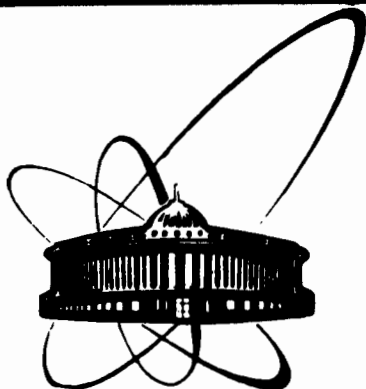


89-378



объединенный
институт
ядерных
исследований
дубна

3-681

P11-89-378

В.Б.Злоказов

АНАЛИЗ
СКРЫТЫХ АНГАРМОНИЧЕСКИХ ПЕРИОДИЧНОСТЕЙ

Направлено в журнал
"Computer Physics Communications"

1989

1. Введение

В работе речь пойдет о поиске скрытых периодичностей самого общего вида в экспериментально регистрируемых функциях и распределениях.

Математическая теория и алгоритмический аппарат выявления скрытых периодичностей имеют уже довольно длинную историю. Практика их применения охватывает весьма большую область прикладной науки, но все описанные в литературе подходы и методы имеют дело исключительно с периодичностями гармонического (синусоидального) характера. Однако данные и распределения, регистрируемые с помощью современной измерительной аппаратуры и описывающие процессы самой разнообразной природы, имеют очень часто ангармонические периодические компоненты. Задача существенно осложняется тем, что в общем случае форма этих компонент неизвестна.

Итак, пусть задано экспериментально измеренное распределение: дискретная функция $f(i)$, определенная на множестве целочисленных индексов $I: i=0, 1, 2, \dots, m$ и являющаяся гистограммным представлением некоторой функции непрерывного аргумента $f(t)$. Будем считать функцию $f(t)$ суммой полипериодической, аperiodической $h(t)$ и шумовой $e(t)$ компонент:

$$f(t) = \sum_{i=1}^n a_i g_i(t, T_i) + h(t) + e(t), \quad (1)$$

где функция $g_i(t, T_i)$ удовлетворяет соотношению

$$g_i(t + T_i) = g_i(t) \quad (2)$$

в области определения $f(t)$.

Методы определения периодичностей, т.е. величин n, a_i и T_i , описанные в литературе [1], могут быть выделены в 2 категории:

- 1/ аппроксимационные, которые аппроксимируют в смысле какой-либо метрики выражение (1) функцией, моделирующей функции g_i и h и зависящей от параметров; значения параметров, получаемые при хорошей аппроксимации, и будут оценками a_i и T_i ;
- 2/ фильтрационные, которые позволяют выделять отдельные компоненты или определять их характерные признаки и определять таким образом интересные нас величины a_i и T_i .

Первые требуют довольно большого объема априорной информации, но дают более качественные результаты; вторые дают меньшую точность, но зато не требуют много априорной информации.

Мы остановимся на вторых, и, в частности, на лучшем из них - методе анализа дискретного преобразования Фурье (ДПФ) функции [2].

Напомним основные определения. Пусть имеет место (1). Для дискретной функции $f(k)$, $k \in I$, ее ДПФ выражается формулой

$$f(v) = \frac{1}{m+1} \sum_{k=0}^m f(k) \exp(-q^{*k} * v), \quad (3)$$

где $q = \sqrt{-1} * 2\pi / (m+1)$, $\pi = 3,14159 \dots$, а $v \in I$ и является целочисленным. В дальнейшем будем $q^{*k} * v$ в экспоненте обозначать как $q^{*k} v$, а константа c , если не оговорено другое, будет обозначать $1 / (m+1)$. Для дискретной функции $g(i)$ запись $g(v)$ будет обозначать ее ДПФ.

2. Анализ и поиск гармонических периодичностей.

Пусть дискретная функция является гистограммным представлением гармонической периодической функции $f(t)$ с периодом T , определенной на отрезке $[t_1, t_2]$, причем $T_L = t_2 - t_1$, где L - целое число:

$$f(t) = a \exp(i 2\pi \left(\frac{t - t_1}{T} \right)), \quad (4)$$

Выберем шаг гистограммирования h и преобразуем отрезок $[t_1, t_2]$ к множеству целочисленных индексов $I: \{0, 1, \dots, m\}$, а под T будем понимать частное T/h (целое число). Очевидно, $m+1 = L \cdot T$.

Дискретное представление (4) имеет вид:

$$f(j) = a \exp(i 2\pi \left(\frac{j * L}{m+1} \right)), \quad (5)$$

Применим теперь к (5) ДПФ. Получим

$$f(v) = \frac{a}{m+1} \sum_{k=0}^m \exp(ckL) \exp(-ckv) = \frac{a}{m+1} \sum_{k=0}^m \exp(ck(L-v)).$$

Легко видеть, что $f(v)$ равна

$$f(v) = \begin{cases} a, & \text{если } L = v. \\ 0, & \text{иначе.} \end{cases}$$

Далее, пусть задана полигармоническая периодическая дискретная функция с периодами T_k и амплитудами a_k , причем $m+1$ кратна всем T_k :

$$f(j) = \sum_{k=1}^n a_k \exp(i2\pi \frac{j * L_k}{m+1})$$

$$m+1 = T_k \frac{L}{k} \quad k=1, \dots, n; \quad (7)$$

ДПФ $f(v)$ будет отличаться от нуля только в точках $v = L_k$, где оно будет равно a_k , и это дает нам простой алгоритм поиска и определения гармонических периодичностей:

1/ построить ДПФ дискретной функции;

2/ найти точки v_k , в которых ДПФ отлична от нуля, и оценить периоды и амплитуды по формулам:

$$T_k = (m+1) / v_k \quad (8)$$

$$a_k = f(v_k).$$

Автоматический поиск пиков в модуле ДПФ целесообразно вести с помощью метода, описанного в работе [3].

Рассмотренная ситуация является слишком идеализированной. Реальная ситуация, как правило, осложнена двумя следующими обстоятельствами:

1/ нарушение условия (7);

2/ наличие неперiodической и шумовой примесей в функции $f(t)$.

Эти два обстоятельства приводят к тому, что контур ДПФ претерпевает более или менее сильные деформации по сравнению с рассмотренным выше. Тем не менее, если влияние осложняющих факторов не является чрезмерно

сильным, в точках v_k модуль ДПФ будет иметь пики, возвышающиеся над окрестными значениями $|f(v)|$.

С учетом сказанного можно следующим образом модифицировать вышеописанный алгоритм для придания ему большей реалистичности:

- 1/ выбрать m такое, чтобы выполнялось (7), и построить ДПФ дискретной функции $f(x)$, $i=0, \dots, m$;
- 2/ найти точки v_k , являющиеся центрами пиков в модуле ДПФ, и оценить периоды и амплитуды гармонических периодичностей по формулам:

$$T_k = (m+1)/v_k \quad (8)$$

$$a_k = f(v_k)$$

Таким образом, обнаружение гармонических периодичностей даст нам набор пар: $\{a_i, T_i\}$ — амплитуд и периодов гармонических периодичностей. Погрешности периодов можно оценить, например, с помощью полуширин пиков: точки по обе стороны пика, в которых амплитуда пика спадает до половины максимума, пересчитываются в периоды по формуле (8) и тем самым дают доверительный интервал для соответствующего периода.

Распознавание гармонических периодичностей по модулю преобразования Фурье является, безусловно, самым мощным и эффективным методом поиска периодичностей и существенно превосходит такие простые методы, как метод Буй-Балло, периодограммные методы Шустера и т. д. [1].

3. Анализ и поиск ангармонических периодичностей

А теперь рассмотрим вопрос о ДПФ от периодических ангармонических гистограмм, т. е. гистограмм, не являющихся линейной комбинацией функций $\langle \exp(ikx) \rangle$ с соизмеримыми периодами, но в то же время удовлетворяющих условию периодичности.

Итак, пусть дана гистограмма $h(x)$, $i=0, \dots, m$ и пусть существует T такое, что

$$h(x+T) = h(x)$$

для любого индекса из области определения. Будем предполагать, что $m+1$ кратно T :

$$m+1 = T * L \quad (9)$$

Пусть P — множество индексов $0, T, 2T, 3T, \dots$ и $h(x)$ — его характеристическая гистограмма. Рассмотрим ДПФ от этой функции.

ЛЕММА 1. $h(v)$ является периодической гистограммой с периодом L и характеристической гистограммой множества индексов $0, L, 2L, 3L, \dots$ с амплитудой, равной $1/T$.

Доказательство следует из таких равенств:

$$h(v) = \frac{1}{m+1} \frac{1 - \exp(-qvTL)}{1 - \exp(-qvT)} = 0 \quad \text{при } vT, \text{ не кратном } m+1;$$

$$h(v) = \frac{1}{m+1} \sum_{i=0}^m h(i) = \frac{m+1}{(m+1)T} = \frac{1}{T} \quad \text{при } vT, \text{ кратном } m+1.$$

Но vT кратно $m+1$ при $v=0, L, 2L, \dots$, откуда следует утверждение ЛЕММЫ 1.

ЛЕММА 2. Если $s(v)$ получена из $h(v)$ сдвигом на k ячеек ($k < T$), то $s(v)$ будет пропорциональна $h(v)$ с некоторым комплексным коэффициентом, равным по модулю 1.

Доказательство. Достаточно рассмотреть случай $k = 1$. На основании (3) мы можем записать

$$s(v) = \frac{1}{m+1} \sum_{k=1}^m h(k-1) \exp(-qv) = \frac{\exp(-qv)}{m+1} \sum_{k=0}^m h(k) \exp(-qv),$$

т.е. $s(v)$ отличается от $h(v)$ на множитель $\exp(-qv)$, равный по модулю 1, что и требовалось доказать.

ТЕОРЕМА 1. ДПФ от периодических ангармонических гистограмм с периодом T , удовлетворяющим условию (9), может отличаться от нуля лишь на множестве индексов $0, L, 2L, \dots, (T-1)L$ (обозначим его Ω).

Доказательство. Так как любая такая периодическая ангармоническая гистограмма $g(v)$ может быть представлена как линейная комбинация T сдвигов характеристической функции множества P , то на основании предыдущих лемм ДПФ такой гистограммы будет равняться 0 вне множества индексов Ω , и линейной комбинации мнимых экспонент на Ω :

$$g(v) = \frac{1}{T} \sum_{k=0}^{T-1} g(k) \exp(-gkv).$$

ПРИМЕЧАНИЕ. Если соотношение (9) не выполняется, ДПФ от периодических ангармонических гистограмм будет иметь в общем случае все компоненты, не равные нулю. Картина будет тем отличнее от рассмотренной, чем больше остаток от деления $m+1$ на T . Тем не менее, качественное сходство сохранится: модуль ДПФ от характеристической функции P будет иметь пики в индексах (или их окрестности) множества Q .

Если гистограмма является суммой периодических ангармонических компонент с периодами T_i , причем $(m+1) = T_i \cdot L_i$, то ДПФ этой гистограммы может отличаться от нуля на объединении множеств индексов

$$Q = \cup Q_i: \quad 0, L_i, 2L_i, \dots, (T-1)L_i.$$

Если гистограмма $h(x)$ кроме периодических ангармонических компонент содержит еще аperiodическую и шумовую компоненты, а также если не выполняется условие кратности $m+1$ периодам, ее ДПФ не будет иметь столь четко выраженной структуры. Но если влияние искажающих факторов окажется не очень сильным, то на множестве индексов Q модуль ДПФ будет иметь более или менее четко выраженные пики.

Это подсказывает алгоритм обнаружения скрытых ангармонических периодичностей.

- 1/ Провести гармонический анализ гистограммы, например, выявив пики в ее модуле ДПФ;
- 2/ Отобрав представляющие интерес периодичности, выбрать оптимальное m , при котором влияние не кратности m периодам минимально.
- 3/ Построить ДПФ гистограммы и попытаться построить множество Q для его модуля.
- 4/ Проанализировать структуру множества Q и проверить правдоподобность гипотез о возможных ангармонических периодах.

При этом неразличимы случаи, когда гистограмма содержит периодическую ангармоническую компоненту с периодом T , и когда гистограмма содержит набор периодических гармонических компонент с периодами, для которых справедлива следующая

ТЕОРЕМА «о соответствии ангармонических и гармонических периодов».

Ангармонический период T обязательно сопровождается набором гармонических периодов T_i , связанных между собой соотношением:

$$T_1 = T$$

$$T_i = T_{i-1} * \frac{i-1}{i} \quad i=2, \dots, L+1$$

(10)

Доказательство следует из предыдущих утверждений.

4. Применение метода

Описанная методика была применена к конкретным данным:

- 1/ временному распределению суточных регистраций концентрации радона в воде глубоких скважин участка зоны Передового разлома Копетдага;
- 2/ временному распределению суточных регистраций атмосферного давления в том же районе;
- 3/ к симулированным тестовым данным.

Длина распределений приблизительно равна 1500 дней и не является степенью двух. Для ускоренного вычисления ДПФ была составлена специальная программа (SFFT - SemiFast Fourier Transform), которая либо использует заданное разбиение числа m на составные множители m_1 и m_2 ($m = m_1 m_2$), либо сама подбирает оптимальные m_1 и m_2 , и реализует алгоритм ускоренного преобразования Фурье числом требуемых операций порядка $m(m_1 + m_2)$ (2).

Результаты анализа вышеуказанных распределений были следующими.

- 1/ Поскольку в автоматическом режиме не опознавалась естественная для таких распределений годовая периодичность, была проведена оптимизация длины распределения. Она была выбрана пропорциональной периоду 365, т.е. 1460 чисел, и модуль ДПФ был пересчитан. На сей раз анализ дал такие оценки.

Вес	Период	Ошибка	Возможные ангармонические "хвосты"															
121	378	347	189	126	96	76	63	54	47	42	36	34						
			32	29	27	25	24	22	21	20	19	18						
22	140	41	70	47	38	28	23	20	17	16	14	13						
30	125	33	63	42	31	25	21	18	16	14	13	11						
32	93	11	47	31	23	19	16	13	12	10	9	8						

32	80	13	40	27	20	16	13	12	10	9	8	7
20	66	9	33	22	17	13	11	9	8	7	6	5
15	58	7	29	20	15	12	10	8	7	6	5	4
10	54	3	27	18	14	11	9	8	7	6	5	4
13	50	5	25	17	12	10	8	7	6	5	4	3
7	43	2	21	14	11	9	7	6	5	4	3	2
18	40	2	20	13	10	8	7	6	5	4	3	2
10	36	1	18	12	9	7	6	5	4	3	2	
12	34	2	17	11	8	7	6	5	4	3	2	
11	32	2	16	11	8	6	5	4	3	2		
27	30	1	15	10	8	6	5	4	3	2	1	
10	29	2	14	10	7	6	5	4	3	2	1	
10	27	1	14	9	7	5	4	3	2	1		
8	26	1	13	9	6	5	4	3	2	1		
12	24	1	12	8	6	5	4	3	2	1		
20	23	1	12	8	6	5	4	3	2	1		
7	19	1	10	6	5	4	3	2	1			
6	18	1	9	6	5	4	3	2	1			
6	17	1	8	6	4	3	2	1				
14	16	1	8	5	4	3	2	1				
7	14	1	7	5	4	3	2	1				
8	13	1	7	4	3	2	1					
11	12	1	6	4	3	2	1					
8	11	1	6	4	3	2	1					
6	10	1	5	3	2	1						
7	9	1	5	3	2	1						
9	8	1	4	3	2	1						
10	6	1	3	2	1							

Эту таблицу можно интерпретировать следующим образом:

- а/ Видна, хотя и очень нечетко, годовая периодичность (378 ± 347) и для всех ее ангармонических "хвостов" есть соответствующие кандидаты в таблице, что доказывает резко ангармонический характер этой периодичности:
- б/ Из-за значительного превышения веса (27) над соседями нелогично считать периодичность (30 ± 1) "хвостом"; это явно самостоятельная периодичность, и для ее ангармонических "хвостов" тоже есть соответствующие кандидаты в таблице, что тоже доказывает ее ангармонический характер.

- с/ Можно предположить, хотя и с меньшей убедительностью, что самостоятельной является и периодичность (23 ± 1) ;
- д/ Периодичности (16 ± 1) и (6 ± 1) являются, по-видимому, все же "хвостами" периодичностей 378 и 30: во всяком случае вопрос о них остается спорным.

2/ По указанным выше причинам длина распределения была положена равной 1460. Результаты анализа модуля ДЛФ таковы.

Вес	Период	Ошибки	Возможные ангармонические "хвосты"												
			183	122	91	73	61	52	46	41	37	33			
81	365	104	30	28	26	24	23	21	20	19	18	17			
26	175	43	88	58	44	35	29	25	22	19	18	16			
24	133	37	67	44	33	27	22	19	17	15	13	12			
29	98	20	49	33	25	20	16	14	12	11	10	9			
16	88	16	44	29	22	18	15	13	11	10	9	8			
17	77	6	38	26	19	15	13	11	10	9	8	7			
18	71	10	36	24	18	14	12	10	9	8	7	6			
15	64	3	32	21	16	13	11	9	8	7	6	5			
17	55	6	28	18	14	11	9	8	7	6	5	4			
14	48	6	24	16	12	10	8	7	6	5	4	3			
18	42	2	21	14	10	8	7	6	5	4	3	2			
13	37	1	19	12	9	7	6	5	4	3	2				
12	34	1	17	11	8	7	6	5	4	3	2				
15	32	2	16	11	8	6	5	4	3	2					
21	29	2	15	10	7	6	5	4	3	2	1				
16	27	2	13	9	7	5	4	3	2	1					
9	26	1	13	9	6	5	4	3	2	1					
14	23	1	12	8	6	5	4	3	2	1					
17	22	1	11	7	6	4	3	2	1						
22	20	1	10	7	5	4	3	2	1						
16	19	1	10	6	5	4	3	2	1						
18	18	1	9	6	4	3	2	1							
18	16	1	8	5	4	3	2	1							
11	15	1	8	5	4	3	2	1							
17	14	1	7	5	3	2	1								
14	13	1	7	4	3	2	1								
13	12	1	6	4	3	2	1								

11	11	1	6	4	3	2	1
14	10	1	5	3	2	1	
15	9	1	5	3	2	1	
15	8	1	4	3	2	1	
12	7	1	3	2	1		
16	6	1	3	2	1		
3	3	1	2	1			

Эту таблицу можно интерпретировать следующим образом:

- а/ Хорошо просматривается годовая ангармоническая периодичность 365 ± 104 ;
- б/ Весьма правдоподобно предположение о наличии ангармонических периодичностей 98 ± 20 , 29 ± 2 ; значительно менее убедительно предположение о периоде 20 ± 2 ; об остальных периодах сказать ничего нельзя.
- э/ Симулировалась сумма следующих компонент:
 - а/ ангармонических периодичностей $(f(t) = A \cdot \max(\sin(ck t), 0)^4)$ с периодами 365, 29, 7;
 - б/ гармонической периодичности с периодом 90;
 - в/ пикообразных функций, константы и случайного шума.
 Результаты были следующие.

Вес	Период	Ошибка	Возможные ангармонические "хвосты"										
			183	122	91	73	61	52	46	41	37	33	
36	365	104	30	28	26	24	23	21	20	19	18	17	
			16	15	14	13	12	11	10	9			
			91	61	46	37	30	26	23	20	18	17	
23	183	25	15	14	13	12	11	10	9	8	7	6	
			46	31	23	18	15	13	12	10	9	8	
113	92	22	39	26	20	16	13	11	10	9	8	7	
20	78	12	34	23	17	14	11	10	9	8	7	6	
12	68	9	14	10	7	6	5	4	3	2	1		
36	29	1	7	5	4	3	2	1					
27	14	0	6	4	3	2	1						
3	13	0	5	3	2	1							
11	10	0	4	2	1								
23	7	0	2	1									
12	4	0	1										
10	2	0											

Литература

1. Серебрянников М. Г., Первозванский А. А.
Выявление скрытых периодичностей. М., Наука, 1963.
2. Бахвалов Н. С. Численные методы. Т. 1. М., Наука, 1975.
3. Злоказов В. Б.
Математические методы анализа экспериментальных спектров и спектрореподобных распределений. ФЭЧАЯ, 1988, т. 16, вып. 5, стр. 1126-1163.

Рукопись поступила в издательский отдел
29 мая 1989 года.