

український
МЕТРОЛОГІЧНИЙ
ЖУРНАЛ

№ 1 2011

науково-технічне видання

ПОЛУЧЕНИЕ ИНФОРМАЦИОННОЙ ИЗБЫТОЧНОСТИ В СИСТЕМАХ ИЗМЕРИТЕЛЬНОГО КОНТРОЛЯ И ДИАГНОСТИКИ ДИНАМИЧЕСКИХ ОБЪЕКТОВ

П.Ф.Щапов, доктор технических наук, профессор Национального технического университета "Харьковский политехнический институт"

О.Г.Аврунин, кандидат технических наук, доцент Харьковского национального университета радиозлектроники



П.Ф.Щапов



О.Г.Аврунин

Рассмотрена возможность применения метода кусочно-линейной регрессионной аппроксимации измерительных сигналов для обработки данных риноманометрической диагностики верхних дыхательных путей человека.

The possibility of using the method of piecewise-linear approximation of measurement signals for processing rhinomanometric diagnostics of human upper airways is considered.

Постановка проблемы. Совершенствование технических средств измерительного контроля и функциональной диагностики, расширение номенклатуры первичных преобразователей без учета динамических свойств входных сигналов не всегда приводят к повышению достоверности принятия решений. Это, в первую очередь, связано с наличием корреляционных связей между составляющими вектора входных сигналов информационно-измерительных средств контроля или диагностики. Увеличение размерности такого вектора новой информации практически не дает. К такому же результату приводит и техническое совершенствование измерительных преобразователей. Повышение их точности практически не снижает неопределенности результатов измерений при неустранимой неопределенности свойств сложных, диффузных объектов контроля, например, биологических и медицинских. В таких объектах подобная неопределенность обусловлена, чаще всего, динамическими свойствами.

Анализ литературы. Повышение количества ожидаемой измерительной информации прямо связано с уменьшением неопределенности результатов измерений [1, 2]. Наиболее эффективными в

этом случае являются структурно-алгоритмические методы [3, 4]. Их применение, в сочетании с информационно-измерительными технологиями преобразования первичной количественной информации в логические решения [5], позволяет уменьшить неопределенность последних.

Цель работы – показать возможности применения регрессионных моделей дисперсионного анализа нестационарных измерительных сигналов для получения дополнительной информации об их динамических свойствах, обусловленных изменениями состояний объекта контроля.

Математическая модель группового регрессионного преобразования. Рассмотрим последовательность результатов измерений физической величины X с указанием моментов времени ее измерения. Такая последовательность представляет упорядоченное по времени множество двумерных наблюдений:

$$X(t) = \{(x_1, t_1), (x_2, t_2), \dots, (x_N, t_N)\}. \quad (1)$$

Пусть Θ_0 и Θ_1 – обозначения функциональных состояний объекта контроля (в "норме" и, соответственно, не в "норме"). Информацию об изменении вида состояния объекта контроля несут информативные параметры измерительного сигнала $X(t)$. Выделение таких параметров связано с задачей синтеза математической модели сигнала $X(t)$, в которой изменение вида состояния объекта приводит к изменению, например, средних значений информативных параметров (коэффициентов модели). В общем виде модель измерительного сигнала содержит m условных параметров $\bar{a}_1(\Theta_r), \dots, \bar{a}_m(\Theta_r), r = 0, 1$, среднее значение которых меняется при изменении состояния объекта контроля:

$$a_l = \begin{cases} \bar{a}_l^{(0)}, & \text{если } r = 0; \\ \bar{a}_l^{(1)}, & \text{если } r = 1, \end{cases}$$

где $l = \overline{1, m}$.

В отличие от статических измеряемых величин, динамические сигналы открывают дополнительную возможность получения информационной избыточности за счет учета корреляционных связей этих сигналов со временем их наблюдения. Корреляция может проявляться в наличии трендов

(первого порядка и выше). Дополнительными информативными параметрами в этом случае будут коэффициенты, входящие в математические модели подобных трендов. Такие тренды являются регрессиями величины X на время t , а остаточная дисперсия такой регрессии может быть использована для оценки, получаемой при контроле информации (информативность тем выше, чем меньше остаточная дисперсия). Рассмотрим аппроксимацию сигнала $X(t)$ последовательностью K частных линейных регрессий со случайными коэффициентами:

$$x_{ji} = A_j + B_j \cdot t_{ji}, \quad j = \overline{1, k}; \quad i = \overline{1, n_j},$$

где k – число групп результатов измерений, для которых построены частные регрессии; n_j – число результатов измерений в j -й группе.

Общее число измерений равно

$$N = \sum_{j=1}^k n_j.$$

Пусть

$$\hat{X}_{ji} = A + B \cdot t_{ji}$$

является общей регрессией X на t , коэффициенты которой определены по всему множеству (1) двумерных результатов наблюдений. Коэффициенты же $\{A_j, B_j\}_1^k$ частных регрессий определяются по результатам соответствующих групповых измерений.

Известно [6], что сумму S квадратов отклонений результатов наблюдений от общего среднего \bar{X}

$$S = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_i - \bar{x})^2$$

можно разложить на пять слагаемых:

$$S = S_0 + S_{WG} + S_G + S_W + S_R, \quad (2)$$

где $S_0 = w_0 B_0^2$;

$$S_{WG} = \frac{w_c w_m}{w_0} (B_c - B_m)^2;$$

$$S_G = \sum_{j=1}^k n_j [\bar{x}_j - \bar{x} - B_m (\bar{t}_j - \bar{t})]^2;$$

$$S_W = \sum_{j=1}^k w_j (B_j - B_c)^2;$$

$$S_R = \sum_{j=1}^k \sum_{i=1}^{n_j} [x_{ji} - \bar{x}_j - B_j (t_{ji} - \bar{t}_j)]^2.$$

В свою очередь,

$$w_m = \sum_{j=1}^k n_j (\bar{t}_j - \bar{t})^2;$$

$$w_c = \sum_{j=1}^k w_j;$$

$$w_0 = w_m + w_c;$$

$$w_j = \sum_{i=1}^{n_j} (t_{ji} - \bar{t}_j)^2,$$

где \bar{x}, \bar{t} – общие средние по множествам $\{x_S\}_1^N$ и $\{t_S\}_1^N$; \bar{x}_j, \bar{t}_j – групповые средние по множествам $\{x_{ji}\}_{j=1}^k$ и $\{t_{ji}\}_{j=1}^k$.

Выбор информативных параметров. Учитывая, что сумма S_R позволяет оценить остаточную дисперсию \bar{S}_R данной регрессионной модели результата измерения

$$\bar{S}_R = \frac{S_R}{N - 2k},$$

выберем в качестве информативных параметров модели статистики:

$$\left\{ \begin{array}{l} F_0 = \frac{S_0}{\bar{S}_R}; \\ F_{WG} = \frac{S_{WG}}{\bar{S}_R}; \\ F_G = \frac{S_G}{\bar{S}_R \cdot (k - 2)}; \\ F_W = \frac{S_W}{\bar{S}_R \cdot (k - 1)}. \end{array} \right. \quad (3)$$

Данные статистики являются отношениями средних квадратов сумм S_0, S_{WG}, S_G и S_W к среднему квадрату остаточной суммы S_R , то есть представляют собой случайные величины с F -распределением (Фишера-Снедекора).

Дисперсионное разложение (2) позволяет рассчитывать F -статистики (3) по реализациям сигнала $X(t)$. Условиями такого разложения являются:

1) нормальность распределения случайного остатка

$$\varepsilon_{ji} = x_{ji} - \bar{x}_j - B_j (t_{ji} - \bar{t}_j), \quad \varepsilon_{ji} \approx \text{NORM}(0, \sigma_\varepsilon^2);$$

2) $M[\varepsilon_{ji}] = 0$;

3) $M[\varepsilon_{ji}^2] = \sigma_\varepsilon^2$;

4) некоррелированность остатков для всех

$$M[\varepsilon_{ji} \cdot \varepsilon_{jz}] = 0 \text{ для всех } i \neq z.$$

Информативность любой из F -статистик (3) определяется количеством информации, которое можно получить по этой статистике о виде состояния Θ_r объекта контроля. Достоинство F -статистик – независимость друг от друга в силу независимости [6] членов дисперсионного разложения (2). Это означает, что статистики (3) можно рассматривать как составляющие вектора

$$\bar{F} = (F_0, F_{WG}, F_G, F_W),$$

являющегося многомерным информативным параметром. Полная информация будет определяться суммой

$$I = I_0 + I_{WG} + I_G + I_W, \quad (4)$$

где слагаемые правой части могут быть рассчитаны независимо друг от друга [1].

Количество информации (4) характеризует параметры, определяемые как составляющие полной дисперсии измерительного сигнала $X(t)$ на интервале наблюдения $(0, t_N)$.

Эта дисперсия является линейной функцией остаточной дисперсии:

$$\bar{S}_R = \frac{S_R}{N - 2K}$$

При нормальном законе распределения измеряемой величины X ее линейное, относительно времени t , преобразование будет характеризоваться независимостью между средними значениями и дисперсией [7] (если не меняется ширина интервала наблюдения). Поэтому информация о смене состояния Θ объекта, полученная по F -статистикам дисперсионного разложения (2), может дополнять информацию, найденную по измерению среднего значения величины X .

Оценка количества дополнительной информации. Пусть $F_{\cdot}^{(0)}$ и $F_{\cdot}^{(1)}$ – статистики (3) дисперсионного разложения (2) (точка заменяет один из индексов “0”, “WG”, “G” или “W”). Данные статистики, как случайные величины, изменяют, в общем случае, нецентральное F -распределение с V_1 и V_2 степенями свободы с параметром нецентральности $\lambda_{\cdot}^{(n)}$ (где $r = \overline{0, 1}$):

$$F_{\cdot}^{(n)} \approx F_{V_1, V_2} \cdot \lambda_{\cdot}^{(n)}$$

Среднее и дисперсия статистики $F_{\cdot}^{(n)}$ соответственно равны [8]:

$$\chi_1^{(n)} = \frac{V_2}{V_2 - 2} \left(1 + \frac{\lambda_{\cdot}^{(n)}}{V_1} \right); \tag{5}$$

$$\chi_2^{(n)} = \frac{2V_2^2}{(V_2 - 2)(V_2 - 4)} \left(\frac{2\lambda_{\cdot}^{(n)}}{V_1} + \frac{(1 + \lambda_{\cdot}^{(n)})^2}{V_2 - 2} \right); \tag{6}$$

Количество информации, получаемой по статистике $F_{\cdot}^{(n)}$, определяется выражением [1]

$$I = \log_2 \sqrt{1 + \left(\frac{\sigma_F}{\sigma_{\Delta F}} \right)^2},$$

где σ_F^2 – дисперсия F -статистики до измерения (контроля); $\sigma_{\Delta F}^2$ – дисперсия F -статистики после измерения.

С учетом выражений (5) и (6) дисперсии σ_F^2 и $\sigma_{\Delta F}^2$ можно определить как

$$\sigma_F^2 \geq \frac{(\chi_1^{(0)} - \chi_1^{(1)})^2}{12};$$

$$\sigma_{\Delta F}^2 \geq 4(\max \chi_2^{(n)}).$$

В таблице представлены результаты дисперсионного анализа измеренных значений сигнала $X(t)$ для состояний Θ_0 и Θ_1 биологического объекта контроля ($N=9, K=3, n_j=n$ для всех $j = \overline{1, 3}$). В таблице также даны значения количества информации (в битах) при $\chi_1^{(0)} = F_{\cdot}^{(0)}$, $\chi_1^{(1)} = F_{\cdot}^{(1)}$, а дисперсия $\chi_2^{(n)}$ представлена как функция среднего $\chi_1^{(n)}$ [8]:

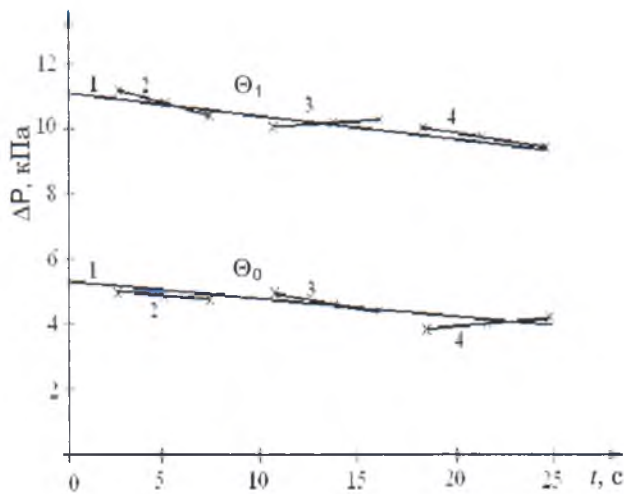
$$\chi_2^{(n)} = \frac{4V_2^2}{V_2 - 4} \cdot \chi_1^{(n)} + \frac{2V_1^2}{V_2 - 4} (\chi_1^{(n)})^2 - \frac{4V_2^2}{(V_2 - 2)(V_2 - 4)}$$

Если учесть, что оценка [1] количества информации по измерению среднего значения сигнала $X(t)$, то есть оценка по $\bar{X}^{(0)}$ и $\bar{X}^{(1)}$, дала величину $I_{\bar{X}} = 2,69$ (бит), то дополнительное приращение $I_{\bar{X}} = 1,09078$ (бит) составляет не менее 40 %, что указывает на эффективность предложенного метода дисперсионного преобразования измерительного сигнала. На рисунке представлены временные ряды измеренных значений сигнала $\Delta P(t)$ для $\Theta = \Theta_0$ и $\Theta = \Theta_1$ с последовательной регрессионной аппроксимацией рядов. Из рисунка наглядно видно изменение угловых коэффициентов частных регрессий при смене состояния объекта контроля.

Практическая реализация. Результаты работы апробированы на экспериментальных данных, полученных в результате проведения задней

№ п/п	Средний квадрат отклонений	Число степеней свободы	Средний квадрат	F-статистика	I (бит)
1	$S_0^{(1)} = 0,736502$ $S_0^{(0)} = 1,857845$	$V_0 = 1$	$\bar{S}_0^{(0)} = S_0^{(0)}$	$F_0^{(0)} = 92,896$ $F_0^{(1)} = 368,25$	0,92
2	$S_{WG}^{(1)} = 0,015786$ $S_{WG}^{(0)} = 0,011872$	$V_{WG} = 1$		$F_{WG}^{(0)} = 1,991$ $F_{WG}^{(1)} = 2,353$	0,00078
3	$S_G^{(0)} = 0,0088707$ $S_G^{(1)} = 0,11395$	$V_G = 1$		$F_G^{(0)} = 1,11888$ $F_G^{(1)} = 22,587$	0,17
4	$S_{\epsilon}^{(0)} = 0,039641$ $S_{\epsilon}^{(1)} = 0,025225$	$V_{\epsilon} = 1$	$\bar{S}_{\epsilon}^{(0)} = 0,007928$ $\bar{S}_{\epsilon}^{(1)} = 0,005045$	–	–

Примечания: для обеспечения условия $V_2 > 4$ суммы S_W и S_R были объединены. Соответственно возросло и число степеней свободы ($V_2 = V_{\epsilon} = 5$). Суммарное количество информации $I_F = 0,92 + 0,00078 + 0,17 = 1,09078$ (бит).



Временные ряды измеренных значений сигнала давления $\Delta P(t)$ с последовательной регрессионной аппроксимацией рядов в норме (состояние Θ_0) и при патологии (состояние Θ_1): 1 – прямые общих регрессий; 2, 3, 4 – прямые частных регрессий

активной риноманометрии на базе оториноларингологического отделения Харьковской областной клинической больницы с помощью разработанного риноманометра ТНДА-ПРХ [9]. При риноманометрии изучается зависимость расхода воздуха, проходящего через носовые ходы от перепада давлений между внутренним выходом из носовой полости (хоанами) и атмосферным. При этом по графикам пиковых значений перепада давлений (см. рисунок), а также расхода воздуха в носовой полости, очевидно, что изучение динамических свойств процесса дыхания может повысить качество диагностических процедур.

Выводы. Использование метода кусочно-линейной регрессионной аппроксимации измерительных сигналов позволило получить дополнительную информацию по изменениям случайных коэффициентов частных линейных регрессий.

Доказано, что дополнительную информацию, помимо коэффициентов частных регрессий, несут четыре члена дисперсионного разложения, что указывает на возможность получения дополнительной информации.

Доказано на практическом примере риноманометрической диагностики, что дополнительное увеличение ожидаемой измерительной информации может достигать 40 % от первоначальной. Последняя была получена по анализу изменений средних значений измеренных сигналов.

Предложенный метод дисперсионного анализа частных линейных регрессий обеспечивает получение дополнительной информации по составляющим дисперсионного разложения сигнала. Такая

процедура эквивалентна процедуре спектрального анализа при отсутствии информации об энергетическом спектре нестационарного по среднему измерительного сигнала, так как исследуемые последовательности результатов измерений являются временными рядами.

Перспективы работы. Данный метод можно рассматривать в рамках совершенствования информационно-измерительных технологий альтернативного контроля и технической диагностики при ограничениях по времени наблюдения (или числу измерений) при априорной неопределенности свойств объекта контроля и диагностики. Этот метод позволяет планировать многократные группированные измерения, полученные на базе нестационарных измерительных сигналов с априори неизвестными спектральными свойствами.

Список литературы

1. Орнатский П.П. Теоретические основы информационно-измерительной техники. – К.: Вища школа, 1983. – 455 с.
2. Щапов П.Ф. Нормализация метрологически неопределенных информационных сигналов для систем измерительного контроля динамических объектов // Механіка та машинобудування. – 2006. – № 1. – С. 280–286.
3. Володарський Є.Т., Кухарчук В.В., Поджаренко В.О., Сердюк Г.Б. Метрологічне забезпечення вимірювань і контролю. – Вінниця: Велес, 2001. – 219 с.
4. Кондрашов С.І. Методи підвищення точності систем тестових випробувань електричних вимірювальних перетворювачів у робочих режимах. – Харків: НТУ «ХПІ», 2004. – 224 с.
5. Малайчук В.П., Мозговой О.В., Петренко О.М. Інформаційно-вимірювальні технології неруйнівного контролю. – Дніпропетровськ: РВВ ДНУ, 2001. – 240 с.
6. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке: Методы планирования эксперимента / Пер. с англ. под ред. Э.К.Лецкого. – М.: Мир, 1981. – 520 с.
7. Жовинский А.Н., Жовинский В.Н. Инженерный экспресс-анализ случайных процессов. – М.: Энергия, 1979. – 112 с.
8. Шефе Г. Дисперсионный анализ / Пер. с англ. В.А.Севастьянова и В.П.Чистякова. – М.: Наука, 1980. – 512 с.
9. Аврунин О.Г. Методика расчета диаметра сопла вентури для устройства по определению перепадно-расходных характеристик носовых проходов // Промислова гідраліка і пневматика. – 2010. – № 2 (28). – С. 62–66.