

ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА

684 / 2-81

9/2-81
P10-80-510

В.Б.Злоказов

**SMOOS - ПРОГРАММА ДЛЯ ФИЛЬТРАЦИИ
НЕСТАЦИОНАРНЫХ СТАТИСТИЧЕСКИХ РЯДОВ**

Направлено в журнал "Computer Physics Communications".

1980

ВВЕДЕНИЕ

К числу важнейших на практике ситуаций, где возникает необходимость фильтрации данных, можно отнести следующие:

А. Временной ряд $f(t)$ /в задачах анализа скрытых периодичностей астрономии, геофизике, геологии и т.д., в задачах факторного анализа, решаемых с помощью исследования спектра мощности ряда и т.д./ представляется в виде такой суммы:

$$f(t) = b(t) + n(t), \quad t = t_0, t_1, \dots, t_n, \quad /1/$$

где $b(t)$ - нестационарный в общем случае тренд, $n(t)$ - стохастический сигнал.

В настоящее время очень хорошо развита техника выделения $n(t)$ для случая стационарного $b(t)$ /или плавно меняющегося/ ^{1,2/}, но значительно хуже дело обстоит в случае существенно нестационарного $b(t)$, особенно для $b(t)$ типа резонансных функций.

Б. Результаты наблюдений физических кривых $f(t)$ /спектры, сечения реакций и т.д./ формально могут быть представлены в виде суммы:

$$f(t) = b(t) + p(t) + n(t), \quad t = t_0, \dots, t_n, \quad /2/$$

где $b(t)$ и $n(t)$ - соответственно низко- и высокочастотные помехи наблюдения, $p(t)$ - информативная компонента. Здесь задача выделения $p(t)$ из $f(t)$ осложнена тем, что границы частотных интервалов, на которых определены $b(t)$, $p(t)$ и $n(t)$, не фиксированы, а довольно сильно размыты, так что в принципе в $p(t)$ могут быть представлены /хотя и с разными весами/ все частоты /например, если $p(t)$ - гауссиан, то ряд Фурье от него по такой системе, как, например, $\cos kx$, не содержит ни одного нулевого коэффициента/. Задачу декомпозиции $f(t)$ можно решать с помощью техники регрессивного анализа, однако не для всех данных удастся успешно строить функцию регрессии с приемлемым числом параметров. Кроме того, для тех случаев, когда такая функция существует, автоматизация обработки больших потоков информации делает актуальной как раз задачу автоматического построения этой функции, для чего необходимо предварительное подавление /хотя бы приближенное/ $b(t)$ и $n(t)$.

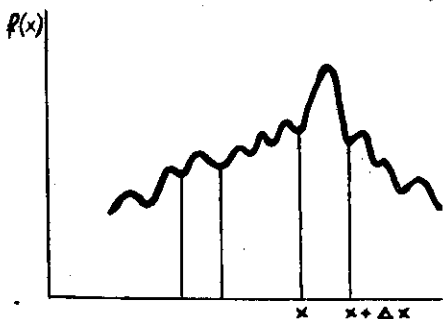


Рис.1. Осциллирующая функция.

Имеется также много других задач, где требуется то или иное сглаживание данных и /или/ устранение тренда. Применяемые для решения перечисленных задач оптимальные цифровые фильтры винеровского типа /свертки, разности и т.д./ имеют высокую эффективность лишь в стационарных или слабонестационарных случаях и не годятся для сильно-нестационарных случаев /например, если тренд имеет резонансную структуру/, так как вносят слишком большие искажения в результат фильтрации. Бо-

лее мощным фильтром является сплайновый фильтр, но и он вносит недопустимые искажения в информационную компоненту /подрезание вершин пиков, заглаживание впадин и других интересных деталей сильнонестационарного тренда/.

Метод, излагаемый в данной работе, базируется на следующих идеях:

- 1/ вводятся понятия меры осцилляций и меры изменчивости дискретных функций;
- 2/ компонента $b(t)$ /в задаче /1// выделяется на основе требования: построить функцию, повторяющую грубый контур функции $f(t)$, но без ее осцилляций, для чего следует минимизировать меру осцилляций искомой функции и одновременно ее расстояние от функции $f(t)$;
- 3/ компонента $b(t)$ /в задаче /2// выделяется на основе требования: построить функцию, снизу огибающую грубый контур функции $f(t)$, но меняющуюся более плавно, чем $f(t)$, для чего следует минимизировать меру изменчивости искомой функции и одновременно ее соответствующим образом взвешенное расстояние от функции $f(t)$.

Таким образом, эти фильтры используют не только частотные, но и амплитудные характеристики данных, что повышает их гибкость и уменьшает вносимые ими искажения в результаты фильтрации.

§ 1. МЕРЫ ОСЦИЛЛЯЦИЙ И ИЗМЕНЧИВОСТИ ФУНКЦИЙ

Рассмотрим сначала непрерывный аналог этих мер. Визуальному восприятию быстрой смены направления осциллирующей кривой

/рис.1/ соответствует быстрая смена знаков и нормы 1-й производной функции этой кривой. С другой стороны, сравнивая два участка кривой, на которых 1-я производная претерпевает одинаковые изменения, мы склонны считать более быстро осциллирующим тот участок кривой, который имеет меньшую длину. Поэтому вполне естественно выбрать в качестве меры осцилляций функции на малом интервале $\{x, x + \Delta x\}$ отношение изменения 1-й производной к длине ее кривой на этом интервале /рис.1/:

$$\bar{c}(x) = \frac{f'(x+\Delta x) - f'(x)}{L(f(x+\Delta x), f(x))}. \quad /1.0/$$

Умножив числитель и знаменатель в $\bar{c}(x)$ на Δx и переходя к пределу при $\Delta x \rightarrow 0$, получим плотность этой меры:

$$\begin{aligned} c(x) &= \lim_{\Delta x \rightarrow 0} \bar{c}(x) = \lim_{\Delta x \rightarrow 0} \frac{(f'(x+\Delta x) - f'(x)) \cdot \Delta x}{\Delta x \cdot L(x, x+\Delta x)} = \\ &= \lim_{\Delta x \rightarrow 0} \frac{f'(x+\Delta x) - f'(x)}{\Delta x} \cdot \frac{1}{\frac{L(x, x+\Delta x)}{\Delta x}} = \frac{f''}{\sqrt{1+f'(x)^2}}. \end{aligned} \quad /1.2/$$

Интегральная мера осцилляций функции на отрезке $\{x_1, x_2\}$ определяется так:

$$\mu(x_1, x_2) = \int_{x_1}^{x_2} c^2(x) dx = \int_{x_1}^{x_2} \frac{f''^2(x)}{1+f'(x)^2} dx. \quad /1.3/$$

Рассмотрим свойства меры /1.3/:

а/ мера μ инвариантна относительно преобразования функции $f(t)$ типа

$$f(x) \rightarrow f(x+T) + C, \quad T - \text{период};$$

б/ пусть $f(x) = P_n(x)$, т.е. является полиномом степени n . Для него /полагая $\{x_1, x_2\}$ равным $\{-\pi, +\pi\}$ / имеем

$$\mu(-\pi, +\pi) = \int_{-\pi}^{+\pi} \frac{P_n''^2(x)}{1+P_n'^2(x)} dx.$$

Сделав замену $z = e^{ix}$ и принимая во внимание, что $dx = \frac{dz}{iz}$, имеем

$$\mu(-\pi, +\pi) = \operatorname{Re} \left\{ \int_{|z|=1} \frac{P_n''^2(z)}{1+P_n'(z)^2} \cdot \frac{dz}{iz} \right\} = \operatorname{Re} \{I\}.$$

Подынтегральное выражение имеет $2n-1$ полюсов: в точке 0 и в корнях уравнения $P_n'(z) = +/\sqrt{-1}$. Обозначим их через $\{z_k\}$. Тогда, по теореме Коши^{/3/},

$$I = 2\pi \sum_{k=1}^{2n-1} \operatorname{res}_{z=z_k} \frac{P_n''^2(z)}{z(1+P_n'(z)^2)}. \quad /1.4/$$

Ограничимся случаем, когда все $z_k \in [-\pi, +\pi]$. Чем больше некратных корней имеет множество $\{z_k\}$ /т.е. чем больше осциллирует полином P_n' /, тем больше членов в сумме в /1.4/ и, следовательно, тем больше значение I . Аналогичное рассмотрение функции $f(x) = \sin(kx)$ устанавливает ту же зависимость μ от k .

Из анализа этих свойств следует, что введенная таким образом мера является вполне адекватной нашим интуитивным представлениям об осцилляциях. Аналогично вводится понятие меры изменчивости функции, которую мы определим как отношение приращения функции к длине приращения аргумента:

$$\bar{d}(x) = (f(x + \Delta x) - f(x)) / \Delta x.$$

Плотность этой меры мы получим, переходя к пределу при $\Delta x \rightarrow 0$:

$$d(x) = f'(x). \quad /1.5/$$

Интегральную меру изменчивости определим аналогично /1.3/:

$$\eta(x_1, x_2) = \int_{x_1}^{x_2} d(x)^2 dx. \quad /1.6/$$

Свойства /1.6/:

- а/ инвариантность при преобразовании $f(x) \rightarrow f(x+T) + C$, $T =$ период;
 б/ если $f(x) = \text{const}$, $\eta = 0$; для $f(x) = ax+b$ $\eta = a^2 \cdot \text{const}$, т.е. мера /1.6/ тоже вполне адекватна нашим интуитивным представлениям об изменчивости функции.

§ 2. ФОРМУЛИРОВКА ВАРИАЦИОННЫХ ЗАДАЧ

Имея меры /1.3/ и /1.6/, мы можем построить искомые фильтры как решения вариационных задач. Пусть $f(x)$ - экспериментально полученная функция на отрезке $\{x_1, x_2\}$.

А. Функцию $g(x)$ - результат сглаживания $f(x)$, т.е. устранения из последней высокочастотных осцилляций, - мы получим как решение следующей задачи: найти функцию, близкую к $f(x)$ в смысле интегральной метрики и одновременно имеющую минимальную меру осцилляций.

Математическая формулировка этой задачи гласит:

$$\text{найти } \min \left\{ \int_{x_1}^{x_2} \frac{g''^2}{1+g'^2} dx + \lambda \int_{x_1}^{x_2} (g(x) - f(x))^2 dx \right\} \quad /2.0/$$

при граничных условиях

$$g(x_1) = f(x_1), g(x_2) = f(x_2), g''(x_1) = 0, g''(x_2) = 0 \quad /2.0'/$$

и некотором заданном $\lambda > 0$.

Прежде чем переходить к дискретному аналогу /2.0/, сделаем одно существенное упрощение этого выражения, ибо нахождение минимума /2.0/ является слишком сложной нелинейной задачей. Так как $g(x)$ есть результат сглаживания $f(x)$, то мы можем подставить в 1-й член в /2.0/ вместо $1+g'^2$ выражение $1+\tilde{g}'^2$, где $\tilde{g}'(x)$ - сглаженная тем или иным /простым/ способом $f'(x)$. Обозначив $1/(1+\tilde{g}'^2(x)) = \epsilon(x)$ и разделив λ на нормирующую величину ϵ_{00} ($\epsilon_{00} = \int_{x_1}^{x_2} \epsilon(x) dx$), мы можем записать дискретный аналог задачи /2.0/, /2.0'/: най-

$$\min \left\{ \sum_{i=1}^{M-1} \epsilon_i (g_{i+1} - 2g_i + g_{i-1})^2 + \frac{\lambda}{\epsilon_{00}} \sum_{i=0}^M (g_i - f_i)^2 \right\} \quad /2.1/$$

при граничных условиях /здесь $\epsilon_{00} = \sum_{i=1}^{M-1} \epsilon_i / (M-2)$ /

$$g_0 = f_0, g_M = f_M, g''_0 = 0, g''_M = 0. \quad /2.1'/$$

Множитель λ не есть аргумент минимизации /как в методе сплайнов/, а задается заранее. Выбирается он из следующих со-

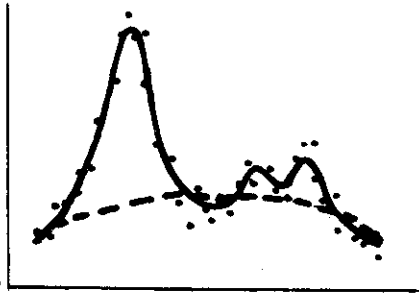


Рис.2. Результаты сглаживания:

- - исходный ряд;
- - сглаженный ряд;
- - огибающая снизу.

образований: если надо увеличить близость сглаженной функции g_i к исходной f_i , следует увеличить λ ; если же надо увеличить гладкость решения g_i , следует уменьшить λ . Благодаря нормировке через ϵ_{00} он слабо зависит от амплитуды функций f_i , и может быть определен с помощью тестов. Разумеется, можно брать λ и из условия минимума /2.1/ по λ , но такой выбор лишен каких-либо преимуществ перед тестовым способом, так как минимум /2.1/, как правило, не совпадает с минимумом расстояния от g_i до математического ожидания f_i . Рис.2 иллюстрирует ход функции g по функции f в общем случае.

Б. Функцию $b(x)$ - сглаженную огибающую $f(x)$ снизу - мы получаем как решение задачи: найти функцию, близкую к $f(x)$ /в большей степени к малым значениям $f(x)$ и в меньшей - к большим/ в смысле интегральной метрики и одновременно имеющую минимальную меру изменчивости.

Математическая формулировка этой задачи имеет вид: найти

$$\min \left\{ \int_{x_1}^{x_2} b'^2(x) dx + \mu \cdot \int_{x_1}^{x_2} w(x) (b(x) - f(x))^2 dx \right\} \quad /2.2/$$

при граничных условиях $b(x_1) = f(x_1)$, $b(x_2) = f(x_2)$, $w(x) = 1/f(x) \cdot /2.2'/$ Здесь μ - величина, аналогичная λ в выражении /2.0/.

Перенормировав μ и записав весовую функцию $w(x)$ в несколько более общем виде, мы получим дискретный аналог /2.2/, /2.2'/: найти

$$\min \left\{ \sum_{i=0}^{M-1} (b_{i+1} - b_i)^2 + \frac{\mu}{w_{00}} \sum_{i=1}^M w_i (b_i - f_i)^2 \right\} \quad /2.3/$$

при граничных условиях $b_0 = f_0$, $b_M = f_M$. /2.3'/

Здесь

$$w_i = 1/(\alpha \cdot |f_i| + \beta \cdot |f_{i+1} - 2f_i + f_{i-1}| + 1), w_{00} = (1 / \sum_{i=0}^M w_i) / (M-2), /2.4/$$

α и β задаются заранее и позволяют при построении весов отдавать предпочтение либо значениям исходной функции, либо ее 2-й производной, либо комбинировать и то, и другое. Величина μ выбирается точно так же, как и λ в /2.1/, и играет аналогичную роль. Заметим, что строить b по f лучше, предварительно сгладив f с помощью фильтра /2.1/.

Рис.2 иллюстрирует ход b по f в общем случае. Завершая этот параграф, можно указать на отличие фильтра /2.1/, /2.1'/ от сплайнового фильтра. В сплайновом фильтре $\epsilon_i \equiv 1$, из-за чего он вносит значительно большие искажения, чем фильтр /2.1/, в такие детали сильно нестационарного тренда, как пики, впадины, точки перегиба и т.д.

§ 3. ВЫЧИСЛИТЕЛЬНАЯ ПРОЦЕДУРА ДЛЯ ФИЛЬТРА /2.3/

Разностное уравнение Эйлера для решения вариационной задачи /2.3/, /2.3'/ имеет вид

$$b_{i+1} - 2b_i + b_{i-1} - \mu \bar{w}_i b_i = -\mu \bar{w}_i f_i, \quad i = 1, \dots, M-1 \quad /3.0/$$

при $b_0 = f_0, b_M = f_M; w_i = w_i / w_{00}$.

Для нахождения его решения применим метод прогонки /4/ как одновременно наиболее быстрый и экономичный в смысле машинной памяти. Представим

$$b_i = C_i b_{i+1} + h_i, \quad C_0 = 0, h_0 = f_0, \quad i = 0, 1, \dots, M-1. \quad /3.1/$$

Подставляя /3.1/ в /3.0/, получим

$$b_{i+1} - 2b_i + C_{i-1} b_i + h_{i-1} - \mu \bar{w}_i b_i = -\mu \bar{w}_i f_i, \quad /3.2/$$

$$b_{i+1} = (2 + \mu \bar{w}_i - C_{i-1}) b_i - \mu \bar{w}_i f_i - h_{i-1}.$$

Сравнивая /3.2/ с /3.1/, найдем

$$C_i = 1/(2 + \mu \bar{w}_i - C_{i-1}), \quad h_i = C_i (\mu \bar{w}_i f_i + h_{i-1}).$$

Последовательно вычисляем все C_i и $h_i, i=1, \dots, M-1$, а затем по формулам /3.1/ /двигаясь в обратном направлении/ все $b_i, i=M-1, \dots, 2, 1$.

§ 4. ВЫЧИСЛИТЕЛЬНАЯ ПРОЦЕДУРА ДЛЯ ФИЛЬТРА /2.1/

Разностное уравнение Эйлера для решения вариационной задачи /2.1/, /2.1'/ имеет 4-й порядок:

$$e_{i+1} g''_{i+1} - 2e_i g''_i + e_{i-1} g''_{i-1} + \bar{\lambda} g_i = \bar{\lambda} f_i, \quad i=1, \dots, M-1 \quad /4.0/$$

при $g_0 = f_0, g_M = f_M, g''_0 = 0, g''_M = 0$.

Здесь

$$g''_i = g_{i+1} - 2g_i + g_{i-1}, \quad \bar{\lambda} = \lambda/\epsilon_{00}, \quad e_i = \epsilon_i.$$

В данном случае построить быстрый и компактный метод для решения задачи несколько сложнее. Обозначим g''_i буквой v_i и перепишем /4.0/ в виде системы разностных уравнений /e/:

$$g_{i+1} - 2g_i + g_{i-1} - v_i = 0,$$

$$e_{i+1} v_{i+1} - 2e_i v_i + e_{i-1} v_{i-1} + \bar{\lambda} g_i = \bar{\lambda} f_i \quad /4.1/$$

при граничных условиях $g_0 = f_0, g_M = f_M, v_0 = 0, v_M = 0$. /4.1'/
Аналог приема /3.1/ для системы /4.1/ имеет вид

$$g_i = a_i^{11} g_{i+1} + a_i^{12} v_{i+1} + b_i^1,$$

$$v_i = a_i^{21} g_{i+1} + a_i^{22} v_{i+1} + b_i^2. \quad /4.2/$$

Из граничных условий /4.1'/ видно, что

$$a_0^{11} = 0, \quad a_0^{12} = 0, \quad a_0^{21} = 0, \quad a_0^{22} = 0, \quad b_0^1 = f_0, \quad b_0^2 = 0.$$

Подставляя /4.2/ в /4.1/ /при значении индекса $i-1$ /, получим

$$g_{i+1} - 2g_i + (a_{i-1}^{11} g_i + a_{i-1}^{12} v_i + b_{i-1}^1) - v_i = 0,$$

$$e_{i+1} v_{i+1} - 2e_i v_i + e_{i-1} (a_{i-1}^{21} g_i + a_{i-1}^{22} v_i + b_{i-1}^2) + \bar{\lambda} g_i = \bar{\lambda} f_i \quad \text{или}$$

$$g_{i+1} = (2 - a_{i-1}^{11}) g_i + (1 - a_{i-1}^{12}) v_i - b_{i-1}^1,$$

$$v_{i+1} = \left(-\frac{e_{i-1}}{e_{i+1}} a_{i-1}^{21} - \frac{\bar{\lambda}}{e_{i+1}} \right) g_i + \frac{2e_i - e_{i-1} a_{i-1}^{22}}{e_{i+1}} v_i + \frac{(\bar{\lambda} f_i - e_{i-1} b_{i-1}^2)}{e_{i+1}}. \quad /4.3/$$

В векторно-матричной записи /4.3/ имеет вид

$$\vec{p}_{i+1} = \vec{A}_i \vec{p}_i + \vec{t}_i, \quad \vec{p}_i = \{g_i, v_i\},$$

откуда

$$\vec{p}_i = \vec{A}_i^{-1} \vec{p}_{i+1} + \vec{A}_i^{-1} \vec{t}_i. \quad /4.4/$$

Сравнивая /4.4/ и /4.2/, получаем рекуррентные формулы для a_i^{kj}, b_i^k :

$$\begin{aligned} b_i^1 &= d_i^{11} b_{i-1}^1 - d_i^{12} (\bar{\lambda} f_i - e_{i-1} b_{i-1}^2) / e_{i+1}, & a_i^{11} &= d_i^{11}, & a_i^{12} &= d_i^{12}, \\ b_i^2 &= d_i^{21} b_{i-1}^1 - d_i^{22} (\bar{\lambda} f_i - e_{i-1} b_{i-1}^2) / e_{i+1}, & a_i^{21} &= d_i^{21}, & a_i^{22} &= d_i^{22}, \end{aligned} \quad /4.5/$$

где $d^{kj} = A_{kj}^{-1}$ - суть элементы обратной матрицы системы

/4.3/:

$$A = \begin{pmatrix} 2 - a_{i-1}^{11} & 1 - a_{i-1}^{12} \\ -(e_{i-1} a_{i-1}^{21} + \bar{\lambda} g_i) / e_{i+1} & (2e_i - e_{i-1} a_{i-1}^{22}) / e_{i+1} \end{pmatrix}.$$

Находим по /4.5/ все a_i^{kj}, b_i^k /движение вперед/, затем по /4.2/ последовательно определяем g_i, v_i /движение назад/.
Функция g_i будет решением, а v_i - его 2-й разностью.

Таблица
Сравнительные характеристики фильтров

критерий статистического типа	χ^2 общий	χ^2 в пик. обл.	Среднее M уклонения	Средняя D уклонения	Дисперсия D	Значение критерия Андерс.
200	0,19	0,2	-0,003	0,58	0,11	-0,48
	0,38	0,4	-0,088	0,74	0,16	-0,30
	0,25	0,2	-0,037	0,82	0,12	-0,71
	0,25	0,2	-0,088	0,65	0,08	-0,53
1000	0,44	0,9	-0,06	0,63	0,23	-0,06
	0,53	1,1	-0,08	0,66	0,23	-0,37
	0,83	1,2	-0,23	1,54	0,90	-0,23
	2,28	3,5	-0,42	2,97	8,54	00,12
2000	0,56	0,8	-0,04	0,52	0,15	-0,22
	0,67	0,9	-0,04	0,57	0,14	-0,29
	1,85	3,0	-0,33	2,17	2,72	-0,01
	5,31	8,1	-0,61	5,46	28,53	0,29
3000	0,45	0,9	-0,04	0,32	0,09	-0,22
	0,75	1,5	-0,13	0,67	0,30	-0,27
	2,58	4,1	-0,39	2,68	7,43	0,11
	7,88	11,6	-0,74	7,53	65,28	0,34
10000	0,47	0,8	-0,03	0,56	0,10	-0,33
	0,53	0,8	-0,07	0,60	0,11	-0,22
	2,83	1,9	-0,38	2,91	7,50	-0,26
	8,93	6,2	-0,77	8,45	73,91	-0,01

Цифры в каждой клетке обозначают соответствующие значения критерия для фильтров /сверху вниз/: ГСФ, ЛСФ, ССФ, ГАФ. Математические ожидания каждого критерия равны нулю, а дисперсия - единице /в случае идеальной фильтрации/.

Кроме описанного глобального фильтра /2.1/, /2.1'/, может использоваться локальный фильтр аналогичной структуры, работающий на основе /4.0/ следующим образом: поочередно составляется уравнение /4.0/ для 5 точек: $i=i_0-2, i_0-1, i_0, i_0+1, i_0+2$ и решается при граничных условиях $g_{i+2}=f_{i+2}; g_{i-2}$ равно уже найденному сглаженному значению /для $i=2, g_0=f_0$ при $i=M-2, g_M=f_M, g_{i-2}$ найденному значению/; так находится g_{i-1}, g_i . Затем i сдвигается на $3(2,1)$ вправо и процедура повторяется, пока не будут исчерпаны все значения i . Каждый раз решается система из 3 уравнений:

$$(4e_{i+1} + e_i + \bar{\lambda})g_{i+1} - 2(e_{i+1} + e_i)g_i + e_i g_{i-1} = \bar{\lambda}f_{i+1} + 2e_{i+1}f_{i+2},$$

$$-2(e_i + e_{i+1})g_{i+1} + (e_{i+1} + 4e_i + e_{i-1} + \bar{\lambda})g_i - 2(e_i + e_{i-1})g_{i-1} =$$

$$= \bar{\lambda}f_i - e_{i+1}f_{i+2} - e_{i-1}g_{i-2};$$

$$e_i g_{i+1} - 2(e_i + e_{i-1})g_i + (e_i + 4e_{i-1} + \bar{\lambda})g_{i-1} = \bar{\lambda}f_{i-1} + 2e_{i-1}g_{i-2}.$$

Решить систему /4.6/ проще всего обращением ее матрицы. Из найденных 3 неизвестных g_{i+1} , g_i , g_{i-1} используются лишь g_{i-1} , g_i . Локальный фильтр экономнее использует память ЭВМ, сравним с глобальным по скорости, но несколько уступает ему по качеству фильтрации.

§ 5. РЕЗУЛЬТАТЫ СРАВНЕНИЯ С ДРУГИМИ ФИЛЬТРАМИ

В таблице проводится сравнение работы фильтров:

- 1/ глобального /4.1/, /4.1'/ /кратко ГСФ/;
- 2/ локального /4.6/ /кратко ЛСФ/;
- 3/ фильтра скользящего среднего /кратко ССФ/;
- 4/ фильтра свертки с функцией типа гауссиан /ГАФ/.

Для проведения такого сравнения были построены серии пуассоновских рядов с известными статистическими характеристиками. Максимумы в сериях составляли соответственно 200, 1000, 2000, 3000, 10000. Все серии имели резонансный тренд, но характер резонансных пиков был различен: острые пики с большой амплитудой для серий с большой статистикой, пологие пики с небольшой амплитудой для серий с малой статистикой, т.е. в последних тренд фактически не был сильно нестационарным. Были вычислены также характеристики:

а/ сумма нормированных на дисперсию квадратов отклонения значения сглаженной серии от истинного значения в каждой точке серии /обозначение в таблице - χ^2 общий/.

б/ то же, но вычисленное лишь в точках серии, содержащих пики тренда /обозначение - χ^2 в пиковой области/.

Далее вычислялось уклонение - разность исходной серии и сглаженной, нормированная на среднеквадратичное отклонение /теоретически в идеальном случае должен получиться белый шум/, и вычислялись характеристики, позволяющие судить о том, насколько уклонение отличается от белого шума. Для этого каждое уклонение разбивалось на 10 участков, для каждого вычислялись среднее, дисперсия и значение критерия Андерсона ^{15/} /по смыслу

близок к выборочному коэффициенту корреляции/ и затем вычислялись средние средних, дисперсий, дисперсий дисперсий и критериев Андерсона для интервалов.

Идеальные значения каждой характеристики - нуль /дисперсии - единица/. Чем больше отличаются характеристики от нуля, тем хуже работает фильтр.

На основе таблицы можно сделать следующие выводы:

1. Фильтры ГСФ и ЛСФ обеспечивают лучшее качество фильтрации во всех отношениях.
2. При фильтрации серий со слабонестационарным трендом ГСФ и ЛСФ хотя и превосходят остальные фильтры по качеству, но все же не дают столь значительных преимуществ, как в случае серий с сильнонестационарным трендом.

В тестовом примере обрабатывался участок реального γ -спектра. Здесь сначала ряд был сглажен по ГСФ, затем по результату сглаживания была построена низкочастотная огибающая. Другими примерами использования указанных выше фильтров были сглаживание спектра и построение фона при реализации алгоритма автоматического поиска пиков в γ -спектрах.

§ 6. ОПИСАНИЕ ПРОГРАММ

Фильтр /4.1./, /4.1'/ реализуется программой SMOOS. Обращение

CALL SMOOS (SA, M, SB, SC, SD, RR, RMODE) .

Здесь SA - массив длиной в M ячеек, содержащий исходный ряд; SB, SC - рабочие массивы;

SD - массив, содержащий сглаженный ряд /длиной M ячеек/, первоначально используется как рабочий массив.

Размерность массивов: SA(2M) , SB(2M) , SC(2M) , SD(2M) .

RR содержит λ /ненормализованное/,

RMODE - указатель режима формирования весов e_i в /2.1/:

RMODE = 0 - обычный режим /по формулам в § 2/,

RMODE = 1. - $e_i \equiv 1$,

RMODE = 2. - $e_i = 1/(1 + r_i^2)$, где массив r_i задается пользователем и располагается в SA вслед за исходным рядом /ячейки M+1, 2M /.

Фильтр /3.0/, а также локальный фильтр /4.6/ реализуется программой SMOSI. Обращение:

CALL SMOSI(SA, M, SB, HA, C8, C9, RR, PR1, F) .

Здесь

SA - массив длиной в M ячеек, содержащий исходный ряд; SB, HA - рабочие массивы. Размерность массивов = SA(M), SB(M), HA(M), SB после работы SMOSI содержит сглаженный ряд.

C8, C9 - коэффициенты α и β в выражении /2.4/.

RR= μ /для фильтра /3.0//, λ - для локального фильтра /4.6/,

PR1 = резервная ячейка для печати,

F = режим работы SMOSI :

F = 0 - реализуется локальный фильтр /4.6/,

F = 1 - реализуется фильтр /3.0/.

В пакет входит вспомогательная программа GRAPD для построения графиков исходного и сглаженных рядов. Обращение к GRAPD:

CALL GRAPD(SA, SB, SC, SD, SE, K, M, RL) .

Здесь

RL = 0 - графики подготавливаются для печати поля шириной в 128 символов /например, на АЦПУ/;

RL = 1 - графики подготавливаются для печати поля шириной в 80 символов /например, при обращении к SMOOS или SMOSI через терминал типа INTERCOM /;

SA, SB, SC - массивы, содержащие ряды, совмещенные графики которых должны быть построены;

K, M - начало и конец /в массиве / ряда для построения графика;

SD, SE - массивы, содержащие дополнительную информацию, которая должна сопровождать каждую точку графика.

§ 7. ТЕСТОВЫЙ ПРИМЕР

В качестве теста программ было осуществлено сглаживание фильтром /4.1/ участка реального γ -спектра $s(x)$ длиной в 100 каналов, в который дополнительно были внесены деформации изолированного типа, повышающие уровень осцилляций участка /обычный участок γ -спектра представляет собой слишком легкий объект для фильтра /4.1//. По результату сглаживания $f(x)$ была построена огибающая снизу $b(x)$, которая, если взять ее проекцию на неравенства

$$b(x) \leq f(x),$$

может рассматриваться как фон, например, в процедуре автоматического поиска пиков в γ -спектре. Совмещенные графики всех 3 функций выданы на АЦПУ в качестве выходной информации.

Обозначения: s - исходный, f - сглаженный ряд, b - огибающая снизу. Исходный ряд распечатан перед графиками, значения f и b в каждом канале печатаются слева от графиков.

ЛИТЕРАТУРА

1. Дженкинс Г.М., Ваттс Д.Г. Спектральный анализ и его приложения. Вып.1,2, "Мир", М., 1972.
2. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов. "Связь", М., 1979.
3. Евграфов М.А. Аналитические функции. "Наука", М., 1968.
4. Бахвалов Н.С. Численные методы. Т.1, "Наука", М., 1975.
5. Wilks S.S. Mathematical Statistics. John Wiley & Sons, Inc. New York, London, 1962.
6. Самарский А.А., Николаев Е.С. Методы решения сеточных уравнений. "Наука", М., 1978.

Рукопись поступила в издательский отдел
15 июля 1980 года.