АКАДЕМИЯ НАУК СССР
СИБИРСКОЕ ОТДЕЛЕНИЕ
АВТОМЕТРИЯ

№ 5 1970

УДК 621.317.080

М. В. САВЕНКОВ
(*Москва*)

ОЦЕНКА СПЕКТРАЛЬНОЙ ПЛОТНОСТИ СЛУЧАЙНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ, ЗАДАННОЙ НЕБОЛЬШИМ ЧИСЛОМ РЕАЛИЗАЦИЙ, ДЛЯ ПОСТРОЕНИЯ ПРОГНОЗА ИЗМЕНЕНИЯ ПАРАМЕТРОВ ТЕХНИЧЕСКИХ ОБЪЕКТОВ В ПРОЦЕССЕ ЭКСПЛУАТАЦИИ

Работоспособность широкого класса технических объектов целиком определяется значениями некоторой совокупности их параметров. Часто удается выбрать параметры, описывающие состояние объектов так, чтобы они были независимы. В этом случае при рассмотрении задачи оценки технического состояния объекта достаточно полагать, что оно характеризуется значением некоторого «определяющего параметра». При профилактике дорогостоящих технических объектов, отказы которых в процессе применения приводят к значительным потерям, целесообразно строить прогноз изменения «определяющего параметра» на время предполагаемого применения и в зависимости от полученных результатов корректировать программу профилактических работ. Применение такой системы эксплуатации технических объектов соответствует сложившейся практике, когда измерение значений определяющего параметра x производится, как правило, через промежутки времени Δt при проведении очередных профилактических работ. Предсказание определяющего параметра конкретного j-го технического объекта на время Δt вперед от момента последнего измерения $t=i\Delta t$ можно осуществить с помощью простейшей линейной формы:

$$x_{i+1}^{j} = \sum_{\nu=1}^{q} \gamma_{\nu} x_{i-\nu+1}^{j}, \tag{1}$$

где γ_ν — коэффициенты линейной формы, определяемые как решение системы уравнений, которые легко выписываются, если известна корреляционная функция R_p наблюдаемой случайной последовательности x_i; $x_{i-\nu+1}^{j}$ — результаты проведенных на j-м объекте измерений определяемого параметра.

Как показано в [1], прогноз по формуле (1) обеспечивает при правильном выборе γ_ν наименьший средний квадрат ошибки предсказания, т. е. если наблюдаемая последовательность x_i гауссовская (что на практике почти всегда выполняется), то прогноз этот неулучшаем. Однако необходимых для определения γ_ν априорных сведений о R_p

обычно нет, и приходится оценивать корреляционную функцию последовательностей x_i, на основании данных, получаемых при эксплуатационном контроле. Используемый при этом исходный материал $\{x_i^j\}$ содержит L случайных временных последовательностей:

$$x^j = \{x_1^j,\ x_2^j, \ldots,\ x_N^j\},$$

где $j=1, 2,..., L$ — номер объекта, у которого N раз контролировался параметр x. Объем $\{x_i^j\}$ чаще всего недостаточен для твердого суждения о виде R_p, что приводит к необходимости привлечения при построении прогноза x_{i+1}^j методов математической статистики.

Расчеты для предсказания значения определяющего параметра x на основании сведений только об $\{x_i^j\}$ становятся очень громоздкими. Однако при оценке состояния дорогих технических объектов это не имеет существенного значения, поскольку время, в течение которого надо принять решение об изменении программы профилактики, измеряется обычно часами, а затраты на получение значений x_i^j, их расшифровку и ввод в запоминающее устройство намного превышают затраты на самую сложную математическую обработку накапливаемых сведений.

Именно построению алгоритма статической обработки результатов эксплуатационного контроля параметров технических объектов для получения оценок спектральных характеристик наблюдаемых последовательностей x_i и построения с их использованием прогноза x_{i+1}^j посвящается данная работа. В дальнейшем будем полагать, что математические ожидания μ_i наблюдаемых последовательностей x_i равны нулю, а их среднеквадратические отклонения σ_i равны единице. Это предположение, значительно облегчающее выкладки, не существенно, поскольку μ_i и σ_i можно оценивать также по имеющимся исходным данным, применив алгоритм, подобный приведенному в [2]. Полагаем также, что спектральная плотность $f(\lambda)$ последовательностей x_i представима в виде произвольной дробно-рациональной функции

$$f(\lambda) = \frac{\left|\sum\limits_{\nu=0}^{h} \alpha_\nu e^{-i\lambda\nu}\right|^2}{\left|\sum\limits_{\nu=0}^{m} \beta_\nu e^{-i\lambda\nu}\right|^2}. \tag{2}$$

В виде (2) может быть представлена спектральная плотность любой случайной последовательности результатов измерения параметров технических объектов, так как при этом на x_i налагается лишь требование, чтобы значения x_i^j в данный момент не зависели от будущих значений. Оценки коэффициентов α_ν и β_ν, определяющие спектральную плотность $f(\lambda)$, могут быть получены по экспериментальным значениям корреляционной функции

$$R_p^j = \frac{\sum\limits_{i=1}^{N^j-p} x_i^j x_{i+p}^j}{\sum\limits_{i=1}^{N^j-p} x_i^j x_i^j},\quad R_p = \frac{1}{L}\sum_{j=1}^{L} R_p^j$$

как решения приводимых в дальнейшем систем уравнений. Но для решения этих систем необходимо знать значения h и m. И их можно оценить на основании имеющегося статистического материала, если проверять последовательно гипотезы о том, что собранным сведениям $\{x_i^j\}$ лучше всего соответствует представление (2) для $f(\lambda)$ при $m=0, 1, 2,...$ и $h=0, 1,..., m-1$. Для проверки этих гипотез необходимо построить специальный статистический критерий, что в дальнейшем будет сделано [см. (14)].

Если случайная последовательность x_i имеет спектральную плотность (2), то справедливо следующее соотношение [3]:

$$\sum_{\nu=0}^{m} \beta_\nu x_{i-\nu}^j = \eta_i^j = \sum_{\nu=0}^{h} \alpha_\nu \varepsilon_{i-\nu}^j, \tag{3}$$

где η_i^j — случайные величины, образующие последовательность типа «скользящего среднего» η_i; ε_i^j — независимые по i случайные величины, образующие чисто регулярную последовательность ε_i.

То же самое соотношение для дальнейшего удобно записать в матричной форме:

$$\begin{Vmatrix} \beta_0 & \beta_1 & \dots & \beta_m & 0 & 0 & \dots \\ 0 & \beta_0 & \dots & \beta_{m-1} & \beta_m & 0 & \dots \\ 0 & 0 & \beta_0 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{Vmatrix} \begin{Vmatrix} x_i^j \\ x_{i-1}^j \\ x_{i-2}^j \\ \dots \end{Vmatrix} = \begin{Vmatrix} \eta_i^j \\ \eta_{i-1}^j \\ \eta_{i-2}^j \\ \dots \end{Vmatrix} =$$

$$= \begin{Vmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_h & 0 & 0 & \dots \\ 0 & \alpha_0 & \dots & \alpha_{h-1} & \alpha_h & 0 & \dots \\ 0 & 0 & \alpha_0 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{Vmatrix} \begin{Vmatrix} \varepsilon_i^j \\ \varepsilon_{i-1}^j \\ \varepsilon_{i-2}^j \\ \dots \end{Vmatrix}$$

или $Bx^j = \eta^j = A\varepsilon^j$. Исходя из приведенного соотношения, нетрудно построить процедуру вычисления оценок для α_ν и β_ν на основании данных об $\{x_i^j\}$. Действительно, при $p>h$ значения корреляционной функции наблюдаемой последовательности x_i зависят лишь от β_ν, что позволяет, умножив левую часть (3) на x_{i+p}^j и просуммировав по всем i, получить:

$$\begin{cases} \sum\limits_{\nu=0}^{m} \beta_\nu R_{p-\nu} = 0; \\ p = h+1, h+2, \dots, h+m. \end{cases} \tag{4}$$

Полученные зависимости (4) образуют систему линейных уравнений для определения β_ν, если положить (без каких-либо ограничений общности) $\beta_0=1$. Получив решения $\beta_1, \beta_2, \dots, \beta_m$ этой системы, можно преобразовать результаты наблюдений $\{x_i^j\}$ так, чтобы исключить из них авторегрессионную составляющую

$$\eta_i^j = \sum_{\nu=0}^{m} \beta_\nu x_{i-\nu}^j, \tag{5}$$

и оценить корреляционную функцию последовательностей η_i

$$\rho_p = \sum_{\nu,\theta=0}^{m} \beta_\nu \beta_\theta R_{p+\theta-\nu}, \tag{6}$$

которая зависит лишь от α_ν. Действительно, для последовательности типа скользящего среднего η_i корреляционная функция ρ_p определяется формулами:

$$\left\{\begin{aligned} &\rho_p = \sum_{\nu=0}^{h-p} \alpha_\nu \alpha_{\nu+p}; \\ &p = 0, 1, \ldots, h, \end{aligned}\right. \tag{7}$$

которые совместно с (6) можно рассматривать как систему уравнений для определения α_ν. К сожалению, это нелинейная система и решение ее при $h>2$ возможно только численными методами на цифровых вычислительных машинах. Неединственность решения системы уравнений (7) в практических расчетах не вызывает больших затруднений, поскольку всегда, пользуясь правилом попарной перестановки симметричных ответов α_ν и $\alpha_{h-\nu}$, можно расположить $\alpha_0, \alpha_1, \ldots, \alpha_h$ в ряд, в порядке убывания их модуля. Выбор именно такого решения имеет наглядную физическую интерпретацию, так как влияние на данное i-е измерение x_i^j членов последовательности $\varepsilon_{i-\nu}$ должно ослабевать с ростом ν.

Вся процедура получения оценок для α_ν и β_ν сильно зависит от назначаемых заранее чисел m и h. Правильно ли назначены были m и h, можно проверить статистически, если убедиться, что из наблюдаемых последовательностей x_i, используя рассчитанные α_ν и β_ν, можно получить последовательности независимых по i случайных величин:

$$\varepsilon^j = A^{-1} B x^j = \Gamma x^j.$$

Специфический вид матрицы A позволяет выписать формулы для элементов γ_{kl} матрицы $\Gamma = A^{-1}B$:

$$\gamma_{kl} = \gamma_{l-k} = \sum_{\nu=\max(l-m,\,k)}^{l} \beta_{l-\nu} \Bigg[\sum_{u=1}^{\nu-k} C_n^u (-\alpha_0)^{-n} \prod_{\substack{v_\eta = 0 \\ \sum_{\eta=1}^{n} v_\eta = \nu - k}}^{n} \alpha_v \Bigg], \tag{8}$$

где u — число различных индексов среди n индексов v_η. Эта громоздкая формула полезна при немеханизированном счете. Последовательности ε^j имеют ковариационную матрицу вида

$$C = \Gamma R \Gamma', \tag{9}$$

где

$$R = \begin{Vmatrix} R_0 & R_1 & R_2 \ldots \\ R_1 & R_0 & R_1 \ldots \\ R_2 & R_1 & R_0 \ldots \\ . & . & . \; . \; . \end{Vmatrix}; \quad \Gamma = \begin{Vmatrix} \gamma_0 & \gamma_1 & \gamma_2 \ldots \\ 0 & \gamma_0 & \gamma_1 \ldots \\ 0 & 0 & \gamma_0 \ldots \\ . & . & . \; . \; . \end{Vmatrix}.$$

Если представление (2) абсолютно точно соответствует спектральной плотности $f(\lambda)$ наблюдаемых последовательностей x_i, то $R_p = 0$ при $p > m+h$ и у полученной, согласно (9), ковариационной матрицы C^0 недиагональные элементы равны нулю [в силу выбора α_ν и β_ν, согласно уравнениям (4), (7)], а диагональные элементы соответствуют дисперсии последовательностей ε^j

$$c_{kk} = D\varepsilon = \sum_{\nu, \theta = 0}^{\infty} \gamma_\nu \gamma_\theta R_{|\nu - \theta|}. \tag{10}$$

В практических расчетах, если даже для $f(\lambda)$ справедливо представление (2), недиагональные элементы матрицы C

$$c_{kl} = \sum_{\nu, \theta = 0}^{\infty} \gamma_\nu \gamma_\theta R_{|\nu + l - \theta - k|} \tag{11}$$

при $l-k>m+h$ отличны от нуля из-за случайных погрешностей определения R_p (а следовательно, α_ν и β_ν), вызванных недостаточностью объема исходных данных. Однако, если (2) при полученных на основе реальной статистики коэффициентах α_ν и β_ν хорошо сглаживает спектральную плотность наблюдаемой последовательности, недиагональные элементы матрицы C при $l-k>m+h$ окажутся значительно меньше диагональных. Это позволяет [4] аппроксимировать функцию правдоподобия Q, соответствующую гипотетической модели (2) для $f(\lambda)$, приближенным выражением

$$Q \approx N L n \frac{|C|}{|C^0|} \approx N \sum_{|l-k|>m+h} c_{kl}^2 (D\varepsilon)^{-2}. \tag{12}$$

Поскольку Q имеет распределение χ^2 с $N-m-h$ степенями свободы [4], постольку на основании (12) может быть предложен следующий статистический критерий для проверки гипотезы о том, что используемое при предсказании изменения определяющего параметра представление (2) его спектральной плотности $f(\lambda)$ соответствует имеющемуся статистическому материалу.

Если статистика

$$S(m, h) = \sum_{j=1}^{L} \sum_{p=m+h+1}^{N^j} (N^j - p + 1) \left[\frac{\sum_{\nu, \theta} \gamma_\nu \gamma_\theta R^j_{|p-\nu-\theta|}}{\sum_{\nu, \theta} \gamma_\nu \gamma_\theta R_{|\nu-\theta|}} \right]^2, \tag{13}$$

где γ определяется согласно (4), (7), (8), укладывается в квантиль распределения χ^2 с $\sum_{j=1}^{L} (N^j - m - h - 1)$ степенями свободы и заданным уровнем доверия, то степени полиномов m и h, аппроксимирующих спектральную плотность $f(\lambda)$ наблюдаемых последовательностей x_i, выбраны верно.

В формулах (10)—(13) суммирование осуществляется по бесконечному набору индексов, что невозможно реализовать при расчетах. Суммирование по бесконечному числу индексов порождается бесконечным числом отличных от нуля элементов в каждой строке матрицы Г. Конечное число $q=m+1$ отличных от нуля элементов в строках Г получается лишь в том случае, когда $h=0$, т. е. когда наблюдается последовательность авторегрессионного типа. И оптимальный линейный прогноз (1) для случайной последовательности, имеющий спектральную плотность вида (2) при $h \neq 0$, строится по бесконечной памяти о прошлом, а знание конечного числа предыдущих значений $x^j_{i-\nu}$ является достаточным при прогнозировании только авторегрессионных случайных процессов.

Физический анализ процессов изменения по времени параметров технических объектов позволяет утверждать, что чем дальше от данного момента i отстоит момент измерения $i-\nu$, тем меньшее влияние оказывает $x^j_{i-\nu}$ на x^j_i, т. е. $\gamma_\nu \to 0$ при $\nu \to \infty$. Это дает основание использовать во всех формулах суммирование по конечному числу индексов. Дать оценку скорости убывания γ_ν по ν не удается, поэтому при выборе максимального индекса $\nu_{\max} = q$, учитываемого в расчетах, приходится опираться лишь на качественные рассуждения. При $\nu \leqslant m+h+1$ получаемые, согласно (9), коэффициенты γ_ν имеют тот же порядок, что и α_ν/α_0. В свою очередь, α_ν убывает по ν тем быстрее, чем быстрее уменьшается по модулю корреляционная функция ρ_p последовательности η_i, значения которой связаны гораздо слабее, чем значения x_i. Поскольку при $\nu > m+h+1$ порядок γ_ν соответствует произведениям отношений α_ν/α_0, которые меньше единицы, постольку коэффициенты γ_ν, начиная с γ_{m+h+2}, делаются незначимы на фоне γ_1.

Итак, при реальном счете для выбора степеней m и h на основании статистики (13), построении прогноза по формуле (1) и оценке точности этого прогноза по формуле (10) следует принимать $\nu_{\max} = q = m+h+1$. Эту рекомендацию подтверждает опыт расчетов, результаты которых приведены в таблице. В качестве исходных данных $\{x^j_i\}$ в экспериментальных расчетах использовались сведения об изменении параметров

Число q зависимых последовательных значений x_i	Значение (в %) максимального добавочного значения γ_ν при $q=m+h+6$	Различие (в %) в определении γ_ν и S при $h=0$ и $h\neq 0$					
		γ_1	γ_2	γ_3	γ_4	γ_5	S
2	1,4	0,4	0,7	—	—	—	0,2
3	0,1	0,3	2,1	6,3	—	—	1,2
4	0,2	0,2	0,3	1,7	9,2	—	0,4
5	2	1,6	3,8	4	10,1	22	0,7

механических систем в процессе их эксплуатации. Были получены значения γ_ν при $q=m+h+2$, $m+h+3\ldots$. Оказалось, что начальные γ_ν при $\nu=1, 2,\ldots, m+h+1$ не меняются, а добавочные γ_ν при $q=m+h+6$ составляют от γ_1 всего 2%, что в статистических расчетах может признаваться несущественным.

Все расчеты, связанные с получением коэффициентов оптимального линейного прогноза, значительно облегчаются, если наблюдаемая последовательность x_i порождается процессом авторегрессионного типа. В этом случае нет необходимости решать нелинейную систему (7), $\gamma_\nu = \beta_\nu$, а статистика (13) превращается в известную статистику Кенвилля ([3], стр. 115). Поэтому интересен вопрос о том, представима ли спектральная плотность наблюдаемого случайного процесса $f(\lambda)$ в виде постоянной, деленной на полином, т. е. допустимо ли полагать $h=0$. Для ответа на этот вопрос можно сравнить значение $S(m+h, 0)$ со значениями $S(m+h-1, 1)$, $S(m+h-2, 2)$, которые нетрудно подсчитать, поскольку решения системы (7) при $h \leqslant 2$ имеют не очень громоздкое аналитическое выражение:

$$\alpha_0 = \frac{1}{4}(z_1 + z_2)\left(1 + \sqrt{1 - \frac{8\rho_2}{(z_1+z_2)^2}}\right); \quad \alpha_1 = \frac{1}{2}(z_1 - z_2);$$

$$\gamma_2 = \frac{1}{4}(z_1 + z_2)\left(1 - \sqrt{1 - \frac{8\rho_2}{(z_1 + z_2)^2}}\right), \tag{14}$$

где $z_1 = \sqrt{\rho_0 + 2\rho_1 + 2\rho_2}$; $z_2 = \sqrt{\rho_0 - 2\rho_1 + 2\rho_2}$.

Если уточнение $S(m, h)$ и коэффициентов γ_ν при $h > 0$ окажется незначительным, то следует принимать гипотезу о том, что наблюдается последовательность авторегрессионного типа. Проведенные нами статистические экспериментальные расчеты, результаты которых приведены в таблице, показывают, что изменение параметров технических объектов в процессе их эксплуатации хорошо аппроксимируется q-связанными марковскими последовательностями с $q=3\div4$. Различие в определении коэффициентов γ_1 и γ_2, а также статистики S при $h=0$ и $h\neq0$ составляет меньше чем 5%. А более значительное расхождение в определении γ_3, γ_4 и γ_5 не имеет большого значения при прогнозировании, поскольку модуль их не превышает обычно 3—8% от модуля γ_1

Предлагаемый алгоритм для прогнозирования изменения параметров j-го объекта позволяет учесть не только результаты контроля этих параметров на j-м объекте, но и статистику, описывающую течение их на L других аналогичных объектах, поскольку коэффициенты γ_ν в (1) определяются на основании данных об $\{x_i^j\}$. Число $m+h$ предыдущих измерений параметра x, которые необходимо учитывать при прогнозировании, также определяется на основе имеющихся данных об $\{x_i^j\}$, а не задается заранее. С ростом m и h обусловленность систем (4) и (7) значительно ухудшается, поэтому для прогнозирования необходимо использовать наименьшие из всех чисел m и h, при которых удовлетворяется статистика (13). Использование для предсказания результатов более ранних измерений параметра будет только увеличивать ошибку прогноза.

При прогнозировании изменения параметров технических объектов в условиях эксплуатирующей организации целесообразно использовать лишь простейшую формулу (1), а все γ_ν пересчитывать периодически в специальном вычислительном центре, обобщающем накопленный опыт эксплуатации всего парка объектов одного типа на основании сведений о результатах их периодического эксплуатационного контроля.

ЛИТЕРАТУРА

1. А. И. Яглом. Введение в теорию стационарных случайных функций.— Успехи математических наук, 1952, т. VII, вып. 5.
2. М. В. Савенков. Алгоритм обработки статистического материала в информационно-измерительной системе, используемой при контроле параметров технических устройств в процессе их эксплуатации.— Вероятностные методы в измерении и контроле, вып. 2. Новосибирск, «Наука», 1969.
3. Э. Хеннан. Анализ временных рядов. М., «Наука», 1964.
4. Д. Лоули, А. Максвели. Факторный анализ как статистический метод. М., «Мир», 1967.

Поступила в редакцию
21 мая 1970 г.