РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
АВТОМЕТРИЯ
2003, том 39, № 4

УДК 519.27

Ч. М. Гаджиев
(Стамбул, Турция)

МЕТОД ОТБРАКОВКИ АНОМАЛЬНЫХ ИЗМЕРЕНИЙ ДЛЯ МНОГОМЕРНЫХ ДИНАМИЧЕСКИХ СИСТЕМ

Разработан оперативный метод отбраковки аномальных измерений для многомерных динамических систем, робастный к систематическим погрешностям в измерениях. Предложенный метод не требует никакой информации о величине и знаке систематических погрешностей, а также о статистических характеристиках аномальных измерений и может быть использован в различных технических отраслях при обработке измерительной информации.

Практика статистической обработки измерительной информации показывает, что среди множества результатов измерений имеются отдельные значения (аномальные результаты измерения), которые сильно отличаются от всех остальных. Появление таких измерений связано с резким нарушением условий работы информационно-измерительной аппаратуры. В большинстве случаев аномальные измерения возникают в процессе измерительных преобразований. Однако они могут появляться также в результате сбоев ЭВМ при предварительной обработке измерений и вследствие сбоев при передаче данных по линиям связи.

Как показывают статистические исследования [1], самые различные научные, промышленные и другие данные содержат, как правило, 5–10 % аномальных измерений, а в случае радиотехнических измерений некоторые сеансы могут содержать более 10 % аномальных измерений [2], существенно снижающих эффективность применения многих классических статистических процедур. Поэтому аномальные измерения, нарушающие статистический характер информации об исследуемом процессе, необходимо своевременно обнаружить и исключить из последующей обработки.

Краткий обзор существующих методов отбраковки аномальных измерений приводится в работе [3]. На основе проведенного анализа отмечено, что в классических алгоритмах наличие систематических погрешностей измерений и их влияние на процедуру принятия решения не учитываются. Подход к конструированию алгоритмов отбраковки аномальных измерений, робастный к систематическим погрешностям в измерениях, предложен в [3] для одномерного случая. В данной работе указанный подход обобщается на многомерный случай.

Постановка задачи. Цель работы – разработка метода отбраковки аномальных результатов в многомерных моделях динамических измерений в условиях нормальности распределения случайных ошибок измерений.

Пусть расчетные значения измерений определяются на основе математической модели многомерной динамической системы, записанной в конечных разностях:

$$x_i = A(\theta)x_{i-1} + B(\theta)y_{i-1}, \tag{1}$$

где x_i – n-мерный вектор состояния системы; y_{i-1} – входное воздействие; θ – k-мерный вектор параметров системы (коэффициентов дифференциального уравнения); $A(\theta)$ и $B(\theta)$ – матрицы размера $n \times n$ и $n \times 1$ соответственно.

При этом истинные значения параметров математической модели динамической системы связаны с их оценками, полученными в результате процесса идентификации, в виде уравнения

$$\theta = \theta_e + \Delta\theta,$$

где θ_e – вектор оценок параметров динамической системы; $\Delta\theta$ – случайная составляющая с нулевым математическим ожиданием и конечной корреляционной матрицей ошибок оценок параметров.

Уравнение нормальных измерений записывается в виде

$$z_i = Hx_i + \xi_i + \lambda_i, \tag{2}$$

где z_i – s-мерный вектор измерений; H – матрица размера $s \times n$ преобразования вектора состояния в s-мерное пространство измерений; λ_i – s-мерный вектор систематической погрешности; ξ_i – случайный s-мерный вектор гауссовских шумов измерений с нулевым математическим ожиданием и корреляционной матрицей $E[\xi_i \xi_j^{\mathrm{T}}] = R_i \delta_{ij}$, E – оператор статистического усреднения, δ_{ij} – символ Кронекера:

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

В измерениях z_i содержатся редкие единичные сбои, моменты появления которых случайны, статистические характеристики неизвестны. Систематические погрешности λ_i также неизвестны. Принимаем, что $E[\lambda_i] = E[\lambda_{i-1}]$, так как систематическая погрешность при повторных измерениях остается постоянной или изменяется медленно.

При вышеизложенных условиях с целью повышения эффективности обработки измерительной информации требуется разработать оперативный метод отбраковки аномальных измерений, робастный к систематическим погрешностям в измерениях.

Алгоритм решения. Учитывая малость значений приращений оценок параметров и начальных значений вектора состояния модели системы, решение системы (1) приближенно запишем в следующем виде:

$$x_i \approx x_{p_i} + \frac{\partial x_i}{\partial (\Delta\theta)^{\mathrm{T}}} \Delta\theta,$$

где x_{p_i} – решение системы (1), полученное на основе оценок параметров.

Поскольку исследуемая динамическая система имеет линейную модель, то можно предположить, что достаточное количество информации о статистической природе расчетных значений выходных координат модели системы будет содержаться в описании их первыми двумя моментами:

$$E[x_i] = \bar{x}_{p_i}; \quad E[(x_i - \bar{x}_{p_i})(x_i - \bar{x}_{p_i})^{\mathrm{T}}] = D_{x_i}, \tag{3}$$

где принимаем, что

$$\bar{x}_{p_i} = \hat{A}(\theta)\bar{x}_{p_{i-1}} + \hat{B}(\theta)\hat{y}_{i-1}. \tag{4}$$

Выражения для определения D_{x_i} можно найти в [4].

В силу предельной теоремы теории вероятности, учитывая множество компонент, влияющих на значение ковариационной матрицы D_{x_i}, распределение истинного значения вектора состояния x_i в i-й момент времени принимаем нормальным $N(\bar{x}_{p_i}, D_{x_i})$. Таким образом, оценка $\bar{x}_{p_i}$, получаемая в соответствии с выражением (4) путем прямой экстраполяции начальных условий (в качестве начальных условий используются оценки параметров и состояния системы, полученные на момент прекращения процесса идентификации), представляет собой расчетные значения выходных координат модели системы.

Как известно, оценка не совпадает с истинным значением, и вследствие этого существует ошибка $\varepsilon_i = x_i - \bar{x}_{p_i}$.

Обозначим

$$\Delta\bar{x}_{p_i} = \bar{x}_{p_i} - \bar{x}_{p_{i-1}}, \qquad \Delta z_i = z_i - z_{i-1} \tag{5}$$

и вычислим невязку

$$\mathrm{v}_i = \Delta z_i - H\Delta\bar{x}_{p_i}. \tag{6}$$

С целью обнаружения аномальных измерений предлагается использовать статистику вида

$$r_i^{*2} = \mathrm{v}_i^{\mathrm{T}} P_{\mathrm{v}_i}^{-1} \mathrm{v}_i, \tag{7}$$

где P_{v_i} – ковариационная матрица невязки v_i.

Отметим, что при формировании введенной статистики неизвестные систематические погрешности взаимно исключают друг друга, т. е. статистика (7) робастна к систематическим погрешностям.

Покажем, что при отсутствии аномальных измерений в канале измерения невязка v_i распределена по нормальному закону с нулевым математическим ожиданием. Для этого с учетом (5) невязку (6) представим в раскрытой форме

$$\mathrm{v}_i = z_i - z_{i-1} + H\bar{x}_{p_{i-1}} - H\bar{x}_{p_i}. \tag{8}$$

Как известно [5], случайные величины, полученные в результате любых линейных преобразований нормально распределенных случайных величин, распределены нормально. На основе этого можно констатировать, что невязка v_i распределена по нормальному закону.

С учетом (2) и (8) определим математическое ожидание невязки v_i:

$$E[\mathrm{v}_i]=E[Hx_i+\xi_i+\lambda_i-Hx_{i-1}-\xi_{i-1}-\lambda_{i-1}+H\bar{x}_{p_{i-1}}-H\bar{x}_{p_i}]=$$

$$=H\{E[x_i-\bar{x}_{p_i}]-E[x_{i-1}-\bar{x}_{p_{i-1}}]\}+\{E[\xi_i]-E[\xi_{i-1}]\}+\{E[\lambda_i]-E[\lambda_{i-1}]\}.$$

Первый член в правой части этого уравнения равен нулю согласно выражению (3). Второй член равен нулю согласно исходным данным о статистических характеристиках шумов измерений. Третий член равен нулю, так как $E[\lambda_i]=E[\lambda_{i-1}]$. Следовательно, имеем $E[\mathrm{v}_i]=0$.

Покажем, что при отсутствии аномальных измерений ковариационная матрица невязки (6) определяется выражением

$$E[\mathrm{v}_i\mathrm{v}_i^{\mathrm{T}}]=P_{\mathrm{v}_i}=2R_i+H(D_{x_i}+D_{x_{i-1}}-D_{x_{i,i-1}}-D^{\mathrm{T}}_{x_{i,i-1}})H^{\mathrm{T}}, \tag{9}$$

где $D_{x_{i,i-1}}$ – перекрестная ковариационная матрица между ошибками двух последовательных расчетных значений выходных координат модели.

Учитывая независимости измерений и расчетных значений выходных координат модели, ковариационную матрицу невязки (6) в общем виде можем представить как

$$E[\mathrm{v}_i\mathrm{v}_i^{\mathrm{T}}]=E\{[\Delta z_i-H\Delta\bar{x}_{p_i}][\Delta z_i-H\Delta\bar{x}_{p_i}]^{\mathrm{T}}\}=$$

$$=E\{[\xi_i-\xi_{i-1}][\xi_i-\xi_{i-1}]^{\mathrm{T}}\}+HE\{[\varepsilon_i-\varepsilon_{i-1}][\varepsilon_i-\varepsilon_{i-1}]^{\mathrm{T}}\}H^{\mathrm{T}}, \tag{10}$$

где $\varepsilon_{i-1}=x_{i-1}-\bar{x}_{p_{i-1}}$.

В силу независимости измерений

$$E\{[\xi_i-\xi_{i-1}][\xi_i-\xi_{i-1}]^{\mathrm{T}}\}=2R_i. \tag{11}$$

Вычисление второго слагаемого выражения (10) представляет сложную задачу, так как ошибки в двух последовательных расчетных значениях выходных координат модели являются коррелированными из-за использования при их вычислении одной и той же модели, а также одних и тех же начальных условий. В результате между указанными ошибками появляются перекрестные ковариации.

Результаты, полученные в работе [6] и применяемые для решения данной задачи, запишем в виде

$$E\{[\varepsilon_i-\varepsilon_{i-1}][\varepsilon_i-\varepsilon_{i-1}]^{\mathrm{T}}\}=E[\varepsilon_i\varepsilon_i^{\mathrm{T}}-\varepsilon_i\varepsilon_{i-1}^{\mathrm{T}}-\varepsilon_{i-1}\varepsilon_i^{\mathrm{T}}+\varepsilon_{i-1}\varepsilon_{i-1}^{\mathrm{T}}]=$$

$$=D_{x_i}+D_{x_{i-1}}-D_{x_{i,i-1}}-D^{\mathrm{T}}_{x_{i,i-1}}, \tag{12}$$

где $D_{x_{i,i-1}} = E[\varepsilon_i \varepsilon_{i-1}^{\text{т}}] = D_{x_{i-1,i}}^{\text{т}}$. Подставляя (11) и (12) в (10), получим доказываемое выражение (9).

На основе изложенного можем констатировать, что предложенная для обнаружения аномальных измерений статистика (7) имеет χ^2-распределение с s степенями свободы (так как v_i имеет гауссовский закон распределения с нулевым средним и ковариационной матрицей P_{v_i}).

Задавая уровень значимости α, из условия

$$P\{\chi^2 \le \chi_\beta^2\} = \beta,$$

где $\beta = 1 - \alpha$ – доверительная вероятность, определяем порог (квантиль) χ^2-распределения – χ_β^2. Тогда правило принятия решения об аномальности измерения на данном шаге i на основе предложенной статистики (7) запишется в следующем виде:

$$\begin{aligned} &\gamma_1:\ r_i^{*2} \le \chi_\beta^2; \\ &\gamma_2:\ r_i^{*2} > \chi_\beta^2. \end{aligned} \tag{13}$$

Приведенный алгоритм обнаружения аномальных измерений робастен к систематическим погрешностям в измерениях, он достаточно прост и легко реализуется на ЦВМ.

Следует отметить, что при использовании данного подхода выявленное аномальное измерение из дальнейшей обработки исключается. Так как в контролируемую статистику (7) входит разность двух последовательных измерений, возникает трудность при формировании следующей разности Δz_{i+1} после отбраковки i-го измерения. В подобном случае в качестве измерения z_i можно использовать соответствующие расчетные значения выходных координат модели $\bar{x}_{p_i}$. Однако это может привести к заниженным значениям статистики r^{*2} на следующем шаге контроля.

Предлагаем другой подход к решению этой задачи.

Так как в случае отсутствия аномальных измерений в измерительном канале невязка v_i имеет нормальное распределение с нулевым средним и ковариационной матрицей (9), уравнение

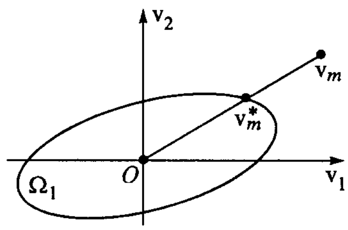
$$\text{v}_i^{\text{т}} P_{\text{v}_i}^{-1} \text{v}_i = \rho^2$$

соответствует эллипсоиду рассеяния (P_{v_i} – положительно-определенная матрица), где ρ – размеры полуосей эллипсоида в средних квадратических отклонениях.

Нетрудно видеть, что уравнение

$$\text{v}_i^{\text{т}} P_{\text{v}_i}^{-1} \text{v}_i = \chi_\beta^2 \tag{14}$$

соответствует допустимому эллипсоиду рассеяния.



В процессе контроля обнаруженное на основе решающего правила (13) аномальное измерение z_i из дальнейшей обработки исключается. Вместо z_i в алгоритме отбраковки аномальных измерений при формировании последующей разности $\Delta z_{i+1} = z_{i+1} - z_i$ предлагается использовать его допустимое граничное значение, определяемое следующим образом:

1) определяется допустимое граничное значение невязки v_i^* как проекция точки v_i на поверхность эллипсоида (14);

2) из выражения (8) при найденном значении v_i^* определяется допустимое граничное значение измерения z_i^* как

$$z_i^* = v_i^* + z_{i-1} - H\bar{x}_{p_{i-1}} + H\bar{x}_{p_i}. \tag{15}$$

При этом проекция точки v_i на поверхность эллипсоида (14) (значение v_i^*) определяется по формуле

$$v_i^* = \frac{v_i \sqrt{\chi_\beta^2}}{[v_i^{\mathrm{T}} P_{v_i}^{-1} v_i]^{1/2}}. \tag{16}$$

Действительно, уравнение проектирующей прямой имеет вид (см. рисунок)

$$v_i = t v_m, \tag{17}$$

так как она проходит через начало координат и точки v_m.

Запишем уравнение допустимого эллипсоида Ω_i с учетом равенства (17):

$$t^2 v_m^{\mathrm{T}} P_{v_m}^{-1} v_m = \chi_\beta^2.$$

Отсюда находим величину

$$t_{1,2} = \pm \frac{\sqrt{\chi_\beta^2}}{[v_m^{\mathrm{T}} P_{v_m}^{-1} v_m]^{1/2}}. \tag{18}$$

Подставив (18) в уравнение (17) с учетом обеспечения

$$\min[(v_m - v_m^*)^{\mathrm{T}} (v_m - v_m^*)]^{1/2},$$

получим значение

$$v_m^* = \pm \frac{v_m \sqrt{\chi_\beta^2}}{[v_m^{\mathrm{T}} P_{v_m}^{-1} v_m]^{1/2}}.$$

После нахождения v_i^* с помощью (15) вычисляется допустимое граничное значение i-го измерения z_i^* и это значение в алгоритме отбраковки аномальных измерений заменяет аномальное измерение z_i.

Необходимо отметить, что предложенный подход к отбраковке аномальных измерений требует последовательной реализации и при формировании разности $\Delta z_i = z_i - z_{i-1}$ должна быть уверенность в том, что измерение z_{i-1} не является аномальным (z_{i-1} – либо нормальное распределение, либо вычисленное с помощью (15) граничное значение). В противном случае использование данного подхода может привести к неверным результатам.

В реальных условиях эксплуатации системы разработанный алгоритм обнаружения и отбраковки аномальных измерений сводится к следующей последовательности вычислений.

1. Определяется допустимое значение (квантиль) χ^2-распределения χ^2_β, соответствующее выбранной доверительной вероятности β.

2. По расчетным значениям входных воздействий на основе выражения (4) определяются расчетные значения выходных координат модели системы $\bar{x}_{p_{i-1}}$ и $\bar{x}_{p_i}$.

3. В момент времени i по поступившим последовательным измерениям z_{i-1} и z_i (должна быть гарантирована нормальность измерения z_{i-1}) и расчетным значениям выходных координат модели системы $\bar{x}_{p_{i-1}}$ и $\bar{x}_{p_i}$ с помощью (8) вычисляется значение невязки v_i.

4. На основе (9) и (7) определяются ковариационная матрица невязки P_{v_i} и значение квадратичной формы r_i^{*2} (контролируемой статистики) соответственно.

5. На основе решающего правила (13) принимается решение о нормальности или аномальности измерения z_i. При этом возможны два случая:

а) при справедливости гипотезы γ_1 измерение z_i считается нормальным и поступает в дальнейшую обработку и последовательность предложенных контрольных измерений повторяется, начиная с п. 2, для следующего момента времени $i+1$;

б) в случае справедливости гипотезы γ_2 измерение z_i отбраковывается как аномальное и с помощью (16) и (15) определяется допустимое граничное значение i-го измерения z_i^*.

6. Последовательность вычислений повторяется, начиная с п. 2, для следующего момента времени $i+1$. При этом в качестве предыдущего измерения используется значение z_i^*.

Основным достоинством разработанного алгоритма отбраковки аномальных измерений является то, что для его работы не требуется информация о величине и знаке систематических погрешностей, а также о статистических характеристиках аномальных измерений. Предложенный алгоритм пригоден для многомерных динамических систем при знании их математической модели и может быть использован с целью обработки измерительной информации в системах автоматического управления летательными аппаратами, радиолокационных системах, инерциальных и спутниковых системах навигации и других.

СПИСОК ЛИТЕРАТУРЫ

1. **Hampel F. R.** Robust estimation: a condensed partial survey // Z. Wahrscheinlichkeits – Theorie and Verw. Geb. 1973. **27**. S. 87.
2. **Брандин В. Н., Васильев А. А., Куницкий А. А.** Экспериментальная баллистика космических аппаратов. М.: Машиностроение, 1984.
3. **Гаджиев Ч. М.** Подход к отбраковке аномальных измерений, робастный к систематическим погрешностям // Автометрия. 2002. **38**, № 4. С. 12.
4. **Гаджиев Ч. М.** Отбраковка аномальных измерений с учетом погрешностей математической модели измеряемого процесса // Метрология. 1998. № 2. С. 3.
5. **Пугачев В. С.** Теория вероятностей и математическая статистика. М.: Наука, 1979.
6. **Brumback B. D., Srinath M. D.** A chi-square test for fault-detection in Kalman filters // IEEE Trans. on Automatic Control. 1987. **AC-32**, N 6. P. 552.

Стамбульский технический университет,
E-mail: cingiz@itu.edu.tr

Поступила в редакцию
13 января 2003 г.