АКАДЕМИЯ НАУК СССР
СИБИРСКОЕ ОТДЕЛЕНИЕ
АВТОМЕТРИЯ

№ 6 1970

УДК 621.391.19

Г. И. САЛОВ

(*Новосибирск*)

О ПРИМЕНЕНИИ ПРОЦЕССА СТОХАСТИЧЕСКОЙ АППРОКСИМАЦИИ В ГИЛЬБЕРТОВОМ ПРОСТРАНСТВЕ К ЗАДАЧЕ КЛАССИФИКАЦИИ СЛУЧАЙНЫХ СИГНАЛОВ

Введение. Постановка задачи. Мы будем рассматривать следующую ситуацию. Пусть многократно наблюдается на конечном интервале времени $T=(a, b)$ вещественный случайный процесс $\xi(t, \omega)$ — функция t, зависящая от элементарного события или исхода ω из множества всех исходов Ω (подробнее см., например, [1]). Каждый раз $\xi(t, \omega)$ представляет собой (в простейшем варианте) случайный процесс $\xi_1(t, \omega)$, где $\omega \in \Omega_1$, или $\xi_2(t, \omega)$, где $\omega \in \Omega_2$ ($\Omega_1+\Omega_2=\Omega$, $\Omega_1\Omega_2=0$). Задача состоит здесь в установлении правила отнесения наблюдаемой выборочной функции (реализации) процесса $\xi(t, \omega)$ к процессу $\xi_1(t, \omega)$ или же к $\xi_2(t, \omega)$. Предполагается, что экспериментатор имеет в своем распоряжении n независимых наблюденных реализаций $\xi(t, \omega_i)$, $i=1, \ldots, n$, процесса $\xi(t, \omega)$ и знает, откуда они происходят, т. е. знает, что ω_i принадлежит именно Ω_1 или Ω_2.

К безошибочным заключениям о $\xi(t, \omega)$ приводило бы использование идеального функционала $F[\xi(t, \omega)] \equiv y(\omega)$, который принимает, например, положительные значения, если $\omega \in \Omega_1$, и 0 или отрицательные значения, если $\omega \in \Omega_2$. К сожалению, возможность построить такой функционал встречается довольно редко, поэтому приходится ограничиваться меньшим.

Один из известных методов построения функционала $F[\xi(t, \omega)]$, на основании которого принимается решение о $\xi(t, \omega)$, состоит (см., например, [2], гл. IV) из выбора конечного числа m функционалов $F_j[\xi(t, \omega)]$ и отыскания вектора $c=c_1, \ldots, c_m$, минимизирующего математическое ожидание

$$I(c)=E\left\{\operatorname{sgn} y(\omega)-\sum_{j=1}^{m} c_j F_j[\xi(t, \omega)]\right\}^2, \tag{1}$$

где, как обычно, $\operatorname{sgn} x=\begin{cases} 0, & \text{если } x=0; \\ \dfrac{x}{|x|}, & \text{если } x \neq 0. \end{cases}$

Вектор c отыскивается с помощью процедуры стохастической аппроксимации в m-мерном (евклидовом) пространстве, где элементами являются m-членные комплексы вещественных чисел. Наибольшую трудность здесь, очевидно, представляет выбор и воспроизведение (реализация) функционалов $F_j\,[\xi\,(t,\ \omega)]$. Естественно желание уменьшить число функционалов до одного и «освободиться» от его выбора.

Часто процессы могут быть записаны, и наиболее просто воспроизвести функционал

$$F\,[\xi\,(t,\omega)] = \int_a^b \xi\,(t,\ \omega)\,x\,(t)\,d\,t. \tag{2}$$

В широком классе случаев (см. ниже) такой вид имеет всякий линейный функционал.

Теперь уже необходимо отыскивать функцию $x(t)=\theta(t)$, порождающую функционал (2) и минимизирующую математическое ожидание, аналогичное (1).

Оказывается, принципиально возможно построение $\theta(t)$, а значит, наилучшего линейного функционала с помощью процедуры стохастической аппроксимации в гильбертовом пространстве. Показать это и является целью данной работы.

Предполагаемые условия. Предварительные сведения. Введем следующие условия. Для всех $t \in T$

$$E\,\{\xi\,(t,\ \omega)\}^2 < \infty \ \text{ и } \ E\left\{\int_a^b \xi^2\,(t,\ \ \omega)\,d\,t\right\}^2 = A^2 < \infty, \tag{3}$$

ковариационная функция процесса $\xi(t,\ \omega)$ положительно определенная, т. е. для всякой функции $\varphi(t)$ с интегрируемым квадратом на интервале $(a,\ b)$ имеет место

$$\int_a^b\int_a^b E\,\{\xi\,(s,\ \omega)\,\xi\,(t,\ \omega)\}\,\varphi\,(s)\,\varphi\,(t)\,d\,s\,d\,t > 0, \tag{4}$$

как только

$$\int_a^b \varphi^2\ \ (t)\,d\,t > 0.$$

Существует интервал в T такой, что почти во всех его точках t

$$E\,\{\xi\,(t,\ \omega)\ \operatorname{sgn}\,y\,(\omega)\} \neq 0. \tag{5}$$

Условие (3) естественно, а (4) выполняется, например, для процессов с невырожденным гауссовским распределением.

Из (3) следует, что почти все, т. е. за исключением множества $\Omega-\Omega_0$ вероятности 0, выборочные функции случайного процесса $\xi(t,\ \omega)$ являются функциями с интегрируемым квадратом на интервале $(a,\ b)$. Мы предположим ради простоты большее, именно интегрируемость всех выборочных функций, однако основные выводы проведем так, что они справедливы для общего случая. Другими словами, пусть теперь все выборочные функции принадлежат вещественному гильбертову пространству $L^2(a,\ b)$ функций с интегрируемым квадратом (в смысле лебеговского интегрирования) на интервале $(a,\ b)$. Как известно, в этом пространстве не различают две функции, совпадающие почти во всех точках t, а скалярное произведение $(f,\ g)$ элементов f и g из $L^2(a,\ b)$ и норма $\|f\|$ элемента f определяются как

$$(f,\ g) = \int_a^b f\,(t)\,g\,(t)\,dt;\ \ \|f\| = \sqrt{(f,\ f)} = \left(\int_a^b f^2\,(t)\,d\,t\right)^{1/2}.$$

Сходимость в этом пространстве — это то, что называют также сходимостью в среднем квадратичном. Согласно теореме Ф. Рисса [3, 4], всякий линейный функционал в $L^2(a, b)$ имеет вид (2), причем $x(t)$ есть функция из $L^2(a, b)$.

Если в (2) $x(t)$ принадлежит $L^2(a, b)$, то при соответствующих условиях измеримости процесса $\xi(t, \omega)$ [5, 6], которые на практике обычно удовлетворяются, по теореме Фубини [3, 4] функционал $F[\xi(t, \omega)]$ является случайной величиной, т. е. измеримой функцией от ω на Ω. Случайной величиной является также норма $\|\xi(t, \omega)\|$. В такой ситуации $\xi(t, \omega)$ можно рассматривать как случайный элемент в $L^2(a, b)$ (см., например, [7]).

Подставив (2) в (1), имеем

$$I[x(t)] = \int_{\Omega} [\operatorname{sgn} y(\omega) - \int_a^b \xi(t, \omega) x(t) dt]^2 dP(\omega), \tag{6}$$

где P — распределение вероятностей, неизвестное наблюдателю. Функционал $I[x(t)]$ определен для всех функций $x(t)$ пространства $L^2(a, b)$. Возьмем произвольный элемент $h \in L^2(a, b)$ и произвольное вещественное число α. Если функционал $I[x(t)]$ достигает экстремума на $x(t) = \theta(t)$, то его первая вариация в этой точке равна нулю при любом приращении $h(t)$. Равенство вариации нулю равносильно тому, что в каждой точке t

$$E\{[\operatorname{sgn} y(\omega) - \int_a^b \xi(s, \omega)\theta(s) ds]\, \xi(t, \omega)\} = 0. \tag{7}$$

Для второй вариации в соответствии с вышепринятым условие (4) влечет неравенство

$$\delta^2 I(x, h) = \frac{1}{2}\left[\frac{d^2 I(x + \alpha h)}{d\alpha^2}\right]_{\alpha = 0} =$$

$$= \int_{\Omega} \left\{\int_a^b \xi(t, \omega) h(t) dt\right\}^2 dP(\omega) = \int_a^b \int_a^b E\{\xi(s, \omega)\xi(t, \omega)\} h(s) h(t) ds\, dt \geqslant 0,$$

независимое от $x(t) \in L^2(a, b)$. Это означает, что в точке $x(t) = \theta(t)$ функционал $I[x(t)]$ имеет абсолютный минимум (см., например, в [8] теорему 3.5.1), т. е.

$$I[\theta(t)] \leqslant I[x(t)].$$

Таким образом, задача об определении наилучшего линейного функционала сводится к решению функционального уравнения (7). При условии (5) это уравнение имеет нетривиальное решение, т. е. $\theta(t) \neq 0$.

Построение оценок функции $\theta(t)$ по реализациям случайного процесса. Стохастическую аппроксимацию в гильбертовом пространстве впервые рассматривал, по-видимому, Л. Шметтерер [9]. Однако условия его результатов труднопроверяемые. Вентеру [10] удалось ослабить условия. Он доказал, в частности, следующее (в действительности предположения его теоремы 3 более слабые, чем нижеприведенные).

Пусть H — абстрактное сепарабельное гильбертово пространство и $\theta \in H$. Пусть $X_1(\omega)$ — произвольный случайный элемент (в указанном выше смысле) в H. Определяется последовательность случайных элементов $\{X_n(\omega)\}$ в H по формуле

$$X_{n+1}(\omega) = X_n(\omega) - S_n(X_n(\omega)) + U_n(\omega). \tag{8}$$

Если

1) S_n есть отображение (функция) H в H, такое, что для любых $x \in H$ и $n \geqslant 1$

$$\| S_n(x) \|^2 \leqslant \beta_n \| x - \theta \|^2,$$

где $\{\beta_n\}$ — последовательность положительных чисел, имеющая свойство $\sum \beta_n < \infty$,

2) для любых $x \in H$ и $\varepsilon > 0$

$$\inf_{\varepsilon \leqslant \|x-\theta\| \leqslant \varepsilon^{-1}} (x - \theta, S_n(x)) = c_n(\varepsilon) \geqslant 0 \text{ и } \sum c_n(\varepsilon) = \infty,$$

3) $\{U_n(\omega)\}$ — последовательность случайных элементов в H, такая, что с вероятностью 1

$$\| E\{U_n(\omega) | X_1(\omega), \ldots, X_n(\omega)\} \| = 0 \text{ и } \sum E\{\| U_n(\omega) \|^2\} < \infty,$$

то при $n \to \infty$

$$P\{\| X_n(\omega) - \theta \| \to 0\} = 1; \quad E\{\| X_n(\omega) - \theta \|^2\} \to 0.$$

Пространство $L^2(a, b)$ является реализацией названного абстрактного пространства H, поэтому мы можем воспользоваться этим результатом. Для любого целого $n \geqslant 1$ и $x(t) \in L^2(a, b)$ положим

$$Z_n[x(t), \omega] = [\operatorname{sgn} y_n(\omega) - \int_a^b \xi_n(s, \omega)\, x(s)\, ds]\, \xi_n(t, \omega), \tag{9}$$

где $\{y_n(\omega)\}$ — последовательность независимых случайных величин с одной и той же функцией распределения, такой же, как у $y(\omega)$: $\{\xi_n(t, \omega)\}$ — последовательность независимых случайных элементов в $L^2(a, b)$ со статистическими свойствами, как у $\xi(t, \omega)$. Теперь ω — сокращение бесконечной последовательности вида $\omega = (\omega_1, \omega_2, \ldots)$, где каждая координата $\omega_i \in \Omega$, так что $\operatorname{sgn} y_n(\omega)$ и $\xi_n(t, \omega)$ зависят только от ω_n. В соответствии со сказанным выше $Z_n[x(t), \omega]$ также является случайным элементом в $L^2(a, b)$.

Пусть $x_1(t, \omega)$ — любой случайный элемент, принадлежащий $L^2(a, b)$, иначе говоря, любая функция из $L^2(a, b)$. Определяем последовательность случайных функций по формуле

$$x_{n+1}(t, \omega) = x_n(t, \omega) - S_n[x_n(t, \omega)] + U_n(t, \omega), \tag{10}$$

где

$$S_n[x(t)] = -a_n E\{Z_n[x(t), \omega]\};$$

$$U_n(t, \omega) = \begin{cases} S_n[x_n(t, \omega)]; & \omega_n \in \Omega - \Omega_0; \\ a_n Z_n[x_n(t, \omega), \omega] + S_n[x_n(t, \omega)]; & \omega_n \in \Omega_0; \end{cases} \tag{11}$$

$\{a_n\} = \left\{\frac{c}{n^{1-\gamma}}\right\}$, $0 \leqslant \gamma < \frac{1}{2}$, — последовательность положительных чисел, удовлетворяющая условию

$$\sum a_n = \infty; \quad \sum a_n^2 < \infty. \tag{12}$$

Чтобы убедиться здесь в сходимости $x_n(t, \omega)$ к $\theta(t)$, достаточно проверить выполнение условий 1)—3). Начнем с условия 1).

Согласно (9) и (7), получаем, применяя дважды неравенство Буняковского — Шварца,

$$\| - a_n E \{Z_n [x(t), \omega]\}\|^2 = a_n^2 \int_a^b E^2 \{[\operatorname{sgn} y_n(\omega) - \int_a^b \xi_n(s, \omega)\, x(s)\, ds] \times$$

$$\times \xi_n(t, \omega)\}\, dt = a_n^2 \int_a^b E^2 \{[\operatorname{sgn} y_n(\omega) - \int_a^b \xi_n(s, \omega)\, (x(s) + \theta(s) -$$

$$- \theta(s))\, ds]\, \xi_n(t, \omega)\}\, dt < a_n^2 E^2 \left\{\int_a^b \xi^2(t, \omega)\, dt\right\} \int_a^b (x(t) - \theta(t))^2\, dt.$$

$$a$$

$$= a_n \int_a^b \int_a^b E\{\xi(s, \omega)\, \xi(t, \omega)\}\, (x(s) - \theta(s))\, (x(t) - \theta(t))\, ds\, dt,$$

так что соотношения (4) и (12) влекут выполнение условия 2).

Проверку первой части условия 3) мы не приводим. Поскольку имеет место

$$E\{\|U_n(t, \omega)\|^2\} = a_n^2 E\left\{\int_a^b \{Z_n [x_n(t, \omega), \omega] + \frac{1}{a_n} S_n [x_n(t, \omega)]^2 dt\right\} \leqslant$$

$$\leqslant a_n^2 \int_a^b E\{Z_n [x_n(t, \omega), \omega]\}^2\, dt = a_n^2 d_n < \infty,$$

то для доказательства сходимости ряда $\sum E\{\|U_n(t, \omega)\|^2\}$ достаточно доказать сходимость ряда с общим членом, равным правой части последнего неравенства. Применим предельную форму признака Раабе [11], основанного на отношении двух следующих друг за другом членов ряда. Для члена с индексом $n+1$ нового ряда справедливо

$$a_{n+1}^2 d_{n+1} = a_{n+1}^2 \int_a^b E\{[\operatorname{sgn} y_{n+1}(\omega) - \int_a^b \xi_{n+1}(s, \omega)\, (x_n(s, \omega) +$$

$$+ a_n Z_n [x_n(t, \omega), \omega])\, ds]\, \xi_{n+1}(t, \omega)\}^2\, dt \leqslant$$

$$\leqslant a_{n+1}^2 d_n \left(1 + 2a_n E^{1/2}\{\int_a^b \xi^2(t, \omega)\, dt\}^2 + a_n^2 E\left\{\int_a^b \xi^2(t, \omega)\, dt\right\}^2\right),$$

что легко получить, применяя неравенство Буняковского — Шварца. Отсюда предел варианты Раабе, согласно (3) и (12), не менее

$$\lim_{n\to\infty} n\left(\frac{a_n^2 d_n}{a_{n+1}^2 d_n (1 + 2a_n A + a_n^2 A^2)} - 1\right) = \lim_{n\to\infty} n\left(\frac{a_n^2}{a_{n+1}^2} - 1\right),$$

т. е. не менее предела для сходящегося ряда (12). Это завершает доказательство. Практически формула (10) выглядит следующим образом:

$$x_{n+1}(t) = x_n(t) + a_n [\operatorname{sgn} y(\omega_n) - \int_a^b \xi(s, \omega_n)\, x_n(s)\, ds]\, \xi(t, \omega_n), \quad (10')$$

где $\{\xi(t, \omega_n)\}$ — последовательность независимых наблюдений случайного процесса $\xi(t, \omega)$.

Рассмотренная процедура и выводы непосредственно переносятся, с очевидными изменениями, на решение известного в теории обнаружения (фильтрации) сигналов интегрального уравнения

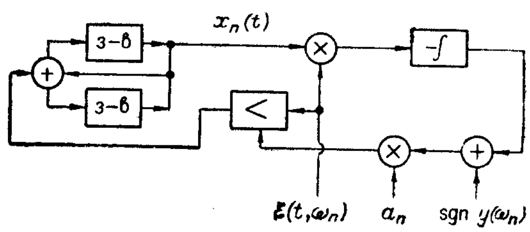
$$\int_0^T E\{\xi(t, \omega)\,\xi(s, \omega)\}\,\theta(s)\,ds = g(t)(E\{\xi(0, \omega)\,\xi(t, \omega)\}),$$

где $g(t)$ — заданная функция из $L^2(0, T)$, отличная от нуля. При этом ограничение (5), естественно, снимается, а последовательность $\{x_n(t)\}$ определяется равенством

$$x_{n+1}(t) = x_n(t) + a_n\left(g(t, \omega_n) - \xi(t, \omega_n)\int_0^T \xi(s, \omega_n)\,x_n(s)\,ds\right),$$

где $g(t, \omega_n) \equiv g(t)\;(\xi(0, \omega_n)\,\xi(t, \omega_n))$.

О реализации на практике процесса построения оценки для $\theta(t)$. Осуществить процесс построения $\theta(t)$ можно, по крайней мере в принципе, по схеме, приведенной на рисунке. Основными элементами схемы являются интегратор, усилитель с управляемым коэффициентом усиления, моделирующий умножение $\xi(t, \omega_n)$ на соответствующее число, и два устройства, записывающих — воспроизводящих функции времени. Пусть в одном из устройств записана и воспроизводится функция времени $x_n(t)$; тогда функция $x_{n+1}(t)$, равная сумме $x_n(t)$ и получающейся добавки, соответствующей $\xi(t, \omega_n), x_n(t)$, $\operatorname{sgn} y(\omega_n)$ и a_n, записывается в другом устройстве и т. д. Если интервал T достаточно короткий, то функцию этих двух устройств может выполнять один магнитный барабан.



Реализации $\xi(t, \omega_n)$ могут браться непосредственно или с магнитной ленты, или с магнитного барабана. Отметим также, что в качестве $\xi(t, \omega_n)$ могут быть взяты, когда это подходит, функции времени на выходе автоматического спектроанализатора фильтрующего типа, описывающие выборочные функции другого исследуемого процесса в частотной области.

Заключение

Рассмотрена задача классификации случайных сигналов. Показано, что наилучшее (в определенном смысле) линейное правило (функционал) классификации может быть найдено по реализациям наблюдаемых случайных сигналов с помощью процесса стохастической аппроксимации. Этот способ позволяет также строить решение интегрального уравнения Фредгольма первого рода в случае, когда ядром является неизвестная ковариационная функция процесса. Отмечена возможность реализации на практике исследуемого подхода.

Автор выражает искреннюю признательность А. Г. Сенину за критические замечания по данной работе.

ЛИТЕРАТУРА

1. Г. Крамер, М. Лидбеттер. Стационарные случайные процессы. М., «Мир», 1969.
2. Я. З. Цыпкин. Адаптация и обучение в автоматических системах. М., «Наука», 1968.
3. Ф. Рисс и Б. Секефальви-Надь. Лекции по функциональному анализу. М., Изд-во иностр. лит., 1954.
4. Р. Эдвардс. Функциональный анализ. М., «Мир», 1969.
5. Дж. Л. Дуб. Вероятностные процессы. М., Изд-во иностр. лит., 1956.
6. И. И. Гихман, А. В. Скороход. Введение в теорию случайных процессов. М., «Наука», 1965.
7. У. Гренандер. Вероятности на алгебраических структурах. М., «Мир», 1965.
8. С. Г. Михлин. Курс математической физики. М., «Наука», 1968.
9. L. Schmetterer. Sur l'iteration stochastique.— Colloques Internationaux du centre national de la recherche scientifique, 1958, 87.
10. J. H. Venter. On Dvoretzky stochastic approximation theorems.— Annals of Mathematical Statistics, 1966, v. 37, № 6.
11. Г. М. Фихтенгольц. Курс дифференциального и интегрального исчисления, т. 2. М., Физматгиз, 1962.

Поступила в редакцию
2 июля 1970 г.