

Г. Я. ВОЛОШИН, Т. Н. ГАДИ
 (Новосибирск)

ИССЛЕДОВАНИЕ ПАРАБОЛИЧЕСКИХ МЕТОДОВ ВЫРАВНИВАНИЯ (СГЛАЖИВАНИЯ) РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ

Пусть в результате каких-либо измерений получены значения функции $f(x)$, соответствующие дискретным значениям аргумента $x_1 < x_2 < x_3 < \dots$. Как известно, результаты измерений сопровождаются погрешностями, обусловленными целым рядом случайных факторов: ошибками измерительных инструментов, случайным характером наблюдаемых реализаций процесса и т. д. Прежде чем производить принципиальную обработку результатов наблюдений (находить промежуточные значения $f(x)$, вычислять производные, некоторые интегралы и т. д.), т. е. выявлять основные закономерности наблюдаемого процесса, необходимо так преобразовать значения $f(x_i)$, чтобы разности

$$\Delta f(x_1) = f(x_2) - f(x_1), \quad \Delta f(x_2) = f(x_3) - f(x_2), \quad \dots$$

и

$$\Delta^2 f(x_1) = \Delta f(x_2) - \Delta f(x_1), \quad \dots$$

изменялись «правильно», закономерно, а не случайным образом. Кроме того, необходимо обеспечить выполнение условия минимального отклонения (по какому-либо критерию) преобразованных значений функции $\{\bar{f}(x_i)\}$ от исходных $\{f(x_i)\}$. Как известно*, такое преобразование называется сглаживанием или выравниванием результатов наблюдений.

Существует целый ряд методов сглаживания. Использование их носит эмпирический характер, т. е. неясно, в каких случаях нужно применить одни сглаживающие операторы, в каких — другие. Этот вопрос приходится решать нестрого, интуитивно, что зачастую приводит к ошибкам в выборе метода сглаживания. В данной работе описаны результаты экспериментального исследования параболических методов сглаживания. Полученные зависимости позволяют оценить возможности тех или иных формул, т. е. определить области их применимости.

Суть параболических методов заключается в том, что последовательно сглаживаются все отсчеты функции, причем для сглаживания очередного отсчета $f(x_0)$ обозревается участок, содержащий $2n+1$ от-

* Э. Уиттекер, Г. Робинсон. Математическая обработка результатов наблюдений. М.—Л., Гостехиздат, 1933.

счетов, а именно: $f(x_{-n}), f(x_{-n+1}), \dots, f(x_0), \dots, f(x_n)$. Через эти $2n+1$ точки проводится парабола заданной степени j

$$\varphi(x) = c_0 + c_1 x + c_2 x^2 + \dots + c_j x^j, \quad (1)$$

коэффициенты которой определяются по способу наименьших квадратов. Величина $\varphi(x)$ в точке $x=x_0$ принимается за выравненное значение $\bar{f}(x_0)$ исходного отсчета $f(x_0)$. Следует заметить, что величина j должна быть меньше, чем $2n+1$, ибо в противном случае эффекта сглаживания не получается, так как парабола пройдет точно через невыравненные значения.

Нетрудно вывести формулы для определения коэффициентов параболы (1). Для этого достаточно минимизировать следующее выражение:

$$\sum_{k=-n}^n [\varphi(x_k) - f(x_k)]^2.$$

В результате получим совместную систему $j+1$ линейных уравнений

$$RC = G, \quad (2)$$

где

$$R = \begin{pmatrix} r_{0+0} & r_{1+0} & \dots & r_{j+0} \\ r_{0+1} & r_{1+1} & \dots & r_{j+1} \\ \dots & \dots & \dots & \dots \\ r_{0+j} & r_{1+j} & \dots & r_{j+j} \end{pmatrix}; \quad C = \begin{pmatrix} c_0 \\ c_1 \\ \dots \\ c_j \end{pmatrix};$$

$$G = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \dots \\ \gamma_j \end{pmatrix}; \quad r_p = \sum_{k=-n}^n x_k^p; \quad \gamma_l = \sum_{k=-n}^n f(x_k) x_k^l.$$

Систему (2) можно решить относительно C либо по правилу Крамера, либо приведя ее к виду

$$C = R^{-1}G,$$

где R^{-1} — матрица, обратная к R .

Для того чтобы изучить свойства параболического оператора сглаживания $S\{f(x)\}$ (или просто S), воспользуемся методикой, широко используемой при анализе электрических цепей, а именно: будем представлять исходную функцию в виде взвешенной суммы*

$$f(x_k) = \sum_i a_i \psi_i(x_k)$$

и исследуем воздействие S на элементарные функции $\psi_i(x_k)$. Зная аргюи (хотя бы ориентировочно) спектр класса функций, к которому

* По аналогии с терминологией, принятой в теории связи, будем называть последовательность коэффициентов a_i спектром функции $f(x)$.

должна принадлежать искомая функция $f_{\text{ид}}(x)$, и наиболее вероятный спектр функции ошибок $\xi(x) = f(x) - f_{\text{ид}}(x)$, а также воздействие S на $\psi_l(x)$, можно подобрать оптимальный алгоритм сглаживания*. Действительно, процесс скользящего сглаживания результатов наблюдений можно трактовать как фильтрацию сигнала, максимально возможное подавление помехи (функции ошибок) при возможно меньшем искажении полезного сигнала. Из теории связи известно, что оптимальный фильтр (в случае аддитивной некоррелированной с полезным сигналом стационарной помехи) должен обладать характеристикой

$$K(l) = \frac{\Phi(l)}{\Phi(l) + \eta(l)}, \quad (3)$$

где $l=0, 1, 2, \dots$, $\Phi(l)$ — спектр полезной функции; $\eta(l)$ — спектр помехи (функции ошибок).

Таким образом, задача экспериментального исследования оператора S сводится к построению семейства кривых $\{K_S(l)\}$, каждая из которых соответствует тому или иному конкретному алгоритму сглаживания.

Функция $K(l)$ характеризует степень подавления l -й элементарной функции $\psi_l(x)$ сглаживающим оператором. Чем выше степень сглаживаемой функции $\psi_l(x)$ по сравнению со степенью сглаживающей параболы, тем сильнее подавляется $\psi_l(x)$, т. е. тем меньше будет ее остаточный средний квадрат. В качестве элементарных функций в рассматриваемом случае целесообразно использовать ортогональные алгебраические полиномы возрастающих степеней, поскольку они хорошо «согласуются» с параболическими формулами сглаживания (легко могут быть получены результаты воздействия S на $\psi_l(x)$, и эти результаты наиболее показательны). Кроме того, использование ортогональных полиномов позволяет осуществлять вычисление спектров $\Phi(l)$ и $\eta(l)$ простыми способами.

В данной работе в качестве $\psi_l(x)$ были использованы полиномы Лежандра.

Вычисление $K(l)$ производилось по следующей схеме. Сначала вычислялся средний квадрат исходной функции $\psi_l(x)$

$$K_{\text{исх}}(l) = \left\{ \int_{-1}^1 \psi_l^2(x) dx \right\}^{1/2}, \quad (4)$$

где (n) обозначает операцию численного интегрирования по n точкам.

Затем $\psi_l(x)$ подвергалась сглаживанию и вновь вычислялось среднеквадратическое значение

$$K_{\text{сгл}}(l) = \left\{ \int_{-1}^1 \tilde{\psi}_l^2(x) dx \right\}^{1/2}. \quad (5)$$

Из (4) и (5) определялось значение

$$K(l) = \frac{K_{\text{сгл}}(l)}{K_{\text{исх}}(l)}.$$

* Такой подход возможен благодаря тому, что параболические операторы сглаживания линейны.

При экспериментальном построении $K(l)$ приходится учитывать тот факт, что оператор S нестационарен. Действительно, качество сглаживания зависит от способа задания отсчетов сглаживаемой функции. Так, если одну и ту же кривую задать сначала равноотстоящими отсчетами (вдоль оси x), а затем неравноотстоящими, то сглаженные кривые окажутся различными. Для «стабилизации» нестационарности оператора S функции $K(l)$ были построены для следующих конкретных случаев:

- а) сглаживаемая функция $\psi_l(x)$ квантовалась с равномерным шагом по оси x (число отсчетов 20, 40 и 60 на интервале $[-1, 1]$);
- б) $\psi_l(x)$ квантовалась с равномерным шагом по оси y (число отсчетов 20, 40 и 60);
- в) $\psi_l(x)$ квантовалась с равномерным шагом по длине дуги (число отсчетов, как и в предыдущих случаях, 20, 40 и 60).

Нетрудно показать, что сглаживающая парабола (1) степени j не искажает составляющих $\psi_l(x)$, если $l \leq j$, т. е. если сглаживается кривая

$$\psi_l(x) = \sum_{m=0}^l b_m x^m,$$

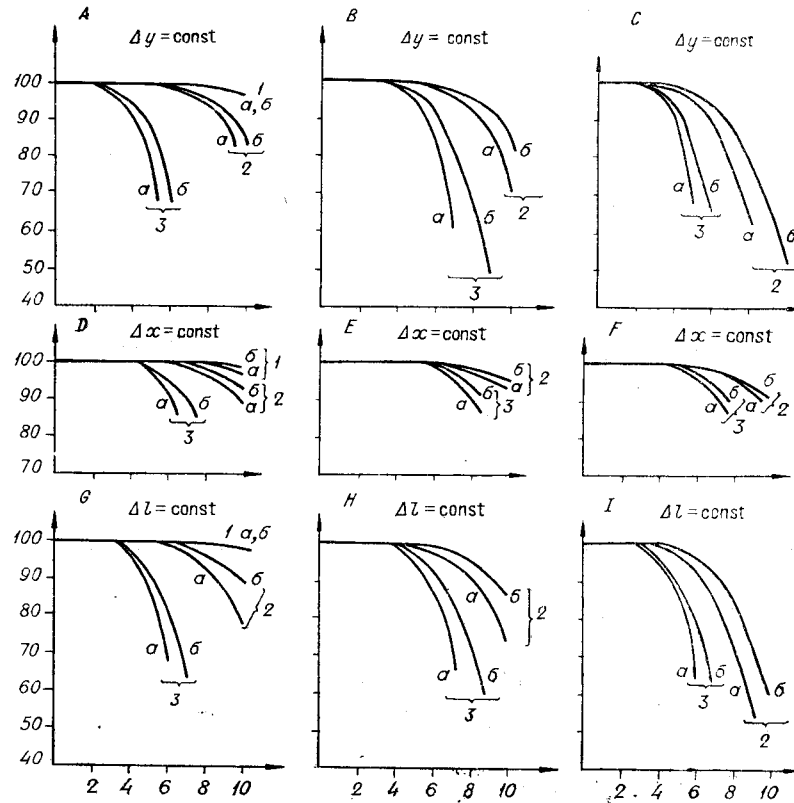


Рис. 1. Характеристики сглаживающих операторов (параболы второй степени):

1 — сглаживаемая кривая задана 20 точками; 2 — сглаживаемая кривая задана 40 точками; 3 — сглаживаемая кривая задана 60 точками; а — 7 итераций; б — 3 итерации; А, D, G — сглаживание по 5 точкам; В, Е, Н — сглаживание по 7 точкам; С, F, I — сглаживание по 9 точкам.

где $l < j$, то коэффициенты c_k сглаживающей параболы (1) получаются равными

$$c_k = b_m \quad \text{при } k=m \quad (l < k; c_k = 0).$$

Таким образом, представляет интерес экспериментальное исследование трех и семи итераций. Каждый из указанных алгоритмов исследован для всех девяти перечисленных выше способов квантования исходной кривой.

О НЕОБХОДИМОМ ЧИСЛЕ ТОЧЕК КВАНТОВАНИЯ

Пусть сглаживаемая кривая задана дискретной последовательностью точек на интервале $[-1, 1]$ (к этому интервалу всегда можно привести произвольный конечный интервал). Возникает вопрос: какое число точек необходимо взять на интервале сглаживания, чтобы уменьшить среднеквадратическое значение ошибок (помех) в заданное число раз? В неявном виде на этот вопрос можно дать ответ в том случае, если нам априори известен характер ошибок.

Пусть для сглаживания данного класса кривых используется оператор S , обладающий характеристикой $K(l)$ (следовательно, нам известно, что вызванные действием оператора S искажения идеальной функции (функции без помех) допустимы). Задан также относительный остаточный уровень помехи (ошибки)

$$\varepsilon = \left\{ \frac{\int_{-1}^1 \bar{\xi}^2(x) dx}{\int_{-1}^1 \xi^2(x) dx} \right\}^{1/2},$$

где $\xi(x)$ — исходная функция ошибки; $\bar{\xi}(x)$ — сглаженная функция ошибки.

Будем полагать, что если функция задана n отсчетами, то в ее спектре нет составляющих со степенями больше, чем $n-1$. Тогда для ортонормированной системы функций $\{\psi_p(x)\}$ мы можем записать**

$$\int_{-1}^1 \xi^2(x) dx = \sum_{p=0}^{n-1} c_p^2 \int_{-1}^1 \psi_p^2(x) dx = \sum_{p=0}^{n-1} c_p^2,$$

где c_p — коэффициенты разложения $\xi(x)$ в ряд по $\{\psi_p(x)\}$.

* «Случай N итераций» — сглаживание какой-либо кривой одним и тем же оператором S N раз подряд.

** Здесь и в дальнейшем погрешности численного интегрирования не учитываются.

Для сглаженной функции ошибки $\bar{\xi}(x)$ можно записать

$$\int_{-\frac{1}{n}}^1 \bar{\xi}^2(x) dx = \sum_{p=0}^{n-1} c_p^2 \int_{-\frac{1}{n}}^1 \bar{\psi}_p(x) dx =$$

$$= \sum_{p=0}^{n-1} c_p^2 K^2(p) \int_{-\frac{1}{n}}^1 \psi_p^2(x) dx = \sum_{p=0}^{n-1} c_p^2 k^2(p).$$

Таким образом,

$$\varepsilon = \left\{ \frac{\sum_{p=0}^{n-1} c_p^2 K^2(p)}{\sum_{p=0}^{n-1} c_p^2} \right\}^{1/2}. \quad (6)$$

Для равномерно распределенной ошибки со статистически независимыми значениями в разных точках наиболее вероятный спектр имеет вид $c_p = \text{const}$ при $1 \leq p \leq n-1$. В этом случае

$$\varepsilon = \left\{ \frac{\sum_{p=0}^{n-1} K^2(p)}{n} \right\}^{1/2}. \quad (7)$$

Формулы (6) и (7) являются основой для вычисления необходимого числа отсчетов. Можно предложить следующую процедуру вычисления.

Выбираем минимально возможное удовлетворяющее нас число отсчетов n . Задавшись некоторым ε и выбрав требуемую характеристику $K(l)$, вычисляем по формуле (6) необходимое n_1 . Если окажется, что $n_1 \leq n$, то задача решена. Если же $n_1 > n$, то впредь нужно брать n_1 отсчетов, а изменение $K(l)$, вызванное увеличением числа отсчетов, необходимо компенсировать увеличением числа итераций при сглаживании.

Краткое описание программы сглаживания. Для проведения экспериментов и дальнейшего практического использования разработана программа (в системе команд машины М-20), позволяющая сглажи-

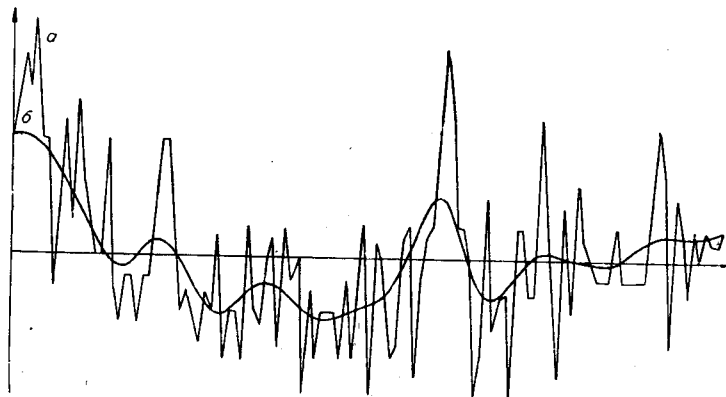


Рис. 2. Пример сглаживания экспериментальной кривой:
 а — исходная кривая; б — результат сглаживания исходной кривой параболой второй степени по 9 точкам 20 итерациями.

вать экспериментальные кривые, заданные дискретной последовательностью точек.

Программа состоит из следующих основных блоков:

1) блок выбора системы координат; позволяет в случае необходимости исходную таблицу значений функции и аргумента перевести из декартовой в полярную систему координат (исходная функция может быть задана как равноотстоящими, так и неравноотстоящими отсчетами);

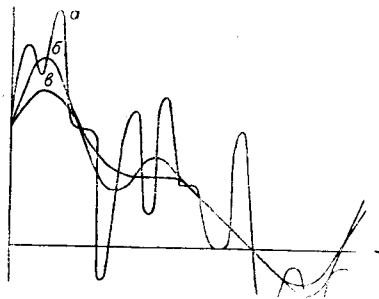
2) блок реорганизации системы отсчетов; позволяет в случае необходимости перестроить систему отсчетов в один из следующих видов: а) постоянный шаг по оси x (произвольное число отсчетов); при пересчете используется кубическая интерполяция; б) постоянный шаг по оси y ; при пересчете используется кубическая интерполяция; в) постоянный шаг по длине дуги; при пересчете используется квадратичная интерполяция;

3) блок сглаживания; позволяет сглаживать кривую с помощью: а) параболы второй степени по 5, 7 или 9 точкам любым числом итераций; б) параболы четвертой степени по 7 или 9 точкам любым числом итераций; в) параболы шестой степени по 9 точкам любым числом итераций.

Режим работы блоков задается с помощью специальной информационной карточки.

Рис. 3. Пример сглаживания экспериментальной кривой:

a — исходная кривая; b — результат сглаживания параболой второй степени по 5 точкам 20 итерациями; c — результат сглаживания параболой второй степени по 5 точкам 40 итерациями.



Для иллюстрации работы программы сглаживания на рис. 2, 3 приведены результаты обработки на М-20 данных некоторых физических измерений.

Поступила в редакцию
14 февраля 1967 г.,
окончательный вариант —
4 мая 1967 г.