Каждое решение, полученное с помощью генетического алгоритма, будет иметь следующую структуру:

1. Точка в пространстве параметров (фенотип):

$$X = (X_1,\ X_2,...,X_M) \in T \text{ из } R^N. \tag{6}$$

2. Бинарная строка s фиксированной длины, однозначно идентифицирующая гиперкуб пространства параметров (генотип):

$$s = (l_1,\ l_2,...,l_I) \in S. \tag{7}$$

где S − пространство представлений − бинарных строк длины I.

3. Скалярная величина μ, соответствующая значению целевой функции в точке t (приспособленность):

$$\mu = \lambda(t). \tag{8}$$

Выводы

Таким образом, применение метода ГАО позволяет достаточно эффективно решить задачу синтеза оптимальной структуры программных модулей.

ЛИТЕРАТУРА

1. Батищев Д.И. Генетические алгоритмы решения экстремальных задач. − Учебное пособие.- Нижний Новгород, 1995. − 63с.

2. Дубинин Н.П. Общая генетика. − М: Наука, 1976. − 572с.

3. Скурихин А.Н. Генетические алгоритмы.// Новости искусственного интеллекта, 1995. − №4. − с.6-46.

4. Исаев С. Популярно о генетических алгоритмах http://home.od.ua /~relayer/algo/neuro/ga-pop/.

5. Teo M.Y., Khoo L.P., Sim S.K. Application of Genetic Algorithms to Optimise Neocognitron Network Parametrs. − In Neural Network World, Vol.7, No 3. 1977. PP. 293-303.

УДК 62-505:504.064

Туленбаев Мурат Сауранбаевич – к.т.н., доцент (Тараз, ТарГУ)

СПЕКТРАЛЬНЫЕ ПРЕДСТАВЛЕНИЯ И ПЕРВИЧНАЯ ОБРАБОТКА СИГНАЛОВ ХИМИКО-АНАЛИТИЧЕСКИХ КОМПЛЕКСОВ ЭКОМОНИТОРИНГА

Информационная система экологического мониторинга обычно осуществляет съем данных об окружающей среде с передвижных и стационарных пунктов наблюдений, оснащенных высокоинформативными химико-аналитическими комплексами (ХАК), включающими хроматографы, спектрометры, рентгенофлуориметры и др. Определяющим в их использовании является интеллектуальная составляющая химического анализа (методики, программный продукт, эргономика). В связи с этим, вопросы разработки математического, программного и метрологического обеспечения ХАК остаются весьма актуальными.

Выходной сигнал аналитического прибора ХАК y(t) в большинстве случаев можно рассматривать как аддитивную смесь полезного сигнала s(t,**l**), помехи n(t) и базисного сигнала (или т.н. дрейфа) d(t)[1]:

$$y(t) = s(t,\mathbf{l}) + n(t) + d(t), \tag{1}$$

где $\mathbf{l_0} = \{l_0, l_1, l_2, \ldots, l_{N-1}\}$ – вектор N параметров сигнала, подлежащих оценке; t – независимая переменная. Параметр l_0 выделен из остальных и назван существенным, так как характеризует положение компонента на оси развертки **l** и позволяет различить компоненты между собой.

Модели информативной (полезной) составляющей сигнала чаще принято представлять в виде суперпозиции М отдельных компонентов:

$$s(t,\mathbf{l_0}) = \sum_{m=1}^{M} A_m f_m(t, l_c, l), \tag{2}$$

где A_m – интенсивность сигнала, мерой которой может служить его амплитуда, площадь или энергия; $f(t_c, l_c, \mathbf{l})$ - функция, описывающая форму сигнала (пика), может быть треугольной, трапецеидальной, Гауссовой, Лоррентцовой или другой более сложной, не исключено и табличное задание.

Обобщением гармонического анализа является применение для представления сигнала s*(t)* в виде ряда или интеграла других систем функций. Такими функциями, в принципе могут быть любые системы линейно независимых функций *{u(k,t), k=1,2,…,N}*. Дискретное преобразование выходного сигнала аналитического прибора *y(t)* будем рассматривать как разложение непрерывного сигнала на конечном интервале времени по системе непрерывных функций *{u(k,t)}*. Совокупность коэффициентов такого разложения является дискретным аналогом сигнала – его спектром.

Задавшись дополнительными требованиями к базисной системе, такими как некоррелированность отсчетов спектра, максимизация отношения сигнал/шум с учетом априорных данных о модели сигнала $f(t_c, l_c, \mathbf{l})$ и шума, можно спроектировать адаптивную базисную систему (АБС) по методике, изложенной в статье [1]. В работе предложен базис близкий к оптимальному базису Карунена-Лоэва, когда систему линейно-независимых функций $\{U(k,t)\}$ выбирают из условия согласованной фильтрации:

$$\int_T B(t,\tau)U(k,\tau)d\tau = f(k,t) \tag{3}$$

На их основе, воспользовавшись процедурой, аналогичной ортогонализации по Грамму-Шмидту [2], конструируется требуемая система базисных функций {φ(k,t)Φ(k,t)}:

$$\Phi(k,t) = U(k,t) - \sum_{k=\tau}^{k-1} \gamma_{k\tau}\Phi(\tau,t) = \sum_{k=\tau}^{k-1} d_{k\tau}U(\tau,t) \tag{4}$$

$$\varphi(k,t) = \frac{1}{\sigma_k^2}\int_T B(t,\tau)\Phi(k,\tau)d\tau = \frac{1}{\sigma_k^2}\sum_{\tau=1}^{k} d_{k\tau} f(\tau,t), \tag{5}$$

где $B(t,\tau)$- корреляционная функция шума, а σ_k^2 имеет смысл дисперсии <u>к</u>-ой спектральной составляющей:

$$\sigma^2{}_k = \iint_T B(t,\tau)\Phi(k,t)\Phi(k,\tau)dtd\tau \tag{6}$$

Коэффициенты ортогонализации $\gamma_{k\tau}$ для рекуррентной схемы вычислений:

$$\gamma_{k\tau} = \int_T U(k,t)\varphi(\tau,t)dt = \frac{1}{\sigma_2^2}\int_T f(k,\tau)\Phi(\tau,\tau)d\tau \quad, k>\tau$$

и для явной схемы:

$$d_{k\tau} = -(\gamma_{k\tau} - \sum_{i=\tau}^{k-1}\gamma_{ki}d_{i\tau}) \;,\; \tau = \overline{1,k-2}\,;\; d_{k,k-1} = -\gamma_{k,k-1} \;;\; d_{kk} = 1 \tag{7}$$

1. Представление сигналов в адаптивных базисных системах (АБС)

С помощью АБС (4)-(5) выходной сигнал аналитического прибора y(t) может быть представлен в виде линейной комбинации:

$$y(t) = \sum_{k=1}^{N} Y(k)\varphi(k,t)\,, \tag{8}$$

где Y(k)- спектр сигнала:

$$Y(k) = \int_T y(t)\Phi(k,t)dt = Y_0(k) - \sum_{\tau=1}^{k-1} \gamma_{k\tau} Y(\tau) \tag{9}$$

или

$$Y(k) = \sum_{\tau=1}^{k} d_{k\tau} Y_0(\tau)\,, \tag{10}$$

где величину $Y_0(\tau)$ назовем обобщенным отсчетом:

$$Y_0(k) = \int y(t)U(k,t)dt \tag{11}$$

Полученные выражения (8)-(11) являются практическими алгоритмами дискретного представления поступающей в обработку смеси.

В указанные формулы входят элементы матриц ортогонализации $\boldsymbol{\gamma}$ и ***d,*** которые можно синтезировать заранее, с использованием универсальных компьютеров, при разработке математического обеспечения по априорным данным о модели компонента сигнала и корреляционной функции шума.

Формула (11) для определения обобщенного отсчета является хорошо известным алгоритмом работы согласованного фильтра или корреляционного приемника. Однако в задаче дискретного представления он дополнен рекуррентными процедурами (9) для вычисления спектра.

2. Функция правдоподобия выборки спектральных составляющих

Статистические методы обработки смеси полезного сигнала с помехой основываются на анализе совокупности значений i-ой выборки из этой смеси, образующих многомерный случайный вектор Y_i[3]. Задачей обработки является принятие решения о наличии в выборке компонента полезного сигнала и оценка его параметров $\vec{l}$. Алгоритмы такой обработки строятся на основе плотности распределения вероятностей W(Y) вектора Y_i .

Вектор Y_i образован спектральными составляющими вида (9) или (10) и включает кроме чисто случайного вектора помехи Ξ_i, также и квази детерминированные векторы сигнала Si и дрейфа Di:

$$Y_i=S_i+D_i+\Xi_i \tag{12}$$

Поэтому в выражение для функции $W(Y_i)$ должен входить центрированный вектор спектральных составляющих $\overset{0}{Y}_i$. Центрирование этого вектора производится путем вычитания из Y_i моделей $F_i(\vec{l})$ и D_i, параметры которых после оценки и принимаются за истинные значения параметров сигнала и дрейфа. Следовательно, при наличии полезного сигнала вектор $\overset{0}{Y}_i$ выражается как $\overset{0}{Y}_i = Y_i - F(\vec{l}) - D_i$, а при отсутствии сигнала - как $\overset{0}{Y}_i = Y_i - D_i$. Так как модель сигнала всегда задана в пределах выборки и сохраняется неизменной от выборки к выборке, индекс у $F(\vec{l})$ опущен.

Для построения функции правдоподобия будем полагать, что распределение случайного вектора Y_i нормально, так как в соответствии с центральной предельной теоремой, преобразование (9) нормализует шум. Тогда функция правдоподобия выборки из n спектральных составляющих в АБС при наличии полезного сигнала и его отсутствии имеет соответственно вид [4]:

$$W(Y/\vec{l}) = [(2\pi)^{n/2}\prod_{k=1}^{n}\sigma_k]^{-1}\exp\{-\frac{1}{2}(Y_i - F(\vec{l}) - D_i)^T B^{-1}(Y_i - F(\vec{l}) - D_i)\},$$

$$W(Y/\vec{l}) = [(2\pi)^{n/2}\prod_{k=1}^{n}\sigma_k]^{-1}\exp\{-\frac{1}{2}(Y_i - D_i)^T B^{-1}(Y_i - D_i)\}, \quad (13)$$

где B- диагональная корреляционная матрица размером n×n с элементами $B_{kk}=\sigma^2_k$; $(Y_i - F(\vec{l}) - D_i)$ - вектор столбец центрированных значений спектральных составляющих; $(Y_i - F(\vec{l}) - D_i)^T$ - транспонированный вектор –столбец (вектор-строка); координаты векторов Y_i , $F(\vec{l})$, и Di являются спектральными составляющими в соответствующих базисных системах.

Практические алгоритмы обнаружения и оценивания используют логарифм отношения функции правдоподобия $LY_i(\vec{l})$, который для рассматриваемого случая с учетом (13) записывается как

$$LY_i(\vec{l}) = F^T(\vec{l})B^{-1}Y_i - \frac{1}{2}F^T(\vec{l})B^{-1}F(\vec{l}) - F^T(\vec{l})B^{-1}D_i \quad (14)$$

В скалярной форме выражения (2.53) и (2.54) примут следующий относительно простой вид:

$$W(Y/\vec{l}) = (\sqrt{(2\pi)^n\prod_{k=1}^{n}\sigma_{i,k}^2})^{-1}\exp\{-\sum_{k=1}^{n}\frac{(Y_i(k) - F(k,\vec{l}) - D_i(k))^2}{2\sigma_{i,k}^2}\},$$

$$W(Y/\vec{l}) = (\sqrt{(2\pi)^n\prod_{k=1}^{n}\sigma_{i,k}^2})^{-1}\exp\{-\sum_{k=1}^{n}\frac{(Y_i(k) - D_i(k))^2}{2\sigma_{i,k}^2}\}$$

$$L(Y_i) = \sum_{k=1}^{n}\frac{1}{\sigma_{i,k}^2}(F(k,\vec{l})y_i(k) - \frac{1}{2}F^2(k,\vec{l}) - F(k,\vec{l})D_i(k)) \quad (15)$$

3. Обнаружение полезного сигнала.

Решение задачи обнаружения полезного сигнала на фоне помех базируется на проверке статистических гипотез. В рассматриваемом случае процедура обнаружения сводится к проверке двух гипотез: о принадлежности выборки Y_i спектральных отчетов смеси полезного сигнала и помехи (гипотеза H_1) и о наличии в этой выборке только помехи (гипотеза H_0). В нашем случае вместе с помехой присутствует остаточный дрейф. Из критериев обнаружения в данном случае наиболее целесообразен критерий Неймана-Пирсона, который не требует наличия априорной информации о потерях при принятии ошибочных решений и знания априорных вероятностей распределения параметров[3]. Гипотеза H_1 в этом случае принимается при выполнении условия

$$L(Y_i) \geq \Lambda_0 \quad (16)$$

где $L(Y_i)$ определяется по (15). Здесь Λ_0 - порог, выбираемый так, чтобы можно было минимизировать вероятность пропуска сигнала $P_{пр}$ при заданной вероятности ложных обнаружении $P_{\Lambda O}$:

$$P_{\Lambda O}=P\{L(Y_i)> \Lambda_O/H_0\}=\beta \quad (17)$$

Подставив выражение $L(Y_i)$ из (15) в (16) и выделив в левой части неравенства член, зависящий от обрабатываемой реализации сигнала, получим:

$$M(\vec{l}) = \int_L U(\vec{l},t)Y(t)dt \geq \frac{1}{2}\int_L U(\vec{l},t)F(\vec{l},t)dt + \int_L U(\vec{l},t)D(t)dt + \Lambda_0 = \Lambda \qquad (18)$$

Значение порога Λ можно найти из этой формулы, воспользовавшись предположением о нормальности помеховой составляющей $M(\vec{l})$ по известным методикам [3]:

$$\Lambda = \aleph_\beta \sqrt{<\Xi^2>} + \int U(\vec{l},t)D(t)dt \qquad (19)$$

где $\aleph_\beta$- процентная точка нормального распределения, а $<\Xi^2>$ и второе слагаемое представляют собой соответственно дисперсию и корреляционную функции помеховой составляющей.

В случае некоррелированности выборочных значений (использование АБС) критерий обнаружения (18) упрощается:

$$\sum_{k=1}^{n} \sigma_k^{-2} F(k,\vec{l})Y_i(k) \geq \Lambda \qquad (20)$$

причем значение порога находится из (19):

$$\Lambda = \aleph_\beta \sqrt{\sum_{k=1}^{n} \sigma_k^2 F^2(k,\vec{l})} + \sum_{k=1}^{n} \sigma_k^{-2} F(k,\vec{l})D_i(k) \qquad (21)$$

Таким образом, процедура обнаружения состоит в вычислении интеграла (суммы) взвешенных значений спектральных составляющих сигнала (18) и (20) на i- той выборке и сравнении полученной суммы с пороговым значением по (19) или (21). Если неравенства (18), (20) не выполняются, выборка сдвигается на шаг Δl, и процедура повторяется. Зона спектра, где эти неравенства выполняются, принадлежит спектру одного компонента полезного сигнала (или нескольких, если они не разделены). Следовательно, процедура обнаружения позволяет найти участки спектров компонентов сигнала последовательно. В левой части (18) выделена функция, являющаяся достаточной статистикой и определяющая ту существенную операцию, которую надо произвести над принятой реализацией, чтобы извлечь всю информацию о неизвестном параметре, содержащуюся в данной реализации y(t). Аналогичность этой функции с (11) позволяет существенно упростить процедуру обнаружения, используя для этого обобщенные отсчеты $\underset{0}{Y}(k)$ как достаточные статистики временной реализации сигнала. Такой подход к обнаружению соответствует применению критерия Неймана-Пирсона в исходной (временной) области существования сигнала.

Действительно, функции правдоподобия временной выборки имеют вид, аналогичный (13), где составляющие векторов Yi, F, Di- временные отсчеты сигнала y(t) , моделей сигнала f(t) и дрейфа g(t): элементы матрицы B определяются корреляционной функцией помехи ξ(t). Соответствующие векторы обозначим строчными буквами. Тогда для логарифма отношения правдоподобия в дискретном случае получим:

$$l(y_i) = Y_{oi} - \frac{1}{2}\sum_{j=1}^{m} U_j(\vec{l}) f_j(\vec{l}) - \sum_{j=1}^{m} U_j(\vec{l}) g_{i,j} \qquad (22)$$

$$l(y_i) = Y_{oi} - \frac{1}{2\sigma_\xi^2}\sum_{j=1}^{m} f_j^2(\vec{l}) - \frac{1}{2\sigma_\xi^2}\sum_{j=1}^{m} f_j(l) g_{i,j}\,, \qquad (23)$$

где m – объем выборки, по которой определяется Y_{oi} : Δt принято равным $\Delta t=1$.

Выражения (22) и (23) справедливы при обработке по дискретным временным отсчетам для коррелированных и некоррелированных выборочных значений соответственно.

Критерий обнаружения примет вид:

$$Y_{oi} \geq \Pi_{в} \tag{24}$$

где $\Pi_в$ найдем аналогично (2.62), (2.63):

$$\Pi_{В} = \aleph_{\beta}\sigma_0 + \sum_{j=1}^{m} U_j(\vec{l})g_{i,j} \tag{25}$$

$$\Pi_{В} = \aleph_{\beta}\sigma_0 + \sum_{j=1}^{m} \frac{1}{\sigma_{\xi}^2} f_j(\vec{l})g_{i,j}\ , \tag{26}$$

где σ_0^2 - дисперсия обобщенного отсчета:

$$\sigma_0^2 = \sum_{j=1}^{m} U_j(\vec{l}) f_j(\vec{l}) \tag{27}$$

В частном случае компенсации дрейфа алгоритмически вторым членом в (25) и (26) можно пренебречь. Тогда, учитывая, что при изменении β в реальных пределах 10^{-3}-2.10^{-2}, величина $\aleph_{\beta}$ изменяется от 3,09 до 2,05 , в предположении некоррелированности временных отсчетов сигнала, получим правило обнаружения компонентов по i-той выборке в следующем простом виде:

$$Y_{oi} \geq (2 \div 3)\sigma_{\xi}^{-1}\sqrt{\sum_{j=1}^{m} f_j^2} \tag{28}$$

Практическая реализация алгоритма обнаружения (24) может быть осуществлена как в дискретном, так и в аналоговом виде. Величина порога при этом определяется выражениями (25) и (26).

Выводы

Таким образом, рассмотрены алгоритмы спектрального представления сигналов ХАК в адаптивных базисах, позволяющие получить некоррелированные и отфильтрованные отсчеты исходного сигнала. Последующая обработка таких отсчетов МНК - процедурами дают возможность вычисления оптимальных оценок искомых параметров. Предложены относительно простые схемы обнаружения полезного сигнала. Результаты оценки параметров выходных сигналов ХАК в спектральной области, надеемся, будут предметом обсуждения дальнейших публикаций.

ЛИТЕРАТУРА

1. Туленбаев М.С. Синтез адаптивных базисов спектрального представления сигналов химико-аналитических комплексов //Вестник ТарГУ имени М.Х. Дулати «Природопользование и проблемы антропосферы», 2009, №3, с.189-196
2. Пугачев В.С. Теория случайных функций. М.: Физматгиз, 1962. 884с.
3. Ван-Трис Г.А.Теория обнаружения, оценок и модуляции. М.: Сов.радио, т.1, 1972. 744с.
4. Левин Б.Р. Теоретические основы статистической радиотехники. М.: Сов.радио, т.1, 1974. 550с., т.2, 1975, 391с.