УДК 519.168

В. П. Долгин, канд. техн. наук
Севастопольский национальный технический университет
(Украина, 99053, Севастополь, ул. Университетская, 33,
тел. (0692) 543570, e-mail: root@sevgtu.sebastopol.ua)

Процедура комбинаторной фильтрации

Изложен метод повышения эффективности процедуры комбинаторной фильтрации, улучшающий отношение сигнал — шум.

Викладено метод підвищення ефективності процедури комбінаторної фільтрації, поліпшуючий відношення сигнал — шум.

Ключевые слова: оптимизация, модулирующая функция, фильтрация, шум, дисперсия, граница погрешности, конечная разность.

Фильтрация, предназначенная для селекции сигнала, сопряжена с преобразованием спектра [1]. При этом помимо преобразования спектра помехи происходит преобразование спектра сигнала, что приводит к его искажению. Качество фильтрации зависит от спектров сигнала и помехи. В случае совпадения спектров сигнала и помехи процесс фильтрации неэффективен. Статистический метод селекции сигнала лишен указанных недостатков.

Как показано в [1], граница и дисперсия погрешности вычислений, возникающей при наличии шумовой составляющей (оценки Хемминга), неуклонно возрастают с увеличением порядка разностных моделей, что ограничивает диапазон их применения. При этом использование традиционных методов фильтрации неэффективно ввиду характерных особенностей разностных методов [1, 2]. В этом случае желаемого эффекта можно достичь при комбинаторном подходе к решению задачи снижения погрешности.

Известны оценки статистических характеристик для отдельной конечной разности [1—3]. Наличие значительной отрицательной корреляции последовательных разностей используется для обнаружения погрешности в таблице данных. Известен алгоритм учета этой погрешности при наличии погрешности отдельной ординаты [1, 2].

Постановка задачи. Для аддитивной случайной помехи любое значение ординаты можно представить в виде $Y_k = y_k + E_k$, где y_k — истинное значение; E_k — величина помехи (погрешность задания ординаты процесса). В общем случае погрешность вычисления конечной разности n-го порядка, отнесенной к k-й ординате Y_k, можно представить в виде

$$E_k^n = (-1)^n \sum_{i=0}^{n} (-1)^i C_n^i E_{k+i}. \qquad (1)$$

При ограничениях на E_{k+i} типа $|E_{k+i}| \le E$ погрешность конечной разности увеличивается с возрастанием ее порядка n по закону

$$|E_k^n| \le 2^n E; \, |E_{k+i}| \le E, \qquad (2)$$

где E — предельная погрешность [1].

Попытаемся улучшить статистические параметры разностного приближения функций, в частности границы и дисперсии погрешности, определяющие отношение сигнал — шум.

Процедура решения. Рассмотрим подход к улучшению статистических оценок вычисления конечных разностей при сохранении ограничений [1] на отдельный шум: погрешности представления данных являются независимыми случайными величинами, ограниченными по абсолютной величине некоторым значением E. Найдем сумму погрешностей m конечных разностей:

$$S^m(E_k^n) = \sum_{j=0}^{m} E_{k+j}^n K_j, \qquad (3)$$

где K_j — коэффициент суммирования j-й конечной разности. С учетом равенства (1) и ограничения (2) выражение (3) можно привести к виду

$$S^m(E_k^n) = K_q \sum_{i=0}^{R} E_{k+R-i} \sum_{j=0}^{m} (-1)^{i-j} C_n^{i-j} k_j. \qquad (4)$$

Здесь k_j — весовой коэффициент, $k_j = K_j / K_q$, где K_q — максимальное значение весового коэффициента; $R = n + m$.

На основании аксиомы треугольника и равенства (4) получим

$$\left| S^m(E_k^n) \right| \le K_q \sum_{i=0}^{R} \left| E_{k+R-i} \sum_{j=0}^{m} (-1)^{i-j} C_n^{i-j} k_j \right|.$$

Для предельного значения погрешности $E \ge |E_{k+R-i}|$ запишем

$$S_n^m(E) = K_q \sum_{i=0}^{R} E \left| \sum_{j=0}^{m} (-1)^{i-j} C_n^{i-j} k_j \right|.$$

ISSN 0204–3572. Electronic Modeling. 2012. V. 34. № 4

Определим с учетом условия (2) границу погрешности сумм конечных разностей, взятых с весовыми коэффициентами:

$$G_n^m(E) = E\sum_{i=0}^{R}\left|\sum_{j=0}^{m}(-1)^j C_n^{i-j} k_j\right|, \tag{5}$$

где $G_n^m(E) \geq \left|S^m(E_k^n)\right| / K_q$. Нетрудно видеть, что при $m=0$ и $k_0 = 1$ оценка Хемминга имеет вид

$$G_n^0(E) = E\sum_{i=0}^{n} C_n^i = E2^n \ .$$

Определим дисперсию. Для этого, поступая по определению [1], выразим дисперсию $D(E)$ погрешности суммы $m+1$ конечных разностей. На основе (4), оставив в силе ограничение, введенное в работе [1], и считая погрешности представления ординат процесса E_k $(k=[0, n+m])$ одинаково распределенными независимыми случайными величинами с дисперсией $D(e)$, после преобразований получим

$$D(E)/D(e) = \sum_{i=0}^{R}\left[\sum_{j=0}^{m}(-1)^j C_n^{i-j} k_j\right]^2 . \tag{6}$$

Эффект фильтрации возрастет, если найти способ уменьшения границы и дисперсии погрешности, например, посредством оптимального выбора весовых коэффициентов k_j, $j=[0, m]$, в формулах (5) и (6) при заданном числе суммируемых последовательных разностей m, т.е. представить процесс выбора k_j, $j=[0, m]$, как некоторую регулярную процедуру в виде модулирующей зависящей от m функции.

По мере поступления данных о процессе появляется возможность улучшения оценок статистических параметров распределения погрешности. Многопараметрическая оптимизация и другие методы [2], в которых используются итерационные процедуры, требуют больших затрат вычислительных ресурсов для определения весовых коэффициентов и часто малоэффективны. Задачу оптимального выбора весовых коэффициентов можно решить с помощью метода минимизации квадрата уклонений погрешности. Покажем это.

Теорема. Среди модулирующих функций $k(j)$ существуют минимизирующие границу и дисперсию погрешности суммы заданного числа m конечных разностей n-го порядка вследствие снижения эффекта воздействия случайной аддитивной составляющей ординат.

Д о к а з а т е л ь с т в о. Известно, что конечная n-я разность полиномиальной функции $(n-1)$-го порядка равна нулю [1]. При наличии случайной аддитивной составляющей в таблице данных n-я разность отлична от нуля. Рассмотрим способ снижения эффекта воздействия указанного фактора.

Описания границы погрешности (5) и дисперсии (6) имеют общее ядро. Поэтому результат вычисления весовых коэффициентов, минимизирующих границу погрешности, будет совпадать с результатом минимизации дисперсии. Найдем производную $F(k)$ функции (6) по аргументу k_q, $q=[0,m]$,

$$F(k)=2\sum_{i=0}^{R} C_n^{i-q} \sum_{j=0}^{m}(-1)^{j+q} C_n^{i-j} k_j,$$

и, приравняв ее нулю, получим

$$\sum_{i=0}^{R} C_n^{i-q} \sum_{j=0}^{m}(-1)^{j+q} C_n^{i-j} k_j = 0,\ q=[0,m]. \tag{7}$$

Изменим порядок суммирования в (7):

$$\sum_{j=0}^{m}(-1)^{j+q} k_j \sum_{i=j}^{R} C_n^{i-q} C_n^{i-j} = 0,\ q=[0,m]. \tag{8}$$

Последовательно меняя значение $q=[0,m]$ в выражении (8), получаем систему $m+1$ уравнений:

$$\begin{gathered}
\sum_{j=0}^{m}(-1)^{j} k_j \sum_{i=j}^{R} C_n^{i} C_n^{i-j} = 0,\ \ q=0,\\
\sum_{j=0}^{m}(-1)^{j+1} k_j \sum_{i=j}^{R} C_n^{i-1} C_n^{i-j} = 0,\ \ q=1,\\
\sum_{j=0}^{m}(-1)^{j+2} k_j \sum_{i=j}^{R} C_n^{i-2} C_n^{i-j} = 0,\ \ q=2,\\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
\sum_{j=0}^{m}(-1)^{j+m} k_j \sum_{i=j}^{R} C_n^{i-m} C_n^{i-j} = 0,\ \ q=m.
\end{gathered} \tag{9}$$

Для получения нетривиального результата решения системы уравнений (9) необходимо ввести уравнение с ненулевым свободным членом, например, задав один из весовых коэффициентов. Существование нетривиального решения системы (9) является доказательством теоремы.

Можно заметить, что матрица коэффициентов для вычисления k_j, $j=[0,m]$, содержит нулевые элементы. Используя свойство симметрии биномиальных коэффициентов и значение символа Кронекера [4, 5], удобно свести матрицу решения системы уравнений (9) к квадратной размерностью $m\times m$. Представляется возможным упростить процедуру вычислений коэффициентов посредством преобразования матрицы к симметрической.

Выполним ортогонализацию матрицы вычисления весовых коэффициентов. Для этого модифицируем выражение (7). Исключим нулевые слагаемые, введя требование $C_n^{i-q}>0$, что предполагает выполнение условий $i-q\ge 0$, $i-q\le n$. Отсюда вытекает

$$\sum_{i=q}^{n+q} C_n^{i-q} \sum_{j=0}^{m} (-1)^{j+q} C_n^{i-j} k_j = 0,\ \ q=[0,m],$$

или

$$\sum_{i=0}^{n} C_n^{i} \sum_{j=0}^{m} (-1)^{j+q} C_n^{i+q-j} k_j = 0,\ \ q=[0,m].$$

Введя требование $C_n^{i+q-j}>0$, запишем условия его реализации: $j\le i+q$, $j\ge i+q-n$. Изменив порядок суммирования, с учетом введенных условий получим

$$\sum_{j=0}^{m} (-1)^{j+q} k_j \sum_{i=0}^{m+n-j} C_n^{i+q-m} C_n^{i+j-m} = 0,\ \ q=[0,m].$$

Теперь введем требования $C_n^{i+q-m}>0$ и $C_n^{i+j-m}>0$, что порождает дополнительные условия $i\le m-j$, $i\ge m-q$, реализация которых, связанная с изменением пределов суммирования по индексу i, приводит к следующему результату:

$$\sum_{j=0}^{m} (-1)^{j+q} k_j \sum_{i=0}^{n+q-j} C_n^{i} C_n^{i+j-q} = 0,\ \ q=[0,m].$$

Учитывая, что [2—5]

$$\sum_{i=0}^{n+q-j} C_n^{i} C_n^{i+j-q} = C_{2n}^{n+q-j},$$

окончательно получаем

$$\sum_{j=0}^{m} (-1)^{j+q} C_{2n}^{n+q-j} k_j = 0,\ \ q=[0,m]. \tag{10}$$

Таблица нормированных значений коэффициентов модулирующей функции

n	$k_j,\ j=[0,m]$, при m						
	0	1	2	3	4	5	6
1	1,00000	0,857143	0,714286	0,571429	0,428571	0,285714	0,142857
2	0,622223	0,933334	1,000000	0,888888	0,666666	0,400000	0,155555
3	0,400002	0,800001	1,000000	0,952381	0,714285	0,399999	0,133333
4	0,300002	0,720002	1,000000	0,999998	0,749997	0,399998	0,119999
5	0,235728	0,642881	0,964303	1,000000	0,749987	0,385701	0,107137
6	0,196420	0,589266	0,937482	1,000000	0,750016	0,375017	0,098221
7	0,170258	0,550043	0,916702	1,000000	0,749972	0,366640	0,091657
8	0,151594	0,519836	0,899860	1,000000	0,750111	0,360103	0,086702

Рекуррентное соотношение (10) позволяет представить в форме квадратной матрицы $(m+1)\times(m+1)$ коэффициенты при $k_j,\ j=[0,m]$:

$$\begin{matrix} C_{2n}^{n} & -C_{2n}^{n-1} & C_{2n}^{n-2} & \cdots & (-1)^m C_{2n}^{n-m} \\ -C_{2n}^{n+1} & C_{2n}^{n} & -C_{2n}^{n-1} & \cdots & (-1)^{m+1} C_{2n}^{n-m+1} \\ C_{2n}^{n+2} & -C_{2n}^{n+1} & C_{2n}^{n} & \cdots & (-1)^{m+2} C_{2n}^{n-m+2}. \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ (-1)^m C_{2n}^{n+m} & (-1)^{m+1} C_{2n}^{n+m-1} & (-1)^{m+2} C_{2n}^{n+m-2} & \cdots & (-1)^{2m} C_{2n}^{n} \end{matrix} \qquad (11)$$

В таблице приведены результаты вычислений весовых коэффициентов $k_j,\ j=[0,m]$, полученные изложенным методом на основе матрицы (11).

На основе представленных результатов можно сделать вывод о том, что использование комбинаторного подхода обработки данных, содержащих аддитивную случайную составляющую, приводит одновременно к сужению зоны неопределенности вследствие сокращения границы и уменьшению дисперсии случайной составляющей погрешности при увеличении числа учтенных разностей m.

В качестве примера приведем результат синтеза модулирующей функции $F(q)$, $q=[0,m]$, для минимизации границы и дисперсии шумовой составляющей разности $n=9$ при $m=n$ и $k_5=1$. Квадратная матрица (11) коэффициентов M и вектор свободных членов V для вычисления $k_j,\ j=[0,m]$, $j\neq 5$, принимают вид

$$M=\begin{bmatrix} 48600 & -43800 & 31800 & -18600 & 8570 & -3060 & 816 & -153 & 18 & -1 \\ -43800 & 48600 & -43800 & 31800 & -18600 & 8570 & -3060 & 816 & -153 & 18 \\ 31800 & -43800 & 48600 & -43800 & 31800 & -18600 & 8570 & -3060 & 816 & -153 \\ -18600 & 31800 & -43800 & 48600 & -43800 & 31800 & -18600 & 8570 & -3060 & 816 \\ 8570 & -18600 & 31800 & -43800 & 48600 & -43800 & 31800 & -18600 & 8570 & -3060 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 816 & -3060 & 8570 & -18600 & 31800 & -43800 & 48600 & -43800 & 31800 & -18600 \\ -153 & 816 & -3060 & 8570 & -18600 & 31800 & -43800 & 48600 & -43800 & 31800 \\ 18 & -153 & 816 & -3060 & 8570 & -18600 & 31800 & -43800 & 48600 & -43800 \\ -1 & 18 & -153 & 816 & -3060 & 8570 & -18600 & 31800 & -43800 & 48600 \end{bmatrix},$$

$$V=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Решение в пакете with(linalg) прикладных программ системы Maple позволило получить значения весовых коэффициентов k = [0,031948; 0,16220; 0,42573; 0,99132; …; 0,77426; 0,44433; 0,17203; 0,34406] модулирующей функции $F(q)$, $q = [0, m]$:

$$F(q)=\begin{cases} -0{,}09830400000+0{,}1302520000\,x, & x<2, \\ -0{,}3648599999+0{,}2635300000\,x, & x<3, \\ -0{,}5603100006+0{,}3286800002\,x, & x<4, \\ -0{,}1932300000+0{,}2369100000\,x, & x<5, \\ 0{,}9479200000+0{,}0086800000\,x, & x<6, \\ 2{,}354439997-0{,}2257399995\,x, & x<7, \\ 3{,}083770001-0{,}3299300002\,x, & x<8, \\ 2{,}622729998-0{,}2722999998\,x, & x<9, \\ 1{,}410646000-0{,}1376240000\,x & \text{в остальном.} \end{cases}$$

На рис. 1 изображены модулирующие функции $F(q)$ и $F(q)=1$; $q=[0, m]$ и соответствующие им графики границ и дисперсии погрешности. Качество комбинаторной фильтрации можно оценить с помощью довери-

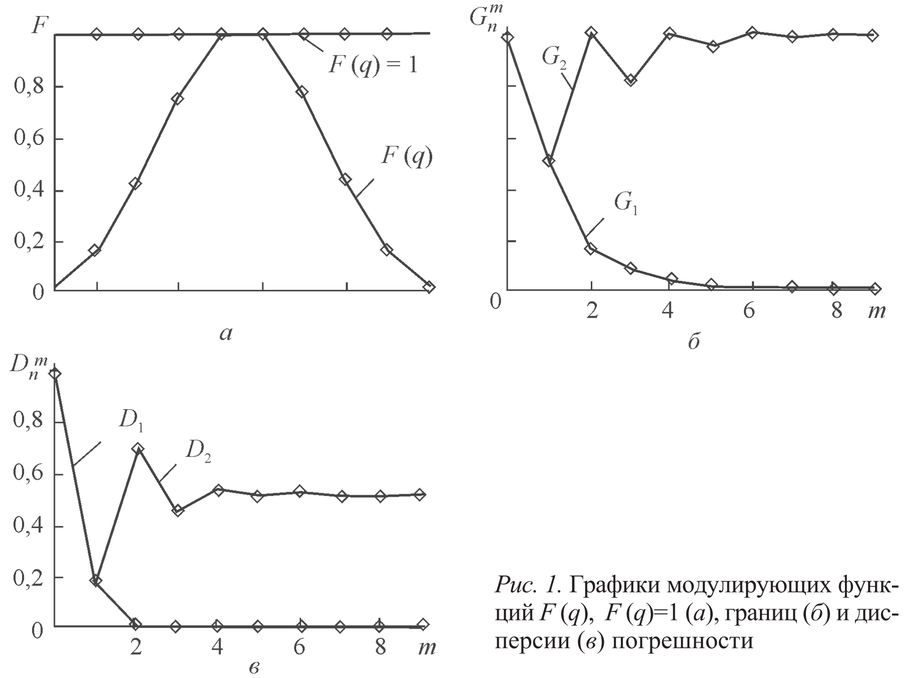

Рис. 1. Графики модулирующих функций $F(q)$, $F(q)=1$ (*а*), границ (*б*) и дисперсии (*в*) погрешности

тельной вероятности [6] того, что погрешность результата фильтрации лежит в интервале $[0, E]$. Для $n=3$ и $m=5n$ вычислены дисперсия ($D_n^m=0{,}00158$) и математическое ожидание выборки ($M_n=1{,}2\cdot 10^{-5}$). При условии, что распределение погрешности вычисления n-й разности соответствует нормальному,

$$f_n(t)=A_n\exp\left(-\frac{(t-M_n)^2}{2D_n^m}\right),\ A_n=\frac{1}{\sqrt{2\pi D_n^m}},$$

вычислены граничные значения $L_n=f_n(t)|_{t=-E}=f_n(t)|_{t=+E}=0{,}425$, $A_n=10{,}03$ и доверительная вероятность при $m=5n$:

$$P_n=\int_{-E}^{+E} f_n(t)\,dt=0{,}988\,.$$

Для сравнения вычислена доверительная вероятность в случае отсутствия суммирования разностей ($m=0$). Дисперсия погрешности в этом случае

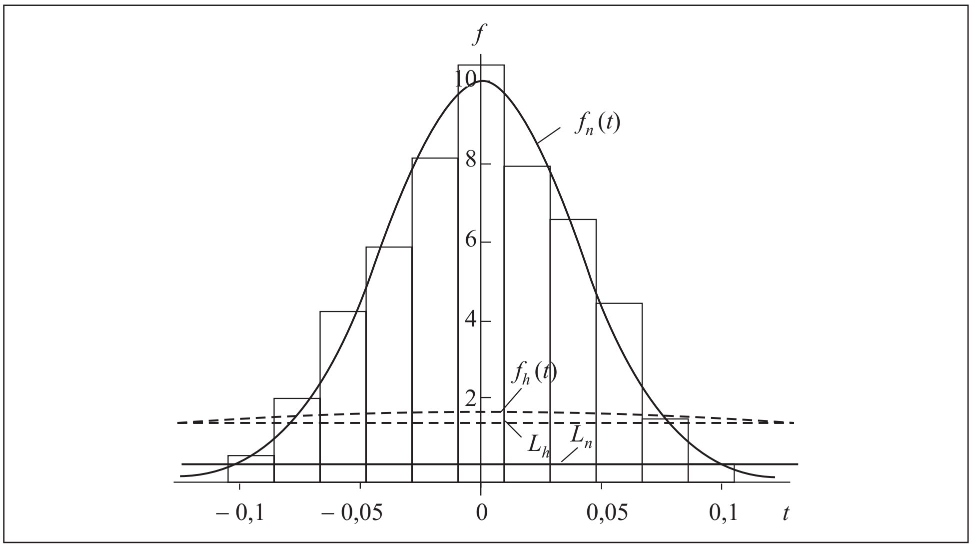

Рис. 2. Кривая, описывающая эффект комбинаторной фильтрации

совпадает с оценкой Хемминга $D_n^0 = EC_{2n}^n = 0{,}0583$ при математическом ожидании $M_h = 0$ и доверительной вероятности

$$P_h = \int_{-E}^{+E} f_h(t)\, dt = 0{,}321,$$

где

$$f_h(t) = A_h \exp\left(-\frac{t^2}{2D_n^0}\right); \; A_h = \frac{1}{\sqrt{2\pi\, D_n^0}},$$

$$L_h = f_h(t)|_{t=-E} = f_h(t)|_{t=+E} = 1{,}52, \; A_h = 1{,}65.$$

Таким образом, качество фильтрации улучшается при суммировании разностей с вычисляемыми весовыми коэффициентами по мере возрастания объема выборки вследствие возможности увеличения числа суммируемых разностей *m*. На рис. 2 представлен пример рассмотренного случая. В численном примере максимальное значение шумовой составляющей принято $E = 0{,}1$. Оценка вероятности появления шумовой составляющей, превышающей значение *E* при использовании процедуры фильтрации составляет $1 - P_n = 0{,}012$, а без наличия фильтрации — $1 - P_h = 0{,}679$. На рис. 2 для наглядности указано максимальное значение $t = E = 0{,}1$.

Выводы

Процедура комбинаторной фильтрации может найти применение в устройствах адаптивного управления стохастическими объектами в виде критериев оценки качества адаптации [7], а также при решении задач, связанных с синтезом моделей диагностики, устройств прогноза развития процесса (например, упредителей, корректоров) в технологических и автотранспортных системах.

The method of raising the efficiency of the procedure of combinatorial filtration has been stated which improves the relation signal—noise.

1. *Hamming R. W.* Numerical Methods for Scientists and Engineers. — N.Y. : McGraw-Hill Book Company, 1973. (Хемминг Р. В. Численные методы. — М. : Наука, 1972. — 400 с.)
2. *Гутер Р. С., Овчинский В. В.* Элементы численного анализа и математической обработки результатов опыта. — М. : Наука, 1970. — 432 с.
3. *Комбинаторный* анализ. Задачи и упражнения / Под ред. К. А. Рыбникова. — М. : Наука, 1982. — 368 с.
4. *Риордан Дж.* Комбинаторные тождества. — М. : Наука, 1982. — 256 с.
5. *Справочник* по специальным функциям / Под ред. М. Абрамовица и И. Стиган. — М. : Наука, 1979. — 830 с.
6. *Вентцель Е. С., Овчаров Л. А.* Теория случайных процессов и ее инженерные приложения. — М. : Высшая шк., 2000. — 383 с.
7. *Бодянский Е. В., Руденко О. Г.* Адаптивные модели в системах управления техническими объектами. — Киев : УМК ВО Украины, 1988. — 202 с.

Поступила 14.11.11

ДОЛГИН Владимир Прохорович, канд. техн. наук, доцент кафедры автомобильного транспорта Севастопольского национального технического университета. В 1958 г. окончил Военно-морское инженерное училище им. Ф. Э. Дзержинского (Ленинград), в 1965 г. — Севастопольский приборостроительный ин-т. Область научных исследований — адаптивные модели в системах управления технологическими объектами.