5. Сафронов А. Н. Исследование условий эффективности фазовой адаптации в когерентной оптике.— Автометрия, 1981, № 2.
6. Репин В. Г., Тартаковский Г. П. Статистический синтез при априорной неопределенности и адаптация информационных систем.— М.: Сов. радио, 1977.
7. Цыпкин Я. З. Адаптация и обучение в автоматических системах.— М.: Наука, 1968.
8. Цыпкин Я. З. Основы теории обучающихся систем.— М.: Наука, 1970.

Поступила в редакцию 16 сентября 1980 г.

УДК 621.391.1 : 518.5

А. Н. ЛАЗАРЧИК

(Минск)

ОЦЕНКА ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ В ПРИСУТСТВИИ ПОМЕХ

Задача восстановления (оценки) плотности распределения вероятности одномерной случайной величины по выборке из независимых наблюдений этой случайной величины рассматривалась в ряде работ [1—4], при этом предполагалось, что элементы выборки не содержат случайных ошибок.

В условиях реального эксперимента, когда помеха достаточно велика и сравнима с исследуемой величиной, полученная в результате измерений выборка, а следовательно и восстановленная по ней плотность распределения, могут значительно отличаться от действительной. В связи с этим возникает задача оценки плотности распределения по выборке, содержащей шумовую составляющую. Этот вопрос частично рассматривался в работе [5].

Пусть имеется случайная величина X. По выборке $(y_1, y_2, \ldots, y_n)$ из независимых значений другой случайной величины Y, представляющей собой результат взаимодействия исследуемой величины X и независимой случайной помехи ξ, требуется оценить плотность распределения случайной величины X.

Вид взаимодействия считается известным и описывается функцией $y = g(x, \xi)$. Будем также считать заданной плотность распределения помехи $w(\xi)$. Обозначим через $p(x)$ плотность распределения величины X, а через $q(y)$ — плотность распределения величины Y. Для того чтобы найти оценку плотности $p(x)$, определим связь между функциями $p(x)$ и $q(y)$. Используя плотность распределения помехи и функцию, определяющую вид взаимодействия сигнала с помехой, можно записать условную плотность распределения y при фиксированном x:

$$p(y|x) = w\left(g^{-1}(y, x)\right) \frac{\partial g^{-1}(y, x)}{\partial y} = K(y, x).$$

Тогда связь между плотностями $p(x)$ и $q(y)$ примет вид

$$q(y) = \int p(x) K(y, x)\, dx. \quad (1)$$

Соотношение (1), если его рассматривать как уравнение относительно $p(x)$, представляет собой интегральное уравнение Фредгольма I-го рода. Как известно [6], задача решения таких уравнений является некорректной, т. е. небольшие ошибки в определении функции $q(y)$ приводят к большим ошибкам решения $p(x)$. Так как оценка распределения $q(y)$, определяемая из эксперимента, всегда содержит некоторую ошибку, то решение $p(x)$ можно найти лишь с помощью метода регуляризации уравнения (1). Различные методы регуляризации сводятся к привлечению определенной априорной информации об искомом решении уравнения (1).

Будем предполагать, что $p(x)$ принадлежит линейному функциональному подпространству Φ_N размерностью N. Выбор линейного подпространства наиболее приемлем, так как интегральный оператор K, определяемый соотношением (1), является линейным. Зададим в Φ_N базис, состоящий из набора линейно независимых функций $(\varphi_1(x), \varphi_2(x), \ldots, \varphi_N(x))$. Тогда плотность распределения $p(x)$ можно представить в виде

$$p(x) = \sum_{k=1}^{N} c_k \varphi_k(x). \tag{2}$$

Подставив выражение (2) в уравнение (1), получим

$$q(y) = \sum_{k=1}^{N} c_k f_k(y), \tag{3}$$

где

$$f_k(y) = \int K(y, x)\, \varphi_k(x)\, dx.$$

Таким образом, чтобы найти оценку плотности распределения $p(x)$, необходимо определить оценку матрицы коэффициентов c в сумме (3) и подставить найденное значение в (2), так как оператор K отображает пространство Φ_N в пространство F_N с базисом $\{f_k(y)\}$, при этом координаты соответствующих векторов не изменяются.

В качестве оценок коэффициентов c_k можно, в частности, взять статистики, рассмотренные в [1]:

$$c^* = \frac{1}{n} \sum_{i=1}^{n} f^{\mathrm{T}}(y_i)\, r^{-1}. \tag{4}$$

Здесь $f^{\mathrm{T}}(y) = (f_1(y), f_2(y), \ldots, f_N(y))$ — матрица-строка; т — символ транспонирования; r^{-1} — матрица, обратная матрице

$$r_{kj} = \int f_k(\tau) f_j(\tau)\, d\tau \quad (k, j = 1, 2, \ldots, N).$$

Соотношение (4) соответствует тому, что за оценку функции $q(y)$ принимается статистика вида

$$u_n(y) = \frac{1}{n} \sum_{i=1}^{n} \delta(y_i, y),$$

где $\delta(y, z)$ — дельта-функция в линейном функциональном подпространстве F_N, определенная в [2] и обладающая следующим свойством: $\int \delta(y, z)\, f(z)\, dz = f(y)$ для любой функции $f(y) \in F_N$.

Оценка (4) является несмещенной и состоятельной для матрицы коэффициентов c. Безусловно, желательно было бы иметь эффективную оценку матрицы параметров, однако при сделанных выше допущениях относительно распределения $p(x)$ ее просто не существует. Можно попытаться найти оценку максимального правдоподобия, которая в данной ситуации будет асимптотически эффективной, но эта задача слишком сложна, поэтому для практических целей, по-видимому, более целесообразно использовать оценку (4). Итак, окончательный вид оценки плотности распределения $p(x)$ по выборке $\{y_n\}$ случайной величины Y таков (в матричной записи):

$$p_n(x) = \frac{1}{n} \sum_{i=1}^{n} f^{\mathrm{T}}(y_i)\, r^{-1} \varphi(x). \tag{5}$$

Следует отметить, что оценка (5) не удовлетворяет в общем случае условию положительности функции $p_n(x)$, однако отрицательные значения этой функции находятся в пределах статистической ошибки и при неограниченном увеличении объема выборки исчезают.

В качестве меры отклонения оценки от действительного значения функции плотности распределения можно принять среднеквадратичный критерий

$$D = \int M\,[p_n(x) - p(x)]^2\,dx, \tag{6}$$

где M — знак математического ожидания по выборке. Используя соотношение (5), получим (6) в виде

$$D = \frac{1}{n}\int\left\{\int\left[f^{\tau}(y)\,r^{-1}\varphi(x)\right]^2 q(y)\,dy - p^2(x)\right\}dx. \tag{7}$$

Рассмотрим случайную величину X, принимающую значения на интервале $(-T,\ T)$, плотность распределения которой представима на этом интервале конечным рядом Фурье. Помеху будем считать аддитивной, т. е. $y = x + \xi$, а ее плотность распределения — симметричной функцией, локализованной на интервале $(-\tau_0,\ \tau_0) \subset (-T,\ T)$. Тогда уравнение (1) примет следующий вид:

$$q(y) = \int\limits_{-T}^{T} K(y - x)\,p(x)\,dx. \tag{8}$$

В качестве ортонормированного базиса функционального пространства Φ_{N+1} возьмем тригонометрическую систему функций, по которой раскладывается плотность $p(x)$:

$$\varphi_0(x) = 1/\sqrt{2T},\quad \varphi_{2m-1}(x) = (1/\sqrt{T})\sin(m\pi x/T),\quad \varphi_{2m}(x) = (1/\sqrt{T})\cos(m\pi x/T),$$
$$m = 1,\ 2,\ \ldots,\ N/2,\ N \text{ — четное.}$$

Непосредственное применение формулы (5) приведет в данном случае к весьма громоздкому выражению, так как система функций $\{f_k(y)\}$ уже не будет ортонормированной. В связи с этим применим следующий искусственный прием, который не отразится на состоятельности оценки, но позволит получить для нее простое выражение.

Преобразуем выборку $\{y_n\}$ в выборку $\{z_n\}$ следующим образом:

$$z_i = \begin{cases} y_i, & \text{если } -T \leqslant y_i \leqslant T;\\ y_i - 2T, & \text{если } y_i > T;\\ y_i + 2T, & \text{если } y_i < -T.\end{cases}$$

Очевидно, что выборка $\{z_n\}$ соответствует плотности распределения $q_1(z)$, связанной с плотностью $p(x)$ интегральным уравнением (8), ядро которого представляет собой функцию $K(x)$, периодически продолженную на всю вещественную ось с периодом $2T$. При этом $q_1(z)$ считается отличной от нуля только на интервале $(-T,\ T)$. Нетрудно показать, что функции $\{\varphi_k(x)\}$ являются собственными функциями интегрального оператора (8) с периодически продолженным ядром. Соответствующие собственные значения равны

$$\lambda_0 = 1,\ \lambda_{2m} = \sqrt{T}\int\limits_{-T}^{T} K(\tau)\,\varphi_{2m}(\tau)\,d\tau,\ \lambda_{2m-1} = \lambda_{2m}.$$

Оценка (5) в этом случае примет следующий вид:

$$p_n(x) = \frac{1}{n}\sum_{i=1}^{n} h^{\tau}(y_i)\,\varphi(x),$$

где матрица $h^{\tau}(y) = \varphi_0(y)/\lambda_0,\ \varphi_1(y)/\lambda_1,\ \ldots,\ \varphi_N(y)/\lambda_N$. Здесь выборка $\{z_n\}$ вновь заменена выборкой $\{y_n\}$, так как в силу периодичности тригонометрических функций $\varphi_k(z_i) = \varphi_k(y_i)$. Аналогично для k-й компоненты матрицы коэффициентов c получим выражение

$$c_k^* = \frac{1}{n\lambda_k} \sum_{i=1}^{n} \varphi_k(y_i).$$

Величину ошибки определим по критерию (6). Принимая во внимание ортонормированность базиса, а также свойства тригонометрических функций, из (7) получим

$$D = \frac{1}{n} \sum_{k=1}^{N} \left(\frac{1}{2T\lambda_k^2} - c_k^2 \right).$$

Отсюда следует верхняя оценка для D, не зависящая от регистрируемого распределения $p(x)$:

$$D < \frac{1}{2Tn} \sum_{k=1}^{N} \frac{1}{\lambda_k^2}.$$

ЛИТЕРАТУРА

1. **Ченцов Н. Н.** Статистические решающие правила и оптимальные выводы.— М.: Наука, 1972.
2. **Айзерман М. А., Браверман Э. М., Розоноэр Л. И.** Метод потенциальных функций в теории обучения машин.— М.: Наука, 1970.
3. **Стратонович Р. Л.** Быстрота сходимости алгоритмов оценки плотности распределения вероятностей.— Известия АН СССР. Сер. Техническая кибернетика, 1969, № 6.
4. **Надарая Э. А.** О непараметрических оценках плотности вероятности и регрессии.— Теория вероятностей и ее применения, 1965, т. 10, вып. 1.
5. **Кораблин М. А.** О синтезе алгоритмов адаптации при наличии помех.— Автоматика и телемеханика, 1971, № 3.
6. **Турчин В. Ф., Козлов В. П., Малкевич М. С.** Использование методов математической статистики для решения некорректных задач.— УФН, 1970, т. 102, вып. 3.

Поступила в редакцию 31 октября 1977 г.;
окончательный вариант — 29 октября 1980 г.

УДК 681.3 : 662.612+533.51

А. Г. ВОРОБЬЕВА, О. П. КОРОБЕЙНИЧЕВ, Л. В. КУЙБИДА,
Л. М. ЛЕВИНА, В. И. МАЛЬЦЕВ, С. В. ПОЛОЗОВ,
И. Н. СКОВОРОДИН

(Новосибирск)

АВТОМАТИЗИРОВАННЫЙ МАСС-СПЕКТРОМЕТРИЧЕСКИЙ КОМПЛЕКС ДЛЯ ИССЛЕДОВАНИЯ СТРУКТУРЫ ПЛАМЕН И ПРОВЕДЕНИЯ ТЕРМИЧЕСКОГО АНАЛИЗА

Для изучения пламен и процессов горения (включая нестационарные быстропротекающие процессы) широкое применение находит динамическая масс-спектрометрия (главным образом время-пролетная) [1]. Вопросы автоматизации целого класса масс-спектрометров с длительностью развертки более 1 мс достаточно подробно разработаны [2]. Этого нельзя сказать об автоматизации время-пролетных масс-спектрометров, имеющих длительность развертки масс-спектров в диапазоне $0{,}01 \div 0{,}1$ мс и ширину масс-спектральных пиков ~ 100 нс. В работе [3] описано устройство «Строб», предназначенное для автоматизированного сбора и первичной обработки информации, поступающей с время-пролетного масс-спектрометра, с последующей выдачей ее на перфоратор. С целью усовершенст-