \quad (22)$$

по полиномам Чебышева ⁽⁵⁾; неизвестные b_k определялись из системы

$$\sum_{k=0}^m A_{jk} b_k = c_j, \quad A_{jk} = \sum_{r,s} K_{rs} P_j(r) P_k(s), \quad j, k = 0, 1, \dots, m. \quad (23)$$

В других работах применялись иные разложения. Такой способ построения решения более прост, поскольку он не требует вычисления специальной системы функций, зависящей от ядра. Отфильтровывание помех здесь производится одновременно для функций u и z путем выбора m (см. ⁽⁶⁾).

Ясно, что этот способ неоптимален, однако он включает наиболее важные операции, свойственные оптимальной обработке, которая рассмотрена выше: сглаживание погрешностей и приближенное решение системы (1), причем то и другое производится по методу наименьших квадратов. Поэтому такой простой способ обычно ведет к хорошим результатам (¹, ⁸), в частности позволяет осуществить численное дифференцирование почти без потери точности (⁷).

Преимущества оптимальных решений перед более простым решением по формулам (22) и (23) можно выяснить, вероятно, только путем численного сравнения. Эти преимущества в ряде случаев будут неощутимы, однако наличие оптимального (точнее, подоптимального) решения важно с принципиальной точки зрения: оно показывает возможности восстановления тех или иных деталей неизвестной функции. Решение по формулам (22) и (23), подобно оптимальным решениям, мало чувствительно к погрешностям вычислений и погрешностям в матрице $\|K_{rs}\|$, а также к числу опорных точек и к выбору квадратурной формулы. Подходящая система ортонормальных функций (полиномы Чебышева и др.) может быть использована для численного построения специальных функций, фигурирующих в оптимальных решениях.

Институт физических проблем им. С. И. Вавилова
Академии наук СССР
Москва

Поступило
28 XII 1971

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- ¹ Л. А. Вайнштейн, ДАН, 204, № 5 (1972). ² В. А. Морозов, Журн. вычислит. матем. и матем. физ., 10, № 4, 818 (1971). ³ В. П. Козлов, Изв. АН СССР, сер. физ. атм. и океана, 2, № 2, 137 (1966); 2, № 12, 1230 (1966). ⁴ Л. П. Грабарь, ДАН, 172, № 4, 767 (1967). ⁵ Л. П. Грабарь, Таблицы полиномов Чебышева, ортонормированных на системе равноотстоящих точек, ВЦ АН СССР, М., 1965. ⁶ Л. П. Грабарь, Радиотехника и электроника, 14, № 9, 1557 (1969) (приложение). ⁷ Л. П. Грабарь, Журн. вычислит. матем. и матем. физ., 7, № 6, 1375 (1967). ⁸ О. С. Кожинский, Л. П. Грабарь, Идентификация, Доклады II Всесоюз. совещ. по статистическим методам теории управления (Ташкент, 1970), «Наука», 1970, стр. 39.