

6. Завьялов Ю. С., Квасов Б. И., Мирошниченко В. Л. Методы сплайн-функций.— М.: Наука, 1980.
7. Васильев Ф. П. Численные методы решения экстремальных задач.— М.: Наука, 1980.
8. Воскобойников Ю. Е. Критерий и алгоритмы выбора параметра при сглаживании сплайн-функциями.— В кн.: Алгоритмы обработки и средства автоматизации теплофизического эксперимента. Новосибирск: ИТФ СО АН СССР, 1978.
9. Golub H. G., Heath M. Generalized cross-validation as a method for choosing a good ridge parameter.— *Technometrics*, 1979, vol. 21, N 2, p. 215—223.
10. Воскобойников Ю. Е. Комплекс программ для сглаживания и дифференцирования экспериментальных данных при помощи В-сплайнов.— В кн.: Алгоритмические и аппаратные средства переработки информации. Новосибирск: ИТФ СО АН СССР, 1981.

Поступила в редакцию 6 января 1982 г.

УДК 621.391.1 : 519.246

Н. Г. ЧЕРНОГУЗ

(Баку)

ОТБРАКОВКА АНОМАЛЬНЫХ ОШИБОК С ПОМОЩЬЮ СТОХАСТИЧЕСКИХ МОДЕЛЕЙ

Существующие подходы к анализу применяемых в информационно-измерительных системах разностных алгоритмов отбраковки аномальных ошибок [1, 2] допускают определенное упрощение задачи и поэтому не дают достаточно полного представления об их эффективности. Целью настоящей работы является анализ алгоритма отбраковки аномальных ошибок при измерении центрированного нормального марковского случайного процесса $\{y_n\}$, $n = 0, 1, 2, \dots$, описываемого стохастическим разностным уравнением авторегрессии 1-го порядка [3]:

$$y_n = \rho y_{n-1} + v_n, \quad (1)$$

где $\rho = E[y_n y_{n-1}] / \sigma_y^2$, $\sigma_y^2 = D[y_n]$ ($E[\]$ — знак математического ожидания, $D[\]$ — дисперсии), $\{v_n\}$ — «белый» гауссов шум.

Рассмотрим аномальные ошибки как стационарный пуассоновский поток импульсов, имеющих величину λ и среднюю частоту ν , причем при появлении аномальной ошибки результат измерения x_n принимает значение $y'_n = y_n + \lambda$, а при ее отсутствии $x_n = y_n$.

Алгоритм отбраковки аномальных ошибок заключается в том, что для проверяемого в текущий момент наблюдения x_n вычисляется прогноз $\hat{y}_n = \rho z_{n-1}$, где z_{n-1} — результат проверки предыдущего наблюдения (причем в общем случае $\rho \neq 1$); и если абсолютная величина ошибки прогноза $u_n = x_n - \hat{y}_n$ превышает некоторый допустимый порог, то проверяемое наблюдение x_n заменяется своим прогнозом, т. е. $z_n = \rho z_{n-1}$, в противном случае $z_n = x_n$.

Пользуясь установившейся для задачи отбраковки терминологией [2], рассмотрим процесс принятия решения при проверке наблюдения x_n как проверку гипотезы $H_0: x_n = y'_n$ против альтернативы $H_1: x_n = y_n$. Ошибка в принятии решения при этом характеризуется двумя условными вероятностями: вероятностью пропуска аномальной ошибки $P(H_1|H_0)$ и вероятностью ложной отбраковки $P(H_0|H_1)$. Качество алгоритма отбраковки можно характеризовать, например, полной вероятностью ошибочных решений [4]

$$L = P(H_1|H_0) P_{H_0} + P(H_0|H_1) P_{H_1}, \quad (2)$$

где $P_{H_0} = \nu$ и $P_{H_1} = 1 - \nu$ — априорные вероятности недостоверных и достоверных наблюдений соответственно. Отметим, что $L_1 = P(H_1|H_0)P_H$ есть безусловная вероятность пропуска аномальной ошибки, а $L_2 = P(H_0|H_1)P_{H_1}$ — безусловная вероятность ложной отбраковки.

Однако непосредственное вычисление $P(H_1|H_0)$ и $P(H_0|H_1)$ в задаче отбраковки оказывается затруднительным. Проанализируем поведение дисперсии ошибки прогноза u_n в процессе работы алгоритма. Заметим, что если предыдущее наблюдение не отбраковано, то $z_{n-1} = x_{n-1}$ и $D[u_n] = \sigma_y^2 [1 - \rho^2 + (\varphi - \rho)^2]$. Если x_{n-1} отбраковано и для прогноза в качестве опорного наблюдения используется x_{n-2} , т. е. $z_{n-1} = \varphi x_{n-2}$, то тогда $D[u_n] = D[y_n - \varphi^2 y_{n-2}] = \sigma_y^2 [1 - \rho^4 + (\varphi^2 - \rho^2)^2]$. Далее, если $i-1$ предыдущих наблюдений подряд отбракованы, то $z_{n-1} = \varphi^{i-1} x_{n-i}$ и

$$D[u_n] = \sigma_y^2 [1 - \rho^{2i} + (\varphi^i - \rho^i)^2] = \sigma_y^2 t^2(i, \varphi, \rho), \quad (3)$$

где $t^2(i, \varphi, \rho) = 1 - \rho^{2i} + (\varphi^i - \rho^i)^2$.

Таким образом, дисперсия ошибки прогноза может в процессе отбраковки меняться, что необходимо учитывать при формировании порога отбраковки.

Предположим, что для гипотезы H_0 принят уровень значимости α . Для того чтобы фактический уровень значимости при обработке не менялся, порог должен возрастать или убывать в соответствии с (3). Сам алгоритм отбраковки при этом принимает вид

$$\begin{aligned} z_n &= \varphi z_{n-1}, \quad i = i + 1; \\ |x_n - \varphi z_{n-1}| &\geq h \sigma_y t(i, \varphi, \rho); \\ z_n &= x_n, \quad i = 1, \end{aligned} \quad (4)$$

где коэффициент h устанавливается в соответствии с α .

В свою очередь, смещение ошибки прогноза $E[u_n]$ зависит от достоверности опорного и проверяемого измерений, а также от величины упреждения при прогнозе i и определяется одним из следующих соотношений:

$$E[y_n - \varphi^i y_{n-i}] = 0, \quad (5)$$

$$E[y'_n - \varphi^i y'_{n-i}] = \lambda, \quad (6)$$

$$E[y_n - \varphi^i y'_{n-i}] = -\varphi^i \lambda, \quad (7)$$

$$E[y'_n - \varphi^i y'_{n-i}] = (1 - \varphi^i) \lambda. \quad (8)$$

Таким образом, в процессе отбраковки ошибка прогноза $u_n = x_n - \hat{y}_n$ распределена по нормальному закону, параметры которого изменяются на каждом шаге в соответствии с (3) и (5)–(8) случайным образом. Поэтому дальнейший анализ будет строиться с помощью условных вероятностей $P(H_c|H_d, \hat{y}_n)$ принятия гипотезы H_c , в то время как верна гипотеза H_d , $c, d = 0, 1$, а прогноз имеет величину \hat{y}_n (где $\hat{y}_n = \varphi^i y_{n-i}$ или $y_n = \varphi^i y'_{n-i}$). Указанные вероятности можно легко определить через соотношения (3), (5)–(8) и интеграл Лапласа $F(b) = (2\pi)^{-1/2} \int_{-\infty}^b \times \exp(-u^2/2) du$. Так, при принятии гипотезы H_1

$$\begin{aligned} P(H_1|H_d, \hat{y}_n) &= F\{(hD^{1/2}[u_n] - E[u_n])/D^{1/2}[u_n]\} - \\ &- F\{(-hD^{1/2}[u_n] - E[u_n])/D^{1/2}[u_n]\}, \quad d = 0, 1. \end{aligned}$$

В частности, для $P(H_1|H_0, \varphi^i y_{n-i})$, учитывая (3) и (6),

$$\begin{aligned} P(H_1|H_0, \varphi^i y_{n-i}) &= F\{h - \lambda/\sigma_y t(i, \varphi, \rho)\} - F\{-h - \lambda/\sigma_y t(i, \varphi, \rho)\} = \\ &= \Phi_h[\theta/t(i, \varphi, \rho)], \end{aligned} \quad (9)$$

где $\Phi_h(b) = F(h - b) - F(-h - b)$ и $\theta = \lambda/\sigma_y$.

Стационарное распределение вероятностей состояний однородной цепи Маркова $\{P_i^{(k)}\}$, $i = 1, 2, \dots, 2k$, может быть найдено с помощью системы уравнений Чепмена — Колмогорова [5]:

$$\sum_{i=1}^{2k} P_i^{(k)} P_{i,j} = P_j, \quad j = 1, 2, \dots, 2k. \quad (16)$$

Решение системы (16) обычными методами при больших k было бы связано с определенными вычислительными трудностями. Однако, учитывая вид матрицы (15), можно свести решение к простому рекуррентному алгоритму. Добавляя к (16) нормирующее соотношение

$$\sum_{i=1}^{2k} P_i^{(k)} = 1, \quad (17)$$

получим следующие уравнения относительно $P_1^{(k)}$ и $P_{k+1}^{(k)}$:

$$\begin{aligned} P_1^{(k)} (1 - P_{1,1} - P_{1,2} P_{2,1} - P_{1,2} P_{2,3} P_{3,1} - \dots - P_{1,2} P_{2,3} \dots P_{k-1,k} P_{k,1}) = \\ = P_{k+1}^{(k)} (P_{k+1,1} + P_{k+1,k+2} P_{k+2,1} + P_{k+1,k+2} P_{k+2,k+3} P_{k+3,1} + \dots \\ \dots + P_{k+1,k+2} P_{k+2,k+3} \dots P_{2k-1,2k} P_{2k,1}); \end{aligned} \quad (18)$$

$$\begin{aligned} P_1^{(k)} (P_{1,k+1} + P_{1,2} P_{2,k+1} + P_{1,2} P_{2,3} P_{3,k+1} + \dots + P_{1,2} P_{2,3} \dots P_{k-1,k} P_{k,k+1}) = \\ = P_{k+1}^{(k)} (1 - P_{k+1,k+1} - P_{k+1,k+2} P_{k+2,k+1} - P_{k+1,k+2} P_{k+2,k+3} P_{k+3,k+1} - \dots \\ \dots - P_{k+1,k+2} P_{k+2,k+3} \dots P_{2k-1,2k} P_{2k,k+1}); \end{aligned} \quad (19)$$

$$\begin{aligned} P_1^{(k)} (1 + P_{1,2} + P_{1,2} P_{2,3} + \dots + P_{1,2} P_{2,3} \dots P_{k-1,k}) + P_{k+1}^{(k)} (1 + P_{k+1,k+2} + \\ + P_{k+1,k+2} P_{k+2,k+3} + \dots + P_{k+1,k+2} P_{k+2,k+3} \dots P_{2k-1,2k}) = 1. \end{aligned} \quad (20)$$

Введем обозначения:

$$M_1^{(k)} = \begin{cases} 1, & k = 1; \\ \prod_{i=2}^k P_{i-1,i}, & k > 1; \end{cases} \quad M_2^{(k)} = \begin{cases} 1, & k = 1; \\ \prod_{i=2}^k P_{k+i-1,k+i}, & k > 1; \end{cases} \quad (21)$$

$$S_1^{(k)} = \sum_{i=1}^k P_{i,1} M_1^{(i)}; \quad S_2^{(k)} = \sum_{i=1}^k P_{k+i,1} M_2^{(i)}; \quad (22)$$

$$S_3^{(k)} = \sum_{i=1}^k P_{i,k+1} M_1^{(i)}; \quad S_4^{(k)} = \sum_{i=1}^k P_{k+i,k+1} M_2^{(i)}; \quad (23)$$

$$S_5^{(k)} = \sum_{i=1}^k M_1^{(i)}; \quad S_6^{(k)} = \sum_{i=1}^k M_2^{(i)}; \quad (24)$$

$$G_1^{(k)} = S_5^{(k)} S_2^{(k)} + (1 - S_1^{(k)}) S_6^{(k)}; \quad (25)$$

$$G_2^{(k)} = S_6^{(k)} S_3^{(k)} + (1 - S_4^{(k)}) S_5^{(k)}. \quad (26)$$

С учетом (21)—(26) из (18)—(20) получаем

$$P_1^{(k)} = S_2^{(k)} / G_1^{(k)}, \quad (27)$$

$$P_{k+1}^{(k)} = S_3^{(k)} / G_2^{(k)}. \quad (28)$$

Соотношения (27), (28) совместно с (25), (26) и выражениями

$$M_1^{(k)} = M_1^{(k-1)} P_{k-1,k}, \quad M_2^{(k)} = M_2^{(k-1)} P_{2k-1,2k}, \quad (29)$$

$$S_1^{(k)} = S_1^{(k-1)} + P_{k,1} M_1^{(k)}, \quad S_2^{(k)} = S_2^{(k-1)} + P_{2k,1} M_2^{(k)}, \quad (30)$$

$$S_3^{(k)} = S_3^{(k-1)} + P_{k,k+1} M_1^{(k)}, \quad S_4^{(k)} = S_4^{(k-1)} + P_{2k,k+1} M_2^{(k)}, \quad (31)$$

$$S_5^{(k)} = S_5^{(k-1)} + M_1^{(k)}, \quad S_6^{(k)} = S_6^{(k-1)} + M_2^{(k)} \quad (32)$$

образуют рекуррентный по k алгоритм для вычисления $P_1^{(k)}$ и $P_{k+1}^{(k)}$.

Очевидно, что алгоритм сходится. Так, для $S_1^{(k)}$, например, имеем

$$S_1 = \lim_{k \rightarrow \infty} S_1^{(k)} = \lim_{k \rightarrow \infty} \sum_{i=1}^k P_{i,1} M_1^{(i)} = P_{1,1} + \lim_{k \rightarrow \infty} \sum_{i=2}^k P_{i,1} M_1^{(i)} < P_{1,1} + \\ + \lim_{k \rightarrow \infty} \sum_{i=2}^k P_{i,1} q^i < P_{1,1} + \lim_{k \rightarrow \infty} \sum_{i=2}^k q^i < \infty,$$

где $q = \max \{P_{i-1,i}\} < 1$, $i = 2, 3, \dots$. Аналогичным образом можно показать сходимость и для величин $S_2^{(k)}$, $S_3^{(k)}$, $S_4^{(k)}$, $S_5^{(k)}$ и $S_6^{(k)}$. Подставляя соответствующие пределы в (25), (26), получим предельные значения $G_1^{(k)}$ и $G_2^{(k)}$, а затем, переходя к (27), (28), пределы искомых вероятностей $P_1^{(k)}$ и $P_{k+1}^{(k)}$. Пусть

$$P_i = \lim_{k \rightarrow \infty} P_i^{(k)}, \quad P_i^* = \lim_{k \rightarrow \infty} P_{k+i}^{(k)}, \quad i = 1, 2, 3, \dots, k. \quad (33)$$

Вероятности $\{P_i, P_i^*\}$, $i = 2, 3, \dots$, можно найти по формулам

$$P_i = P_1 M_1^{(i)}, \quad P_i^* = P_1^* M_2^{(i)}. \quad (34)$$

Теперь вычислим через полученные величины значение вероятности L , задаваемое соотношением (2). Полная вероятность ложной отбраковки $L_2^{(k)}$ с учетом $2k$ состояний имеет вид

$$L_2^{(k)} = \sum_{i=1}^k [P_i^{(k)} P(H_0 | H_1, \varphi^i y_{n-i}) (1 - \nu) + P_{k+i}^{(k)} P(H_0 | H_1, \varphi^i y'_{n-i}) (1 - \nu)]. \quad (35)$$

Учитывая, что $P(H_0 | H_1, \varphi^i y_{n-i}) (1 - \nu) = [1 - P(H_1 | H_1, \varphi^i y_{n-i})] (1 - \nu) = 1 - \nu - P(H_1 | H_1, \varphi^i y_{n-i}) (1 - \nu) = 1 - \nu - P_{i,1}$ и аналогично что $P(H_0 | H_1, \varphi^i y'_{n-i}) (1 - \nu) = 1 - \nu - P_{k+i,1}$, а также принимая во внимание (17) и (16), получим

$$L_2^{(k)} = \sum_{i=1}^k [P_i^{(k)} (1 - \nu - P_{i,1}) + P_{k+i}^{(k)} (1 - \nu - P_{k+i,1})] = \\ = (1 - \nu) \sum_{i=1}^k (P_i^{(k)} + P_{k+i}^{(k)}) - \sum_{i=1}^k (P_i^{(k)} P_{i,1} + P_{k+i}^{(k)} P_{k+i,1}) = 1 - \nu - P_1^{(k)}. \quad (36)$$

Таким образом, полная вероятность ложной отбраковки L_2 — это просто разность между априорной вероятностью достоверных измерений P_{H_1} и безусловной вероятностью их пропуска, соответствующей P_1 . Аналогичным образом можно найти и безусловную вероятность пропуска аномальных ошибок L_1 , которая соответствует вероятности P_1^* :

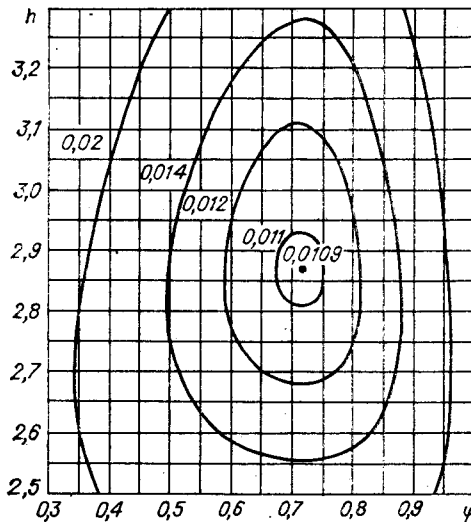
$$L_1 = P_1^*. \quad (37)$$

В итоге

$$L = L_1 + L_2 = P_1^* + 1 - \nu - P_1. \quad (38)$$

При расчетах оценки L быстро сходятся. Так, например, для случая, когда $\rho = 0,8$, $\varphi = 0,65$, $\theta = 3$, $h = 3$ и $\nu = 0,1$, за 7 итераций достигается точность порядка 10^{-4} (см. таблицу).

k	1	2	3	4	5	6	7
$L^{(k)}$	0,4574	0,0684	0,0454	0,0431	0,0429	0,0428	0,0428



Сходимость еще более улучшается при уменьшении φ , ρ , ν или θ ; необходимая точность при этом может быть достигнута всего за 2—3 итерации. С ростом указанных величин сходимость ухудшается, однако поскольку объем вычислений связан с числом итераций k линейно, то это всегда позволяет достичь «разумной» с вычислительной точки зрения точности.

Необходимо подчеркнуть, что рассмотренная схема вероятностного анализа применима (в отличие от упрощенных способов [1, 2]) и для таких условий, когда сама отбраковка, казалось, теряет смысл, например, при очень малых θ или при больших ν . Тем не менее вероятность принятия ложного решения L можно достаточно точно оценить и для таких случаев и, что самое главное, соответствующим выбором φ и h свести к минимуму. При этом достигаемый выигрыш может оказаться весьма существенным, а L может достигнуть приемлемой для практики величины.

В качестве конкретного примера приведем рассчитанные описанным выше способом значения L (см. рисунок) при $\rho = 0,8$, $\theta = 3$ и $\nu = 0,4$. Здесь минимальное значение $L \approx 0,0109$ достигается при $\varphi \approx 0,72$ и $h \approx 2,875$. Для сравнения отметим, что при $\varphi = 1$ (что соответствует часто используемому при коррелированных измерениях конечно-разностному алгоритму отбраковки) и $h = 3$ (правило «три сигма») величина L увеличивается почти в три раза.

ЛИТЕРАТУРА

1. Голубков В. С., Коротаев В. П. Отбраковка недостоверных результатов телеизмерений.— Метрология, 1973, № 2.
2. Новоселов О. Н., Фомин А. Ф. Основы теории и расчета информационно-измерительных систем.— М.: Машиностроение, 1980.
3. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление.— М.: Мир, 1974.
4. Ван Трие Г. Теория обнаружения оценок и модуляции.— М.: Сов. радио, 1972, т. 1.
5. Розанов Ю. А. Лекции по теории вероятностей.— М.: Наука, 1968.

Поступила в редакцию 13 августа 1979 г.;
окончательный вариант — 22 мая 1980 г.

УДК 621.398 : 621.391.274

А. Л. АЛИМОВ, А. Е. ЩАДИЛОВ

(Ленинград)

ОПТИМАЛЬНОЕ АДАПТИВНОЕ СЖАТИЕ ЦИФРОВЫХ СООБЩЕНИЙ ПО АЛГОРИТМУ КУСОЧНО-ЛИНЕЙНОЙ АППРОКСИМАЦИИ

При цифровой телеметрии статистически плохо определенного процесса обычно наблюдается информационная избыточность потока дискретных отсчетов контролируемого параметра. Для устранения избыточности, сжатия сообщений применяются специальные алгоритмические и