УДК: 519.6

ФЕНОМЕНОЛОГИЧЕСКИЙ ПОДХОД К МОДЕЛИРОВАНИЮ ЯВЛЕНИЙ

З.Е. Филер, А.Н. Дреев
Кировоградский государственный педагогический университет
Кировоградский национальный технический университет

Введение. При первичном изучении явления, опираясь на данные наблюдений, принимают ту или иную модель в виде регрессионного уравнения зависимости значений числовой характеристики явления от времени или от другой величины. Широкий набор математических пакетов (Maple, MathCAD, MatLab, Mathematic, …) используют различные методы приближения, в том числе и минимизацию среднеквадратического отклонения на заданном базисе функций. Однако, набор функций для базиса построения регрессионного уравнения, даже определённого класса, должен вручную определять пользователь. Автоматизировать процесс удалось для узкого набора функций; например, широко используемый пакет EXCEL предлагает использовать часто встречающиеся законы: показательный, степенной, логарифмический, в виде многочлена (от первой до шестой степени). Другие пакеты позволяют также производить гармонический анализ, который даёт уравнения сумм кратных гармоник, которые циклично продолжают экспериментальные данные. Однако, все такие формы регрессионного уравнения не всегда пригодны к изучению сложных почти периодических процессов, на один-два основных периода.

Авторы разработали алгоритм, обобщающий метод рядов Фурье. В классическом методе уравнение регрессии удовлетворяет требованию минимальности квадратического отклонения значений временной последовательности от n-ой частичной суммы тригонометрического ряда с данной основной частотой ω и кратными частотами $k\omega$ на интервале длиной T=2π/ω.

Он основан на последовательном выделении моногармонических составляющих типа

$$g(x) = A + B\sin(\omega x) + C\cos(\omega x)$$

с искомыми числами A, B, C.

В общем случае условия минимальности суммы квадратов отклонений

$$S_n(A,B,C,\omega)=\sum_{i=1}^{n}(g(x_i)-y_i)^2$$

нелинейные относительно ω. В методе Фурье минимальная (основная) частота ω задаётся, и уравнения для A, B, C линейны.

Предлагается такая *процедура* нахождения A, B, C и ω:

- из априорных соображений (физического или экономического смысла) определяется диапазон возможных значений частот (ω_0; Ω); он делится на равные части точками ω_k, количество частей N выбирается из условия $N=3\frac{\Omega-\omega_0}{2\pi}(t_n-t_0)$; при этом учитываются локальные минимумы возле частот, кратных времени наблюдений (рис. 1);

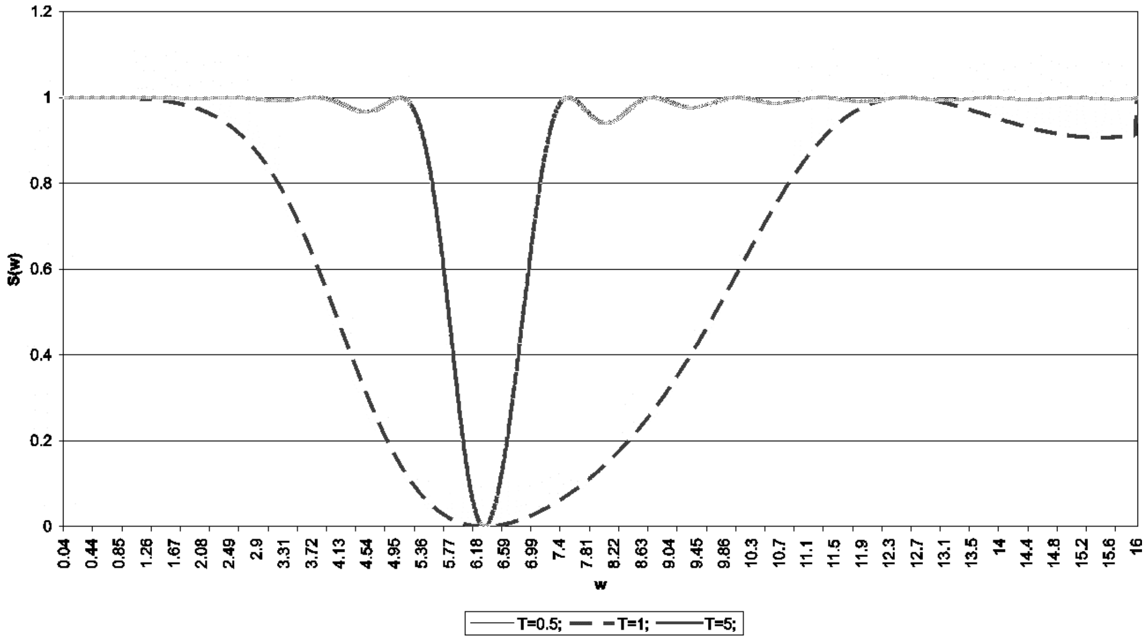

Рис. 1. Зависимость минимумов Фурье-частот от времени наблюдения

- для каждого значения ω_k решаются нормальные уравнения

$$\frac{\partial S}{\partial A}=\frac{\partial S}{\partial B}=\frac{\partial S}{\partial C}=0;$$

- для найденных A, B, C и принятого ω_k определяется минимальная сумма S_n^1;

- выбирается наименьшая из найденных сумм, которая и определяет первую из гармоник искомого регрессионного уравнения; ω_k уточняется как численное решение задачи $S_n^1(\omega)\to\min$;

- таблица экспериментальных значений y_k заменяется новой таблицей $z_k = y_k - g(x_k)$.

С найденной таблицей значений z_k всё повторяется сначала до определения новой гармоники типа $g(x)$. Процесс последовательного определения гармоник $g_k(x)$ прекращается при практическом повторении суммарных квадратических отклонений S_n^k и прекращении их существенного уменьшения.

На основе описанного алгоритма А.Н.Древым была разработана *программа* EXTRAPOL с *инструкцией* по её применению для пользователей, доступной даже для студентов младших курсов. Она совместима с массивами, заданными в EXCEL и позволяет строить как графики заданных массивов, так и их приближений с помощью построенного тригонометрического тренда.

Верхняя частота Ω определяется из шага h таблицы аргумента (критерий Котельникова-Найквиста), т.к. наименьший период $T_{мин} \geq 4h$, $\Omega \leq 2\pi/T_{мин}$. Низшая частота выбирается из условия $\omega_0 \geq 2\pi/L$, где L – 4…6 длительностей диапазона наблюдений. Это соответствует тому, что этот диапазон принимается за «длинный» период. Наличие непериодического тренда (в среднем монотонного) можно рассматривать как отрезок не менее ¼ «длинного периода».

Программа сама рассчитывает значимое число гармоник и сообщает их пользователю как рекомендуемое. Однако, он может и увеличить число искомых гармоник, добиваясь большей близости отыскиваемого тренда от заданного массива. При наличии в анализируемом сигнале большой случайной составляющей, полученное регрессионное уравнение даёт малонадёжные высокочастотные колебания для экстраполяции – прогноза, найденного с помощью EXTRAPOL.

Ясно, что неявно для прогноза предполагается *эргодичность* случайного процесса и его почти периодичность, когда частоты почти детерминированы, а соответствующие амплитуды являются случайными величинами. Тогда прогноз хорошо отражает *фазы* процесса (этапы нарастания, спада, максимума и минимума).

Инструкция для пользователя

Исходные данные - две колонки чисел: 1 - значения аргумента (числа!); 2 – значения элементов массива, который анализируют. Пара «аргумент-значение» в текстовом формате должна занимать отдельную строку и разделяться пробелами или/и знаком табуляции.

Последовательность работы с программой:

1. Выделить исходные данные и скопировать их в буфер обмена.

2. Перейти к программе для анализа, и в левом окне вставить исходные данные. (правой кнопкой мыши - "вставить")

3. Внизу, ползунком задаём желаемое число гармоник.

4. Кликаем кнопку "Аналізувати". Ждём результата.

5. Для построения прогноза кликаем на кнопку "Розрах. прогн."

6. Устанавливаем начальное и конечное значения периода прогнозирования, а также шаг между последовательными значениями.

7. Кликаем "пуск".

8. Выделяем результат, через буфер обмена вставляем его в Excel.

9. Закрываем программу.

Оценка погрешности экстраполяции

В процессе построения регрессионного уравнения усредняются значения частот и амплитуд гармоник исследуемого сигнала. Поэтому прогноз может иметь значительную ошибку, которую необходимо оценить.

Процесс построения базиса для регрессионного уравнения даёт значения циклических частот колебаний, на основе которых находиться характеристическое уравнение. Соответствующее дифференциальное уравнение содержит в множестве решений и построенное приближение:

$$y^{(n)} + b_{n-1}y^{(n-1)} + ... + b_2 y'' + b_1 y' + b_0 y = 0 \qquad (1)$$

Соответственно, по обобщённой формуле Тейлора [1] для найденного приближения можно построить остаточный член:

$$R_n(t) = \int_0^{t-t_i} u_n(\tau) y^*(t-\tau) d\tau, \qquad (2)$$

где $u_n(0) = 0, u'_n(0) = 0, ..., u_n^{(n-1)}(0) = 0, u_n^{(n)}(0) = 1$ – частное решение (1), $y^*(t)$ – результат действия дифференциального оператора (1). Из (2) можно получить оценку погрешности интерполяции:

$$R_n(t) \le \max_{t \in [t_i; t_{i+1}]} |u_n(t)| \cdot \max_{t \in [t_i; t_{i+1}]} |y^*(t)| h,$$

и экстраполяции на время Δt:

$$R \le \max_{t \in [t_1; t_n]} |u_n(t)| \cdot \max_{t \in [t_1; t_n]} |y^*(t)| \Delta t. \qquad (3)$$

Оценка погрешности справедлива при малых изменениях за время Δt свойств объекта исследования (условие эргодичности), и учитывает наличие шума.

Концептуальный прогноз, при котором допускаются сравнительно большие сдвиги по фазе (момента наступления события), имеет значительно меньшую погрешность, чем найденная при помощи оценки (3).

Отличительной чертой метода последовательного извлечения тригонометрических трендов есть возможность обнаружить периоды в сигнале, которые длиннее области наблюдения в несколько раз.

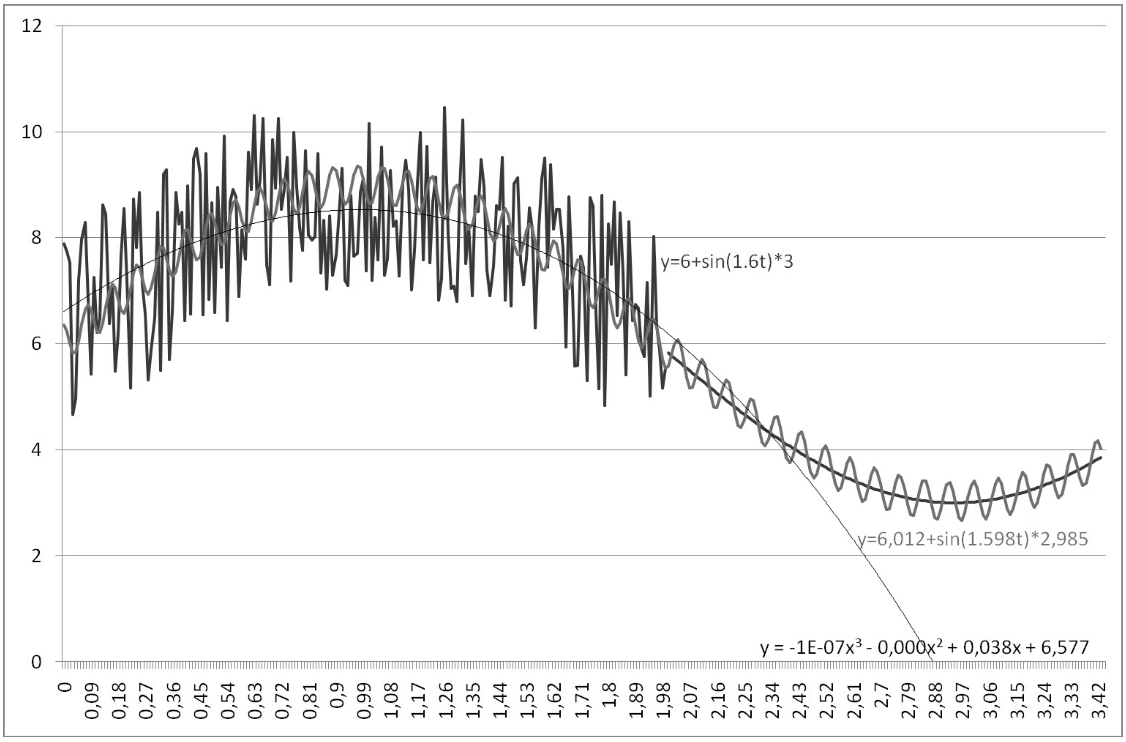

Рис. 2. Построение прогнозов зашумлённого сигнала

Для иллюстрации возможности метода последовательного извлечения тригонометрических трендов построен график рис.2, на котором изображена функция y=6+3sin(1,6t). На промежутке [0;2] сигнал сильно зашумлён. Для зашумлённого участка в Exel построен тренд-прогноз (выбрано по степеням наилучшее соответствие), программой Exstrapol построено тригонометрическое продолжение, которое показывает значительно лучшее соответствие не только зрительно, но и по значениям коэффициентов первой гармоники.

Література

1. Филер З.Е. Об одном обобщении формулы Тейлора и её применении к решению дифференциальных уравнений//УМЖ, 1981, т. 33, в.1. – С. 123 – 128.
2. Дрєєв О.М., Філер З.Ю. Спектральний аналіз майже періодичних сигналів// Теорія та методика навчання математики, фізики, інформатики. Збірник наукових праць. Випуск 4. Том 3. - Кривий Ріг. Видавничий відділ НМетАУ, 2004. – С.64-68.
3. Дрєєв О.М., Філер З.Ю. Частотно-амплітудний аналіз майже періодичних процесів та його застосування для прогнозів// Теорія та методика навчання математики, фізики, інформатики. Збірник наукових праць. Випуск 5. Том 3. - Кривий Ріг. Видавничий відділ НМетАУ, 2005. – С.78-82.

Получено 27.05.09