КУЙБЫШЕВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ им.С.П.КОРОЛЁВА
Труды, выпуск 77, 1975 г.

А.Ф.Бочкарев, И.В.Белоконов

АНАЛИЗ ВЛИЯНИЯ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВОЗМУЩЕНИЙ НА РАЗБРОС ПАРАМЕТРОВ ИССЛЕДУЕМОГО ПРОЦЕССА

При многократном статистическом анализе сложных систем, столь часто встречающихся в авиационной технике, важное значение имеет вопрос о разделении большой совокупности действующих на объект исследования случайных возмущений на существенные и несущественные, о выделении факторов, по которым можно пользоваться линейной моделью. Все это позволяет упростить и существенно снизить трудоемкость исследования.

Известно, что любую случайную функцию можно представить в виде канонического [I] или неканонического [2] разложения относительно случайных величин. Поэтому можно рассматривать в качестве входных возмущений только случайные величины, относительно которых проводятся все исследования, что существенно упрощает моделирование исследуемого процесса. В работе [3] при определении значимости возмущений принимались гипотезы о независимости случайных факторов и о равенстве нулю их нечетных центральных моментов.

Допущение о центрированности случайных величин не снижает общности результатов, в то время как равенство нулю внешних моментов накладывает существенные ограничения на вид закона распределения.

Покажем, что снятие последнего упрощения не усложняет анализ, и он также проводится по простым аналитическим соотношениям

Пусть исследуемый процесс можно записать в виде функции $y=\Phi(X)$ m - мерного вектора центрированных случайных величин X.

В результате проведения N экспериментов с исследуемым объек - том (в динамике полета, например, под экспериментом понимают од- кратное интегрирование уравнений математической модели в случае анализа рассеивания параметров движения) получена выборка векто- ра случайных величин $X^{(1)},\dots X^{(N)}$ и соответствующая ей выборка функции $y^{(1)},\dots y^{(N)}$.

Задача исследования значимости случайных величин $x_i\ (i=\overline{1,m})$ в формировании рассеивания функции y относится к задачам мно- гофакторного анализа, в основе которого лежит получение некото - рой модели функции $\tilde{y}$ из условия её наилучшего приближения к исходной функции y .

Для статистического исследования в первом приближении обыч- но ограничиваются представлением функции y в виде полиномов первой и второй степени вида

$$\tilde{y}^{(1)}=b_0^{(1)}+\sum_{i=1}^{m} b_i^{(1)} x_i\ , \tag{1}$$

$$\tilde{y}^{(2)}=b_0^{(2)}+\sum_{i=1}^{m} b_i^{(2)} x_i+\sum_{i\le j=1}^{m} b_{ij}^{(2)} x_i x_j \tag{2}$$

(индексы "1" и "2" указывают на степень уравнения регрессии).

Следуя методике, развитой в [3] , запишем соотношения для ма- тематических ожиданий и дисперсий рассматриваемых моделей:

$$M[\tilde{y}^{(1)}]=b_0^{(1)}\ , \tag{3}$$

$$\sigma^2[\tilde{y}^{(1)}]=\sum_{i=1}^{m} b_i^{(1)^2} M[x_i^2]\ , \tag{4}$$

$$M[\tilde{y}^{(2)}]=b_0^{(2)}+\sum_{i=1}^{m} b_{ii}^{(2)} M[x_i^2]\ , \tag{5}$$

$$\begin{aligned}\sigma^2[\tilde{y}^{(2)}]=\sum_{i=1}^{m}[b_i^{(2)^2} M[x_i^2]+b_{ii}^{(2)^2}(M[x_i^4]-M^2[x_i^2])+\\ +2b_i^{(2)}b_{ii}^{(2)}M[x_i^3]]+\sum_{i<j=1}^{m} b_{ij}^{(2)^2} M[x_i^2]M[x_j^2]\ .\end{aligned} \tag{6}$$

Отсюда видно, что вклад i -ого фактора в формирование дис- персии функции y определяется так:

$$\eta_i^{(1)}=b_i^{(1)^2} M[x_i^2]\ , \tag{7}$$

$$\eta_i^{(2)} = b_i^{(2)^2} M[x_i^2] + b_{ii}^{(2)^2} (M[x_i^4] - M^2[x_i^2]) +$$
$$+ 2 b_i^{(2)} b_{ii}^{(2)} M[x_i^3] + M[x_i^2] \sum_{\substack{j=1 \\ j \neq i}}^{m} \frac{b_{ij}^{(2)^2}}{2} M[x_j^2]. \tag{8}$$

Тогда коэффициент значимости, позволяющий разделить все случайные факторы на существенные и несущественные, можно найти следующим образом:

$$\delta\eta_i^{(1)} = \frac{\eta_i^{(1)}}{\sigma^2[y]}, \qquad \delta\eta_i^{(2)} = \frac{\eta_i^{(2)}}{\sigma^2[y]} \tag{9}$$

Несущественные факторы в исследованиях можно не учитывать.

Из числа существенных факторов выделяются такие, по которым можно пользоваться линейной моделью и которые требуют построения нелинейных зависимостей.

Для оценки этого вводится коэффициент нелинейности

$$\mu_i = \frac{\eta_i^{(2)} - \eta_i^{(1)}}{\eta_i^{(2)}} \tag{10}$$

Таким образом, значение $\eta_i^{(1)}$, $\eta_i^{(2)}$, μ_i позволяет определить степень влияния независимых случайных факторов на разброс параметров исследуемого процесса.

Из вышесказанного видно, что вся трудность анализа заключается в определении коэффициентов уравнений регрессии.

Зная выборку реализации X и Y, коэффициенты полиномов можно определить в рамках регрессионного анализа с помощью метода наименьших квадратов, из условия минимума выражения $L = M[\varepsilon^2]$, где $\varepsilon = y - \bar{y}$.

В [3] получены коэффициенты регрессии при условии равенства нулю нечетных высших моментов случайных величин. Снятие этого допущения не изменяет выражения для коэффициентов линейного уравнения регрессии

$$b_0^{(1)} = M[y] \tag{11}$$

$$b_i^{(1)} = \frac{M[y x_i]}{M[x_i^2]}. \tag{12}$$

Получим аналогичные соотношения для квадратного уравнения регрессии из условий

$$
\begin{aligned}
&\frac{1}{2}\,\frac{\partial L}{\partial b_0^{(2)}} = M\Big[\varepsilon\,\frac{\partial \varepsilon}{\partial b_0^{(2)}}\Big] \\
&\frac{1}{2}\,\frac{\partial L}{\partial b_i^{(2)}} = M\Big[\varepsilon\,\frac{\partial \varepsilon}{\partial b_i^{(2)}}\Big], \quad i = 1, m \\
&\frac{1}{2}\,\frac{\partial L}{\partial b_{ij}^{(2)}} = M\Big[\varepsilon\,\frac{\partial \varepsilon}{\partial b_{ij}^{(2)}}\Big], \quad i \le j = 1, m.
\end{aligned}
\tag{I3}
$$

Выражения (I3) образуют линейную систему уравнений относительно $b_0^{(2)}$, $b_i^{(2)}$, $b_{ii}^{(2)}$, которую можно представить в следующем виде:

$$C \cdot B = A\,, \tag{I4}$$

где

$$
C = [C_{ij}]_{(2m+1)\times(2m+1)} =
\begin{bmatrix}
1 & 0 & \cdots & 0 & \cdots & 0 & M[x_1^2] & \cdots & M[x_i^2] & \cdots & M[x_m^2] \\
 & M[x_1^2] & \cdots & 0 & \cdots & 0 & M[x_1^3] & \cdots & 0 & \cdots & 0 \\
 & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & M[x_i^2] & \cdots & 0 & 0 & \cdots & M[x_i^3] & \cdots & 0 \\
 & & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & & & & M[x_m^2] & 0 & \cdots & 0 & \cdots & M[x_m^3] \\
 & & & & & & M[x_1^4] & \cdots & M[x_1^2]M[x_i^2] & \cdots & M[x_1^2]M[x_m^2] \\
 & & & & & & & \cdots & \cdots & \cdots & \cdots \\
 & & & & & & & & M[x_i^4] & \cdots & M[x_i^2]M[x_m^2] \\
 & & & & & & & & & \cdots & \cdots \\
 & & & & & & & & & & M[x_m^4]
\end{bmatrix}
$$

$$
B = \begin{bmatrix} b_0^{(2)} \\ b_1^{(2)} \\ \vdots \\ b_m^{(2)} \\ b_{11}^{(2)} \\ b_{22}^{(2)} \\ \vdots \\ b_{mm}^{(2)} \end{bmatrix}, \qquad
A = \begin{bmatrix} M[y] \\ M[yx_1] \\ \vdots \\ M[yx_m] \\ M[yx_1^2] \\ M[yx_2^2] \\ \vdots \\ M[yx_m^2] \end{bmatrix}
$$

Решение (I4) можно в матричном виде записать так:

$$B = A \cdot C^{-1} \tag{I5}$$

Операция обращения матрицы C сложна, поэтому для определения искомых коэффициентов проще воспользоваться методом Гаусса [4].

В таком случае получаются простые аналитические соотношения

$$b_0^{(2)} = M[y] - \sum_{i=1}^{m} b_{ii}^{(2)} M[x_i^2], \tag{I6}$$

$$b_i^{(2)} = \frac{M[yx_i]}{M[x_i^2]} - b_{ii}^{(2)} \frac{M[x_i^3]}{M[x_i^2]}, \quad i = 1,m \tag{I7}$$

$$b_{ij}^{(2)} = \begin{cases} \dfrac{M[yx_i x_j]}{M[x_i^2]M[x_j^2]}, & i < j \\ \dfrac{M[yx_i^2] - M[y]\cdot M[x_i^2] - \dfrac{M[x_i^3]}{M[x_i^2]} M[yx_i]}{M[x_i^4] - M^2[x_i^2] - \dfrac{M^2[x_i^3]}{M[x_i^2]}}, & i = j \end{cases} \tag{I8}$$

Во всех приведенных здесь формулах, если нет никакой информации о законах распределения случайных величин, вместо математических ожиданий от выражений, образованных из y и X, ставятся их статистические оценки, полученные на основании имеющейся выборки реализаций.

Например, $M[x_i^4] \approx \frac{1}{N}\sum_{j=1}^{N} x_i^{(j)4}$, $M[yx_i] \approx \frac{1}{N}\sum_{j=1}^{N} y^{(j)} x_i^{(j)}$ и.т.д.

Ошибку вероятностной аппроксимации можно оценить по соотношениям:

$$L^{(1)} = M[y^2] + b_0^{(1)2} - 2b_0^{(1)} M[y] + \sum_{i=1}^{m} b_i^{(1)} (b_i^{(1)} M[x_i^2] - 2M[yx_i]) \tag{I9}$$

$$L^{(2)} = M[y^2] + b_0^{(2)2} - 2b_0^{(2)} M[y] + \sum_{i=1}^{m} (b_i^{(2)} M[x_i^2] + b_{ii}^{(2)2} M[x_i^4] - \tag{20}$$

$$- 2b_i^{(2)} M[yx_i] + 2b_0^{(2)} M[x_i^2] + 2b_i^{(2)} b_{ii}^{(2)} M[x_i^3]) + 2\sum_{i<j=1}^{m} b_{ii}^{(2)} b_{jj}^{(2)} M[x_i^2] M[x_j^2] -$$

$$- 2\sum_{i\le j=1}^{m} b_{ij}^{(2)} M[yx_i x_j].$$

Все выражения, выведенные здесь для квадратичного уравнения регрессии, совпадают с полученными в [3] для симметричного закона распределения, если положить $M[x_i^3] = 0$, $i = 1,m$.

Таким образом, для независимых случайных факторов вышеприведенные соотношения обеспечивают быстрое проведение статистического анализа, который позволяет в дальнейших исследованиях не учитывать те факторы, которые оказывают малое влияние на дисперсию выходной величины, а также дает возможность обоснованно подходить к подбору модели связи между входными и выходными величинами.

Литература

I. Пугачев В.С. Теория случайных функций. М., Физматгиз, 1962.

2. Чернецкий В.И. Анализ точности нелинейных систем управления. М., Машиностроение, 1968.

3. Школьный Е.П., Майборода Л.А. Атмосфера и управление движением летательных аппаратов. Л., Гидрометеоиздат, 1973.

4. Березин Е.С., Жидков Н.П. Методы вычислений. М., Физматгиз 1960.