Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Министерство высшего и среднего специального образования СССР
Министерство высшего и среднего специального образования РСФСР
Сибирское отделение АН СССР
Томский государственный университет им. В. В. Куйбышева
Сибирский физико-технический институт им. В. Д. Кузнецова
Красноярский ВЦ СО АН СССР

VI СОВЕЩАНИЕ-СЕМИНАР ПО НЕПАРАМЕТРИЧЕСКИМ И РОБАСТНЫМ МЕТОДАМ СТАТИСТИКИ В КИБЕРНЕТИКЕ

(тезисы докладов)

ЧАСТЬ I

Томск — 1987

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

боковые улицы терялись в старом
среди тротуаров, и где-то слышалась
слушал кукушку, удивляясь, что ей
, и шум был покойный, мягкий, почти
сидел где-нибудь на лавочке, устроен-

нова шел по улицам-просекам.
отуарам текла нарядная толпа, на траве
было много детей и цветов. Гурнов поду-
наверное, удобнее, чем в тесной, всегда пе-
стоящий москвич неисправим. Если, учась
ижникам Красной площади и, став взрослым,
л в бурливой, неповторимой, единственной в
ие, то отовсюду тебя потянет в Москву, как ни
ибирский рай. Люди испытывают силы тяготения
их слабостей и многого другого, чего даже не

лся, гуляя по улицам городка, и ему казалось, что
о необычное. Он посмеивался над своим ожида-
есть лет Гурнов убедился, что чудес не бывает, но
езаботно гулять, сидеть, смотреть на женщин и

кинотеатра, тоже новенького, изящной современной

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Министерство высшего и среднего специального образования СССР
Министерство высшего и среднего специального образования РСФСР
Сибирское отделение АН СССР
Томский государственный университет им. В.В.Куйбышева
Сибирский физико-технический институт им. В.Д.Кузнецова
Красноярский ВЦ СО АН СССР

1-660936

VI СОВЕЩАНИЕ - СЕМИНАР ПО НЕПАРАМЕТРИЧЕСКИМ И РОБАСТНЫМ МЕТОДАМ СТАТИСТИКИ В КИБЕРНЕТИКЕ

(тезисы докладов)

Часть I

НАУЧНАЯ БИБЛИОТЕКА ТГУ

Томск - 1987

007:519.2(063)

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Материалы УI Всесоюзной школы-семинара по непараметрическим робастным методам статистики в кибернетике. Часть I//Под ред. Ф.П.Тарасенко.- Томск, 1987.- 239с.-Ip.90к. 200экз.

В сборнике рассматриваются вопросы по разработке и исследованию статистических процедур обработки экспериментальных данных при весьма слабых априорных предположениях.

Для статистиков-теоретиков и многочисленных пользователей прикладной статистики.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

С О Д Е Р Ж А Н И Е

Р.А.Абусев (Пермь). О групповой статистической классификации с помощью непараметрических оценок плотности..4-9

А.Л.Агафонов, Е.Л.Даниленко (Омск). Рекуррентные оценки линейной регрессии при нарушении классических предпосылок I0-I7

С.Е.Ададуров, А.В.Макшанов (Ленинград). Устойчивая фильтрация в динамических системах случайной структуры I8-2I

В.К.Анциферов, И.А.Пестунов (Красноярск). Об одном критерии выбора оптимальных признаков, основанном на понятии дифференциальной энтропии 22-26

Э.К.Атаянц, В.П.Рынин (Рязань). Эффективная модификация алгоритма К-ближайших соседей 27-34

И.И.Балтрунас, Р.С.Вальтерите (Каунас). Непараметрическая идентификация нелинейных стохастических процессов с трендом 35-4I

И.И.Балтрунас, В.Ю.Рудзекене (Каунас). Непараметрические оценки нелинейной функции авторегрессии 42-44

Ю.Н.Благовещенский (Москва). Непараметрические оценки по случайно цензурированным справа выборкам 45-65

В.З.Борисов (Якутск). Об асимптотической оптимальности двухэтапной процедуры оценивания 65-67

Т.В.Варядченко, О.Ю.Кульчицкий (Ленинград). Поиск статистических закономерностей в обучающих массивах данных 68-79

Т.В.Варядченко, О.Ю.Кульчицкий (Ленинград). Непараметрическая оценка спектральной плотности случайного процесса методом нелинейного резонансного фильтра.... 80-92

Н.О.Вильчевский, Г.Л.Шевляков (Ленинград). Аксиоматический подход к выбору критериев оценивания 93-97

С.Э.Воробейчиков, В.В.Конев (Томск). О последовательном методе обнаружения скачкообразных изменений параметров случайных процессов рекуррентного типа 98-II6

Л.С.Воробьев (Минск). Исследование устойчивости $C(\alpha)$-критерия при нарушении условия независимости случайных величин II7-I2I

А.Б.Глаз (Рига). О построении эффективных решающих правил на коротких обучающих выборках в задачах распознавания образов I22-I24

В.А.Гныря, А.А.Светлаков (Томск). Алгоритм устранения аномаль-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ных результатов измерений 125-131
С.Ю.Городецкий, Ю.И.Неймарк (Горький). Построение процедур поиска решения на основе адаптивных стохастических моделей 132-136
В.В.Губарев, Е.Н.Наумова, А.А.Наумов (Новосибирск). Исследование возможности применения бутстрэп-метода в задачах статистического анализа 137-147
А.М.Дубров, Н.Д.Вылегжанин, В.Б.Королев (Москва). Робастное оценивание в многомерном статистическом анализе 148-150
А.А.Иванилов (Красноярск). Алгоритмы управления динамическими объектами с запаздыванием в условиях непараметрической неопределенности 151-154
В.А.Ильин, Н.А.Харламова (Томск). Синтез робастной процедуры распознавания и определения ориентации объектов по зашумленным двумерным изображениям в системе технического зрения робота 155-164
Е.В.Ильин (Красноярск). Некоторые модификации непараметрических алгоритмов оптимизации 165-171
В.А.Йоала (Таллин). Применение разностного оператора для анализа временных рядов 172-173
М.П.Каминский (Ленинград). Регрессия квантилей как непараметрическая процедура прикладной статистики. 174-179
С.А.Кислицин, М.П.Ангелов, В.М.Мохов (Томск). Обработка изображений на основе модели множественной регрессии . . . 180-187
А.В.Китаева (Томск). Оценки медианного типа для коэффициентов линейной регрессионной модели 188-191
М.М.Коган, Ю.И.Неймарк (Горький). Об особенностях сходимости рекуррентных процедур типа метода наименьших квадратов. 192-196
Т.Х.-А.Колло (Тарту). Исследование асимптотического поведения функций от математического ожидания и ковариационной матрицы 197-201
В.В.Конев, С.М.Пергаменщиков (Томск). К последовательному оцениванию параметров стохастических динамических систем... 202-214
В.И.Котюков, В.А.Калинкин (Новосибирск). Структурно устойчивые многофакторные модели 215-220
Г.М.Кошкин (Томск). О среднем и дисперсии функций от статистик 221-227
И.А.Кожевникова (Москва). Среднеквадратическое отклонение непараметрических спектральных характеристик с временным сдвигом 228-233
Э.А.Клёкис (Каунас). Методы оптимальной и адаптивной фильтрации в случае помех, содержащих выбросы 234-238

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

О ГРУППОВОЙ СТАТИСТИЧЕСКОЙ КЛАССИФИКАЦИИ С ПОМОЩЬЮ НЕПАРАМЕТРИЧЕСКИХ ОЦЕНОК ПЛОТНОСТИ

Р.А.Абусев

Обобщенная задача групповой классификации состоит в следующем.

Пусть класс объектов π_i описывается обучающей последовательностью π_{i0} случайных матриц $X_{i1}, X_{i2}, \dots, X_{in_i}$, длины n_i, каждая из которых имеет k строк и m столбцов $i = 1,2,\dots,M$. Пусть далее π_{00} – последовательность (группа) длины n_0 случайных матриц $X_{01}, X_{02}, \dots, X_{0n_0}$ той же размерности. Требуется отнести группу π_{00} к одной из совокупностей $\pi_1, \pi_2, \dots, \pi_M$ как единое целое.

Обозначим $p(x/\pi_i)$ условную плотность распределения объектов в совокупности π_i и $p(\pi_{00}/\pi_i)$ – условную плотность распределения выборки π_{00} при гипотезе H_i, т.е. в предположении, что π_{00} извлечена из совокупности π_i, $i = 1,2,\dots,M$.

Построение правила классификации группы π_{00} геометрически может быть интерпретировано как разбиение пространства E всевозможных групп объема n_0 на подпространства $E_1, E_2, \dots, E_M$ такие, что $E_i \cap E_j = \emptyset$ при $i \neq j$, $\cup E_i = E$. Тогда если π_{00} принадлежит E_j, то принимается решение, что π_{00} извлечена из совокупности π_j, $j = 1,2,\dots,M$.

Рассмотрим решающее правило

$\pi_{00} \prec \pi_j$, если для всех $z \neq j$

$$\sum_{i=1}^{M} c_{ji} q_i p(\pi_{00}/\pi_i) < \sum_{i=1}^{M} c_{zi} q_i p(\pi_{00}/\pi_i), \qquad (1)$$

$z, j = 1,2,\dots,M$, где c_{ji} – неотрицательная цена оши-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

бочного отнесения группы π_{oo} из совокупности π_i к совокупности π_j , q_i – априорная вероятность совокупности π_i , $i,j = 1,2,\ldots,M$. Можно доказать, как обобщение леммы Неймана–Пирсона [1] , что правило (1) является оптимальным в том смысле, что оно минимизирует байесовский риск R равный

$$R = \sum_{i=1}^{M} \sum_{j=1}^{M} c_{ji} q_i \alpha_{ji} , \qquad (2)$$

где α_{ji} – вероятность ошибочного отнесения группы π_{oo} из класса π_i к классу π_j , $i,j = 1,2,\ldots, M$.

Рассмотрим частные случаи. Если $m = k$ и матрицы X_{i1} , X_{i2} , ... , X_{in_i} независимы, положительно определены, симметричны и каждая имеет одно и то же распределение Уишарта $W_{\Sigma_i}(X/\ell)$ с матрицей Σ_i и числом степеней свободы ℓ , $\ell > k$ [1] , тогда $P(\pi_{oo}/\pi_i) = P_{\Sigma_i}(X/\ell)$ – плотность распределений Уишарта. В реальных задачах вышеуказанные матрицы могут быть, например, корреляционными или ковариационными. Если $m = 1$ и случайные матрицы-столбцы X_{ij} независимы , $j = 1,2,\ldots, n_i$, $i = 0,1,\ldots, M$, то приходим к задаче групповой классификации, сформулированной в работе [3].

Если $m = 1$, $n_o = 1$ и все случайные матрицы-столбцы X_{ij} $j = 1,2,\ldots,n_i$ $i = 0,1,2,\ldots, M$ независимы, то приходим к хорошо известной традиционной задаче классификации одного k –мерного объекта [1,2] , которую еще называют задачей поточечной классификации [3–5] .

В случае $M = 2$, $m = 1$, $C_{11} = C_{22} = 0$ из (1) получим оптимальный байесовский метод

$$\pi_{oo} \subset \pi_1, \text{ если } \frac{\omega_1 P(\pi_{oo}/\pi_1)}{\omega_2 P(\pi_{oo}/\pi_2)} \geq 1 \qquad (3)$$

классификации группы π_{oo} из n_o k –мерных объектов x_{o1} , x_{o2} , ... , x_{on_o} , принадлежащей одной из двух совокупностей π_1 или π_2 с величиной риска

$$R = \omega_1 \int_{E_2} P(\pi_{oo}/\pi_1)\, d\pi_{oo} + \omega_2 \int_{E_1} P(\pi_{oo}/\pi_2)\, d\pi_{oo} , \qquad (4)$$

где $\omega_1 = q_1 c_{21}$; $\omega_2 = q_2 c_{12}$; $E_1 \cup E_2 = E$.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Использование непараметрических оценок плотности для решения задачи групповой классификации

Пусть $P(x/\pi_i)$ и $P(\pi_{00}/\pi_i)$ не известны, $i = 1,2$. Воспользуемся непараметрическими оценками [7,8]:

$$P_{n_i}(x/\pi_i) = \frac{1}{n_i h_i^k} \sum_{j=1}^{n_i} K\left(\frac{x - x_{ij}}{h_i}\right) \qquad (5)$$

плотности $P(x/\pi_i)$, $i = 1,2$, где $K(y)$ - ядра Парзена.

Рассмотрим следующее простое правило групповой классификации:

$$\pi_{00} < \pi_1, \text{ если } \frac{\omega_1}{n_0} \sum_{j=1}^{n_0} P_{n_1}(x_{0j}/\pi_1) \geq \frac{\omega_2}{n_0} \sum_{j=1}^{n_0} P_{n_2}(x_{0j}/\pi_2). \qquad (6)$$

Если взять в качестве $K\left(\frac{x - x_{ij}}{h_i}\right)$ ядро нормального типа [8,10], то последнее запишется

$$\pi_{00} < \pi_1, \text{ если } \frac{\omega_1 (n_0 n_1)^{-1}}{h_1^k |\Sigma_1|^{1/2}} \sum_{i=1}^{n_0} \sum_{j=1}^{n_1} \exp\left\{-\frac{1}{2h_1^2}(x_{0i} - x_{1j})' \Sigma_1^{-1} (x_{0i} - x_{1j})\right\} \geq$$

$$\geq \frac{\omega_2 (n_0 n_2)^{-1}}{h_2^k |\Sigma_2|^{1/2}} \sum_{i=1}^{n_0} \sum_{j=1}^{n_2} \exp\left\{-\frac{1}{2h_2^2}(x_{0i} - x_{2j})' \Sigma_2^{-1} (x_{0i} - x_{2j})\right\}. \qquad (7)$$

Здесь матрицы Σ_i порядка $k \times k$ предполагаются заданными, в противном случае заменяются на выборочные матрицы S_i, $i = 1,2$.

Рассмотрим частные случаи (7). Положим $h_1 = h_2 = 1$, $\Sigma_1 = \Sigma_2 = I$, $\omega_1 = \omega_2 = 1$. Тогда из (7) получим правило

$$\pi_{00} < \pi_1, \text{ если } \frac{1}{n_0 n_1} \sum_{i=1}^{n_0} \sum_{j=1}^{n_1} \exp\left\{-\frac{1}{2}\rho^2(x_{0i}, x_{1j})\right\} \geq$$

$$\geq \frac{1}{n_0 n_2} \sum_{i=1}^{n_0} \sum_{j=1}^{n_2} \exp\left\{-\frac{1}{2}\rho^2(x_{0i}, x_{2j})\right\}, \qquad (8)$$

где I - единичная матрица, $\rho^2(x,y)$ - просто квадрат евклидова расстояния. Далее воспользуемся неравенством $\exp(-z^2) \leq \frac{1}{1+z^2}$ и из (8) получим

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\tilde{x}_{00} \subset \tilde{x}_1, \text{ если } \frac{1}{n_0 n_1} \sum_{i=1}^{n_0} \sum_{j=1}^{n_1} \frac{1}{1 + \frac{1}{2} \rho^2(x_{0i}, x_{1j})} \geq$$

$$\geq \frac{1}{n_0 n_2} \sum_{i=1}^{n_0} \sum_{j=1}^{n_2} \frac{1}{1 + \frac{1}{2} \rho^2(x_{0i}, x_{2j})} . \qquad (9)$$

Очевидно, что правила (8), (9) совпадают с хорошо известными методами потенциальных функций [II] для решения задачи автоматической классификации. Из них же при $n_0 = 1$ следуют методы потенциальных функций [II] для решения задач поточечной классификации [2-6].

Верхние границы для байесовского риска при групповой классификации

В случае поточечной классификации с помощью непараметрических оценок оценки для верхних границ вероятности ошибок классификации получены в [IO], [6].

Подобным образом в случае групповой классификации с помощью правила (3) верхняя граница $\bar{R}$ для риска R имеет вид

$$R = \int \min \left[\omega_1 \rho(\tilde{x}_{00}/\tilde{x}_1), \omega_2 \rho(\tilde{x}_{00}/\tilde{x}_2) \right] d\tilde{x}_{00} \leq$$

$$\leq \omega_1^s \omega_2^{1-s} \int \rho^s(\tilde{x}_{00}/\tilde{x}_1) \rho^{1-s}(\tilde{x}_{00}/\tilde{x}_2) \, d\tilde{x}_{00} = \bar{R} .$$

Возьмем оценки $\rho_{n_i}(\tilde{x}_{00}/\tilde{x}_i) = \prod \rho_{n_i}(x_{0j}/\tilde{x}_i)$ для $\rho(\tilde{x}_{00}/\tilde{x}_i)$, $i = 1.2$. Тогда оценка $\bar{R}^*$ для $\bar{R}$ может быть получена из последнего равенства при $s = 1/2$:

$$\bar{R}^* = (\omega_1 \omega_2)^{1/2} \int \left[\prod_{j=1}^{n_0} \rho_{n_1}(x_{0j}/\tilde{x}_1) \rho_{n_2}(x_{0j}/\tilde{x}_2) \right]^{1/2} dx_{01} \dots dx_{0n_0} =$$

$$= (\omega_1 \omega_2)^{1/2} \prod_{j=1}^{n_0} \int \left[\rho_{n_1}(x_{0j}/\tilde{x}_1) \rho_{n_2}(x_{0j}/\tilde{x}_2) \right]^{1/2} dx_{0j} . \qquad (10)$$

Нахождение оценки $\bar{R}^*$ и границ для $\bar{R}$ сводится к вычислению интегралов в правой части (10).

Одну границу $\bar{R}_1'$ для $\bar{R}$ получим из (10), используя результатн работы [10].

Другая граница $\bar{R}^*_n$ для $\bar{R}$ может быть получена используя неравенство $\frac{1}{n} \sum_{i=1}^{n} a_i \geq \left(\prod_{i=1}^{n} a_i \right)^{1/n}$, $a_i \geq 0$, $i = 1, 2, \dots, n$, и конкретные типы ядер.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Далее границу $\overline{R_3^*}$ для $\bar{R}$ можно получить на основании неравенства Иенсена, которое в [6] использовано в виде

$$\int_{R^K} [P_{n_1}(x/\pi_1)\cdot P_{n_2}(x/\pi_2)]^{1/2} dx \le \sqrt{A}\left[\int_Q P_{n_1}(x/\pi_1)\cdot P_{n_2}(x/\pi_2)\,dx\right]^{1/2},$$

где Q – часть пространства R^K, в которой $P_{n_1}(x/\pi_1)\cdot P_{n_2}(x/\pi_2)$ не пренебрежимо мало, а A – объём области Q.

ЛИТЕРАТУРА

1. Андерсон Т. Введение в многомерный статистический анализ. – М.: Физматгиз, 1963. – 500 с.
2. Айвазян С.А., Бежаева З.И., Староверов О.В. Классификация многомерных наблюдений. – М.: Статистика, 1974. – 240 с.
3. Абусев Р.А., Лумельский Я.П. Несмещённые оценки и задачи классификации многомерных нормальных совокупностей. – Теория вероятностей и её применения, 1980, № 2, с. 381 – 389.
4. Абусев Р.А., О вероятности ошибок при групповой классификации. – В кн.: Применение статистических методов в производстве и управлении. Тезисы докладов Всесоюзной научно-технической конференции.: Пермь, 1984, с. 71 – 73.
5. Абусев Р.А. О сравнении поточечной и групповой классификации в случае многомерного нормального распределения. – В кн.: Статистические методы, Пермь, 1982, с. 3 – 7.
6. Фукунага К. Введение в статистическую теорию распознавания образов. – М.: Наука, 1979. – 367 с.
7. Parzen E. On estimation of a probability density function and mode. – Ann. Math. Statist., 1965, v. 36, N 4, p. 1174 – 1184.
8. Епанечников В.А. Непараметрическая оценка многомерной плотности вероятности. – Теория вероятностей и её применения, 1969, т. 14, вып. 1, с. 156 – 161.
9. Абусев Р.А. О непараметрических оценках в групповой классификации. – В кн.: Статистические проблемы управления, Вильнюс, 1978, вып. 27, с. 91 – 96.
10. Абусев Р.А. К вопросу о непараметрических оценках вероятности ошибок классификации. – В кн.: Адаптивные системы и их приложения. Новосибирск: Наука, 1978, с. 103 – 108.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

II. Айзерман М.А., Браверман Э.М., Розоноэр Л.И. Методы потенциальных функций в теории обучения машин.- М.: Наука, 1970.- 384с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

РЕКУРРЕНТНЫЕ ОЦЕНКИ ЛИНЕЙНОЙ РЕГРЕССИИ ПРИ НАРУШЕНИИ КЛАССИЧЕСКИХ ПРЕДПОСЫЛОК

А.Л.Агафонов, Е.Л.Даниленко

Рассмотрим линейную регрессионную модель $y=F(x)\theta+\varepsilon$ при классических предпосылках [1]: А) $\theta\in R^p$; В) вектор ε – случайный; С) $M\varepsilon=0, cov(\varepsilon)=\sigma^2 I_n$; Д) $F(x)$ – детерминированная матрица; Е) $rang\,F(x)=p\leq N$. Наиболее сложной вычислительной задачей при нахождении оценки $\hat{\theta}$ и планировании эксперимента является вычисление дисперсионной матрицы $D=(F^TF)^{-1}$ и основных функционалов (определителя, следа) от нее. Изменение числа наблюдений и числа оцениваемых параметров (например, при последовательном планировании эксперимента или последовательном оценивании) ведет к изменению матрицы F. Если изменяется число наблюдений, то меняется число строк, если изменяется число параметров, то меняется число столбцов.

Пусть F_i – матрица, полученная из матрицы F добавлением $(i=1)$ или изъятием $(i=3)$ строки снизу, добавлением $(i=2)$ или изъятием $(i=4)$ столбца справа, так что F_i – матрица полного ранга. Основная задача состоит в получении экономичных с вычислительной точки зрения формул вида

$$D_i=\varphi_1(D, f_{kj}^{(i)}),\ det\,D_i=\varphi_2(det\,D, f_{kj}^{(i)}),\ Sp\,D_i=\varphi_3(Sp\,D, f_{kj}^{(i)}),$$

где $f_{kj}^{(i)}$ – элементы матрицы F_i, не входящие $(i=1,2)$ в матрицу F (или элементы матрицы F, не входящие $(i=3,4)$ в матрицу F_i). Для решения этой задачи достаточно получить рекуррентные формулы для частных случаев:

$$F_1=\left\|\frac{F}{f^T}\right\|;\quad F_2=\|F\,|\,g\|;\quad F=\left\|\frac{F_3}{f^T}\right\|;\quad F=\|F_4\,|\,g\|,$$

где f и g – некоторые вектор-столбцы соответствующих размеров.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Для матриц D_1 и D_2 такие формулы получены в статьях [2, 3]. Здесь получены рекуррентные формулы для матриц D_3, D_4 и в случае нарушения предпосылок А –Е регрессионного анализа.

Введем обозначения: $d = f^T D f \neq 0$, I_N – единичная матрица $N \times N$, $a = g^T (I_N - F_4 D_4 F_4^T) g \neq 0$, $C = I_p - (1-d_1)^{-1} d_p h^T$,

$K = I_p - (1+d_2)^{-1} S C_p^T$, C, K – неособенные матрицы,

d_p – p-й столбец матрицы $D, h^T = g^T F - \|0:1\|$,

$S^T = \left\| \sum_{k=1}^{p-1} d_{k1} h_k \dots \sum_{k=1}^{p-1} d_{kp} h_k \right\|$, C_p^T – p –я строка матрицы

$C, d_1 = h^T d_p$, $d_2 = C_p^T S$, $D = \|d_{ij}\|$, h_k – элементы вектора

h, $(D F_4^T g)^T = \|\beta_1 \dots \beta_p\|$, $(Df)^T = \|\alpha_1 \dots \alpha_p\|$.

Т е о р е м а I.

$$D_3 = D + (1-d)^{-1} Df (Df)^T; \quad \det D_3 = (1-d)^{-1} \det D, \; d \neq 1;$$

$$DKC = \left\| \begin{matrix} D_4 & 0 \\ 0 & 1 \end{matrix} \right\|, \quad \det D \det K \det C;$$

$$S_p D_3 = S_p D + (1-d)^{-1} \sum_{k=1}^{p} \alpha_k^2, \; S_p D_4 = S_p D - a^{-1} \left(1 + \sum_{k=1}^{p-1} \beta_k^2\right).$$

Д о к а з а т е л ь с т в о. Простыми преобразованиями получаем первую формулу

$$D_3 = (F_3^T F)^{-1} = D(I_p - f f^T D)^{-1} = D\left(I_p + f\left(1 - f^T D f\right)^{-1} f^T D\right) =$$

$$= D + (1-d)^{-1} Df (Df)^T, \quad \text{если} \quad d \neq 1.$$

Вторая формула верна в силу известного равенства

$$\det \left\| \begin{matrix} D & \vdots & -(d-1)^{1/2} Df \\ \cdots & \cdots & \cdots \\ -(d-1)^{1/2} (Df)^T & \vdots & 1 \end{matrix} \right\| = \det D (1-d)^{-1} = \det \left(D - (d-1)^{-1} Df (Df)^T\right)$$

Пусть $H = \left\| \begin{matrix} 0 & \vdots & F_4 g \\ \cdots & \cdots & \cdots \\ g^T F_4^T & \vdots & g^T g - 1 \end{matrix} \right\|$, h^T – p –я строка матрицы. Тогда третья формула следует из соотношений

$$\left\| \begin{matrix} D_4 & 0 \\ 0 & 1 \end{matrix} \right\| = \left\| \begin{matrix} (F_4^T F_4)^{-1} & 0 \\ 0 & 1 \end{matrix} \right\| = \left(F^T F - H\right)^{-1} = D(I_p - HD)^{-1} =$$

$$= D\left(\left(I_p + h d_p^T\right) + \left\| \begin{matrix} 0 \\ S^T \end{matrix} - \right\|\right)^{-1} = D\left(I_p + C_p S^T\right)^{-1} C = DKC, \text{ если } K \text{ и } C \text{ – неособенные}$$

матрицы.

Четвертая формула следует из третьей.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Поскольку функционал $Sp(\cdot)$ - линейный, то из соотношения для D_3 следует пятая формула. Заключительное соотношение получаем из соотношения для SpD_2 [2,3], если осуществить замену матрицы D_2 на матрицу D и матрицы D на матрицу D_4.

I°. Если на оцениваемые параметры наложены ограничения, то предположение А не выполняется. Допустим, что выполняется предположение $A_1: R\theta = r$, где R - известная матрица размера $k \times p$, $k \leq p$, $rang\, R = k$, $r \in R^k$, а остальные предпосылки выполнены.

Несмещенной и эффективной линейной оценкой вектора неизвестных параметров θ в этом случае является оценка [I]:

$$\hat{\theta}_R = \hat{\theta} + DR^T W(r - R\hat{\theta}), \quad \text{где } W = (RDR^T)^{-1},$$

дисперсионная матрица которой $D_R = D(I_p - R^T WRD)$.

При изменении числа наблюдений или числа состояний дисперсионная матрица и функционалы от нее принимают вид

$$D_{Ri} = D_i (I - R_i^T W_i R_i D_i);$$

$$\det D_{Ri} = \det D_i \det (I - R_i^T W R_i D_i);$$

$$Sp\, D_{Ri} = Sp\, D_i - Sp\, D_i R_i^T W_i R_i D_i$$

Рекуррентное оценивание этих соотношений заключается в использовании результатов работ [2,3], теоремы I и следующих замечаний.

З а м е ч а н и е I. Для $i = 1,3$ матрица $R_i = R$, поскольку изменение числа наблюдений не влияет на априорные ограничения, однако матрица $W_i = W + (1 \pm d)^{-1}(1 - d_i)^{-1} WRDf(WRDf)^T$,

$$d = f^T Df, \quad d_i = (1 \pm d)^{-1}(RDf)^T RDf.$$

З а м е ч а н и е 2. При изменении размерности вектора состояний меняются априорные ограничения: при $i = 2$ матрица $R_2 = \|R_i : \rho\|$ и справедливо равенство $R_2\theta = r$, при $i = 4$ матрица $R = \|R_4 : \rho\|$ и справедливо равенство $R_4\theta = r$, ρ - вектор-столбец коэффициентов, соответствующих добавлению $(i = 2)$ или изъятию $(i = 4)$ параметра в модели сложной системы, $rang\, R_i = k$, $i = 2,4$.

Для $i = 2$ матрица $W_2 = (R_2 D_2 R_2^T)^{-1}$, а выражение для матрицы D_2 дает теорема I.

2°. Известно, что случай, когда $M\varepsilon \neq 0$, сводится к случаю

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$M\mathcal{E}=0$ с помощью введения параметра θ_{p+1}. При этом вид оценки $\hat{\theta}$ не меняется. Но иначе обстоит дело, если нарушается предположение C для ковариационной матрицы: $cov\,\mathcal{E}=\sigma^2 I_N$. Такое нарушение приводит к изменению вида оценки $\hat{\theta}$.

Выдвинем предположение $C_1: cov\,\mathcal{E}=\Omega$, где матрица Ω предполагается известной положительно определенной матрицей. Остальные предпосылки пусть выполнены. Тогда метод наименьших квадратов дает взвешенную, несмещенную и эффективную оценку в классе несмещенных оценок $\hat{\theta}=(F^T\Omega^{-1}F)^{-1}F^T\Omega^{-1}y$, а дисперсионная матрица этой оценки $D_\Omega=(F^T\Omega^{-1}F)^{-1}$.

При изменении числа наблюдений (например, увеличении) матрица $cov(\mathcal{E})=\Omega$, а следовательно и корреляционная матрица, будет иметь размерность, большую чем исходная матрица. В случае $i=1$ матрица Ω имеет вид $\Omega_1=\left\|\begin{matrix}\Omega & \ell\\ \ell^T & \omega\end{matrix}\right\|$, здесь $\ell^T=\|cov(\varepsilon_1,\varepsilon_{N+1})\dots$ $\dots cov(\varepsilon_N,\varepsilon_{N+1})\|$, $\omega=cov(\varepsilon_{N+1},\varepsilon_{N+1})$. В случае $i=3$ ковариационная матрица Ω_3 отличается от матрицы Ω тем, что не содержит N -ю строчку и N -й столбец матрицы Ω. В случаях $i=2$, $i=4$ матрица Ω не меняется. Предположим, что существуют дисперсионные матрицы $D_{\Omega i}=(F_i^T\Omega_i^{-1}F_i)^{-1}$, $i=\overline{1,4}$.

Рекуррентные формулы в рамках указанных предположений получены в теоремах 2-5.

Введем обозначения $h^T=\|-(\Omega^{-1}\ell)^T\,|\,1\|$, $E^T=(D_\Omega F_1^T h)^T=\|\beta_1\dots\beta_p\|$,

$$a=\omega-\ell^T\Omega^{-1}\ell,\quad B=1-a^{-1}h^TF_1D_\Omega F_1^Th.$$

Т е о р е м а 2.

$$D_{\Omega 1}=D_\Omega-(aB)^{-1}EE^T,\ \det D_{\Omega 1}=\det D(1-(aB)^{-1}h^TF_1D_\Omega F_1^Th),$$

$$Sp\,D_{\Omega 1}=Sp\,D_\Omega=(aB)^{-1}\sum_{k=1}^{p}\beta_k^2,\quad aB\neq 0.$$

Д о к а з а т е л ь с т в о. Первая формула следует из соотношений

$$D_{\Omega 1}=(\|F^T\,|\,f\|\Omega_1^{-1}\left\|\begin{matrix}F\\ f^T\end{matrix}\right\|)^{-1}=(F^T\Omega^{-1}F+a^{-1}F_1^Thh^TF_1)^{-1}=$$

$$=D_\Omega(Ip+a^{-1}F_1^ThH^TF_1D_\Omega)^{-1},$$

в которых учтено известное равенство для блочных матриц

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\Omega^{-1}=\left\|\begin{matrix}\Omega & \ell\\ \ell^T & \omega\end{matrix}\right\|^{-1}=\left\|\begin{matrix}\Omega^{-1} & 0\\ 0 & 0\end{matrix}\right\|+a^{-1}hh^T,\ a\neq 0\ .$$

В силу известного соотношения для определителя блочной матрицы

$$\det\left\|\begin{matrix}D_\Omega & (ab)^{-1/2}D_\Omega F^T h\\ (ab)^{-1/2}(D_\Omega F^T h)^T & 1\end{matrix}\right\|=1\cdot\det(D_\Omega-(ab)^{-1}EE^T)=$$

$$=\det D_\Omega\,(1-(ab)^{-1}h^TF_1D_\Omega F_1^Th),\ ab\neq 0,$$

справедлива вторая формула.

С учетом линейности $Sp(\cdot)$ из первой формулы следует третья.

З а м е ч а н и е. Если $\Omega=\left\|\begin{matrix}\Omega_3 & \ell\\ \ell^T & \omega\end{matrix}\right\|$, тогда справедлива формула

$$\left\|\begin{matrix}\Omega_3^{-1} & 0\\ 0 & 1\end{matrix}\right\|=\Omega^{-1}\left(I_N-\frac{CS^T}{1+S^TC}\right)\left(I_N+\frac{\ell\omega_N^T}{1-\omega_N^T\ell}\right),$$

если матрицы, входящие в произведение, не вырождены, здесь ω_N - N -й столбец матрицы Ω^{-1}, C - N -й столбец матрицы $I_N+(1-\omega_N^T\ell)^{-1}\ell\omega_N^T$,

$$S^T=\left\|\sum_{K=1}^{N-1}\ell_K\omega_{KN}\cdots\sum_{K=1}^{N-1}\ell_K\omega_{KN}\right\|,\ \ell_K-$$

элемент вектора ℓ, $\|\omega_{Ki}\|=\Omega^{-1}$. Доказательство этого соотношения проводится аналогично доказательству подобного соотношения в теореме 2.

Введем обозначения: $h^T=\|-(\Omega_3^{-1}\ell);1\|$; $a=\omega-\ell^T\Omega_3^{-1}\ell$, $b=1-a^{-1}h^TFD_\Omega F^Th$; $E_3^T=(DF^Th)^T=\|\beta_1\cdots\beta_p\|$. Аналогично доказательству теоремы 2 доказывается

Т е о р е м а 3.

$$D_{\Omega 3}=D_\Omega+(ab)^{-1}E_3E_3^T,\ \det D_{\Omega 3}=\det D_\Omega\,(1+(ab)^{-1}h^TFD_\Omega F^Th)^{-1},$$

$$Sp\,D_{\Omega 3}=Sp\,D_\Omega+(ab)^{-1}\sum_{K=1}^{p}\beta_K^2.$$

Несколько изменим обозначения: $a=g^T(\Omega^{-1}-\Omega^{-1}FD_\Omega F^T\Omega^{-1})g$,

$h^T=\|-(D_\Omega F^T\Omega^{-1}g)^T;1\|$, $\|\alpha_1\ldots\alpha_p\|=(D_\Omega F^T\Omega^{-1}g)^T$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Т е о р е м а 4.

$$D_{\Omega 2}=\begin{Vmatrix} D_\Omega & 0 \\ 0 & 0 \end{Vmatrix}+a^{-1}hh^T,\ a\neq 0,\ \det D_{\Omega 2}=a^{-1}\det D_\Omega,$$

$$S_p D_{\Omega 2}=S_p D_\Omega+a^{-1}\Big(1+\sum_{k=1}^{p}\alpha_k^2\Big).$$

Д о к а з а т е л ь с т в о. Первая формула получается после применения соответствующего результата из работ [2,3] к блочной матрице

$$D_{\Omega 2}=\begin{Vmatrix} F^T\Omega^{-1}F & F^T\Omega^{-1}g \\ g^T\Omega^{-1}F & g^T\Omega^{-1}g \end{Vmatrix}.$$

При доказательстве второй формулы используем стандартный прием, изложенный в теоремах I и 2. Третье соотношение следует из первой формулы и линейности функционала $S_p(\cdot)$.

Пусть d_p - p -й столбец матрицы D_Ω, $g^T\Omega^{-1}F_4=\|0:1\|=h^T=\|h_k\|$, $C=I_p+(1-h^Td_p)^{-1}d_ph^T$, C_p - p -й столбец матрицы C,

$$D_\Omega=\|d_{ik}\|,\ S^T=\Big\|\sum_{k=1}^{p-1}d_{1k}h_k\ldots\sum_{k=1}^{p-1}d_{pk}h_k\Big\|,\ K=I_p+(1+C_p^TS)^{-1}\times$$

$$\times SC_p^T,\ a=g^T(\Omega^{-1}F_4D_{\Omega 4}F_4^T\Omega^{-1})g,\ \|\alpha_1\ldots\alpha_{p-1}\|=(D_{\Omega 4}F_4^Tg)^T.$$

Т е о р е м а 5.

$$CKD_\Omega=\begin{Vmatrix} D_{\Omega 4} & 0 \\ 0 & 1 \end{Vmatrix},\ \det D_{\Omega 4}=a^{-1}\det D_\Omega,\ a\neq 0,$$

$$S_pD_{\Omega 4}=S_pD_\Omega-a^{-1}\Big(1+\sum_{k=1}^{p-1}\alpha_k^2\Big).$$

Доказательство данной теоремы аналогично доказательству подобных соотношений в теореме I.

3°. Нарушение предпосылки Е означает линейную зависимость столбцов матрицы $F(x)$. Однако на практике ситуация, когда $rang\,F<p$ не встречается, поскольку независимые переменные являются результатом измерений и поэтому содержат ошибки. Это ведет к тому, что даже если истинные значения векторов-столбцов матрицы F линейно зависимы, то информационная матрица F^TF невырождена, хотя и плохо обусловлена.

В случае плохой обусловленности информационной матрицы оценки метода наименьших квадратов становятся малоэффективными, то есть

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

дисперсия оценок будет большой.

Отказавшись от несмещенности, можно построить эффективные оценки с меньшим квадратом ошибки. К таким оценкам относятся ридж-оценки. В классе оценок с фиксированной длиной ридж-оценка $\hat{\theta}(k)=(F^TF+kI)^{-1}Fy$ минимизирует сумму квадратов отклонений $e=y-F\hat{\theta}(k)$, где $k\geq 0$ - параметр ридж-оценки, способы выбора которого описаны, например в [I].

Дисперсионные матрицы, соответствующие ридж-оценкам, имеют вид $D_{RI}=\sigma^2(I+kD)^{-2}D$, где $D=(F^TF)^{-1}$. Здесь σ^2 - известная величина.

Обозначим $R=(I+kD)^{-1}$, тогда $D_{RI}=\sigma^2R^2D$ и рекуррентные формулы для матрицы D_{RI} следуют из рекуррентных формул для матриц R и D.

Приведем рекуррентные соотношения для матрицы R. Предполагаем, что матрицы $R_i=(I+kD_i)^{-1}$, $i=\overline{1,4}$, существуют.

Введем новые обозначения:

$$C_i=1-k(Df)^TRDf(1\pm d)^{-1},\ \|\beta_1\dots\beta_p\|=(RDf)^T=E_R^T,$$

знак (+) соответствует $i=1$ знак (-) соответствует $i=3$.

Т е о р е м а 6. Для $i=1,3$

$$R_i=R+k(1\pm d)^{-1}C_iE_RE_R^T,\ \det R_i=\det R(1-k(1\pm d)^{-1}\times$$

$$\times C_iE^T(I+kD)E),\ S_pR_i=S_pR+k(1\pm d)^{-1}C_i\sum_{k=1}^{p}\beta_k^2.$$

Д о к а з а т е л ь с т в о. Известные выражения для матриц подставляем в соотношение $R_i=(I+kD_i)^{-1}$, откуда после стандартных преобразований, использованных в теореме I, получаем первую группу формул. Доказательства второй и третьей групп формул аналогичны доказательству подобных формул из теоремы I.

Пусть $F^+=DF^T$, $h^T=\|-(F^+g)^T;1\|$, $E_{R2}^T=(\left\|\begin{matrix}R & 0\\ 0 & 1\end{matrix}\right\|h)^T=$

$$=\|\alpha_1\dots\alpha_{p+1}\|,\ b=g^Tg+k-g^TFDRF^Tg,\ a=g^Tg-g^TFDF^Tg,$$

$$r=1+kah^TE_{R2}.$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Теорема 7.

$$R_2 = \left\| \begin{matrix} R & 0 \\ 0 & 1 \end{matrix} \right\| - h(ar)^{-1} E_{R2} E_{R2}^T, \quad \det R_2 = \det R (ab)^{-1},$$

$$S_p R_2 = 1 + S_p R - h(ar)^{-1} \sum_{k=1}^{p+1} \alpha_k^2, \quad abr \neq 0.$$

Введем обозначения: r^T – p –я строка матрицы R; $h^T = \|g^T F_1 g^T g + h - 1\|$, $C = I_p + (1 - r^T h)^{-1} h r^T$; C_p – p –й столбец матрицы C, $S^T = \left\| \sum_{k=1}^{p-1} h_k r_{k1} \dots \sum_{k=1}^{p-1} h_k r_{kp} \right\|$, $m = I_p - (1 + S^T C_p) C_p S^T$,

$$a = g^T (1 - F_4 D_4 F_4^T) g, \quad b = g^T g + h - g^T F_4 R_4 F_4^+ g, \quad \left(\left\| \begin{matrix} R_4 & 0 \\ 0 & 1 \end{matrix} \right\| h \right)^T = \|\alpha_1 \dots \alpha_p\|.$$

Теорема 8.

$$R C m = \left\| \begin{matrix} D_4 R_4 & 0 \\ 0 & 1 \end{matrix} \right\|, \quad \det R_4 = (Ba)^{-1} \det R,$$

$$S_p R_4 = S_p R + h(ar)^{-1} \sum_{k=1}^{p} \alpha_k^2 - 1, \quad abr \neq 0.$$

Литература

I. Демиденко Е.З. Линейная и нелинейная регрессии.-М.:Финансы и статистика,I98I.-300с.

2. Даниленко Е.Л., Романовский Р.К. Рекуррентные формулы для критериев оптимальности.-Заводская лаборатория,I98I,т.47, № 3, с.57-58.

3. Даниленко Е.Л. Планирование последовательного контроля состояний сложной системы.-Планирование эксперимента. М.,I985,с.I3I-I34.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

1-660936

УСТОЙЧИВАЯ ФИЛЬТРАЦИЯ В ДИНАМИЧЕСКИХ СИСТЕМАХ СЛУЧАЙНОЙ СТРУКТУРЫ

С.Е.Ададуров , А.В.Макшанов

Введение

В настоящее время всё большее применение в системах поиска и обнаружения радиотехнических сигналов находят обнаружители оценочно-корреляционного типа[I] ,позволяющие вести квазикогерентную обработку. Однако использование обычных алгоритмов оценочно-корреляционного обнаружения(ОКО)в радиотехнических системах (РТС)для ситуаций,характеризующихся сложной помеховой обстановкой,часто бывает невозможно.К таким ситуациям следует отнести случаи,когда полезный сигнал пропадает вследствие воздействия излучения мешающих РТС со структурой сигнала,совпадающей по спектру с полезным сигналом,и одновременным увеличением уровня шумов в приемном тракте.

В данной ситуации время появления сигнала становится случайным.Увеличение мощности шума в радиоканале ведёт при неравномерной огибающей спектра шума к денормализации закона распределения ошибок параметров частоты и задержки сигнала.Кроме того, наличие в радиоканале мешающего сигнала,аналогичного по структуре полезному сигналу,приводит к значительному количеству аномальных измерений.Заметим,что аномальные измерения,как правило, обусловлены выбросами взаимокорреляционных функций полезного и мешающих сигналов при условии приблизительного равенства их мощностей.

Таким образом,большой практический интерес имеет задача совершенствования алгоритмов ОКО для РТС,работающих в сложной помеховой обстановке.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Синтез устойчивых алгоритмов фильтрации в ОКО

Одним из возможных путей построения ОКО, работающих в сложной помеховой обстановке, является построение сглаживающих цепей обнаружителя-оценивателя на основе общей теории динамических систем случайной структуры (ДССС) [2].

Пусть состояние системы описывается дифференциальным уравнением вида

$$\left.\begin{aligned} \dot{X}^{(1)} &= FX \ ; \\ \dot{X}^{(2)} &= 0 \ , X^{(2)} = 0, \end{aligned}\right\} \qquad (1)$$

где X -вектор, характеризующий параметр сигнала,

F -матрица динамики.

Уравнение канала наблюдения даётся в виде

$$Z = HX + \xi, \qquad (2)$$

где H -матрица наблюдения;

ξ -белый шум с корреляционной матрицей $Q = \frac{1}{2} N \delta(t-\tau)$.

Тогда уравнения совместного обнаружения и оценивания при условии

$$\left.\begin{aligned} H &= [1\,0\,0 \ldots 0]; \\ P_1(t_0) &= 0 \ ; \\ P_2(t_0) &= 1 ; \end{aligned}\right\} \qquad (3)$$

$$\nu_{12}(X^{(1)}, t) = 0 \ ; \quad \nu_{21}(X^{(2)}, t) = \begin{cases} 0, & |X_1^{(2)}| < \Delta ; \\ 1, & |X_1^{(2)}| \geq \Delta , \end{cases} \qquad (4)$$

где P_1, P_2 -вероятности состояний;

ν_{12}, ν_{21} -интенсивности переходов;

Δ -постоянная, даются выражениями:

$$y(t) = \int S_{вх}(X, t)\, S_{оп}(X, t-\tau)\, d\tau \ ; \qquad (5)$$

$$\dot{\hat{X}}^{(1)} = F\hat{X}^{(1)} + R^{(1)} H^T Q^{-1} (Z - H\hat{X}^{(1)}) + P_2 P_1^{-1} \cdot 2\,(2\pi)^{-1/2} \cdot (2R_{11}^{(1)})^{1/2} \times$$
$$\times \left(\exp\left(-(\Delta - \hat{X}_1^{(2)})^2 (2R_{11}^{(1)})^{-1}\right) + \exp\left(-(-\Delta - \hat{X}_1^{(2)})^2 \cdot (2R_{11}^{(1)})^{-1}\right)\right) \qquad (6)$$
$$\times P_2 P_1^{-1} \left(\Phi\left(\frac{-\Delta - \hat{X}_1^{(2)}}{2R_{11}^{(1)}}\right) + 1 - \Phi\left(\frac{\Delta - \hat{X}_1^{(2)}}{2R_{11}^{(1)}}\right)\right)\left(\hat{X}^{(2)} - \hat{X}^{(1)}\right) ;$$

$$\dot{\hat{X}}^{(2)} = R^{(2)} H^T Q^{-1} (Z - H\hat{X}^{(2)}) - 2\,(2\pi)^{-\frac{1}{2}} \sqrt{R_{11}^{(2)}} \left(\exp\left(-(-\Delta - \right.\right.$$
$$\left.\left. - \hat{X}_1^{(2)})^2 (2R_{11}^{(2)})^{-1}\right) + \exp\left(-(\Delta - \hat{X}_1^{(2)})^2 (2R_{11}^{(2)})^{-1}\right)\right) ; \qquad (7)$$

$$\dot{R}^{(1)} = FR^{(1)} + R^{(1)} F^T - R^{(1)} H^T Q^{-1} H R^{(1)} + P_2 P_1^{-1} \left((R^{(2)} - \right.$$
$$\left. - R^{(1)}) + (\hat{X}^{(2)} - \hat{X}^{(1)})(\hat{X}^{(2)} - \hat{X}^{(1)})^T\right) ; \qquad (8)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\dot{R}^{(2)} = R^{(2)} H^T Q^{-1} H R^{(2)} + (R^{(1)} - R^{(2)}) + (\hat{X}^{(1)} - \hat{X}^{(2)})(\hat{X}^{(1)} - \hat{X}^{(2)})^T; \quad (9)$$

$$\dot{\hat{P}}_2 = -\hat{P}_2 \left(\Phi\left(\frac{-\Delta - \hat{X}_1^{(2)}}{R_{11}^{(1)}} \right) + \Phi\left(\frac{\Delta - \hat{X}_1^{(2)}}{R_{11}^{(1)}} \right) \right); \quad (10)$$

$$\hat{P}_1 = 1 - \hat{P}_2; \quad (11)$$

$$\hat{X}^{(1)}(t_0) = X_0^{(1)}, \; X^{(2)}(t_0) = 0. \quad (12)$$

где

$S_{вх}(\cdot)$ -входной сигнал;

$S_{оп}(\cdot)$ -опорный сигнал; знак $\wedge$ обозначает оценку параметра.

Однако моделирование алгоритмов ОКО показало слабую устойчивость (6)- (12) при воздействии помех, вызывающих асимметрию распределения погрешностей измерений. Кроме того, наличие достаточного количества аномальных измерений вследствие помех ведёт к выходу обнаружителя из заданной ячейки просмотра области неопределённости по параметрам сигнала.

Проведённые экспериментальные исследования показали, что одним из возможных выходов из указанной ситуации является робастификация процедур (6) - (12). С этой целью слагаемые в (6) и (7)

$$R^{(1)} H^T Q^{-1} (Z - H\hat{X}^{(1)}),$$
$$R^{(2)} H^T Q^{-1} (Z - H\hat{X}^{(2)})$$

были преобразованы в

$$R^{(1)} \psi(H^T Q^{-1} (Z - H\hat{X}^{(1)})), \quad (13)$$
$$R^{(2)} \psi(H^T Q^{-1} (Z - H\hat{X}^{(2)})), \quad (14)$$

где $\psi(\cdot)$ -функция контраста. Для моделирования процесса обнаружения использовалась модель погрешностей измерений типа

$$F(x) = (1-\varepsilon) N(0, G^2) + \varepsilon N(\mu, G^2) + \gamma N(\mu, 2G^2), \quad (15)$$

где $\varepsilon, \gamma \in [0; 1]$,

$N(\mu', G^{2'})$ -нормальное распределение с параметрами μ' и $G^{2'}$.

Результаты моделирования показали, что при величинах $\varepsilon = 0{,}15$ и выше, $\gamma = 0{,}08$ и вышесглаживающие цепи (6) - (12) оказываются неработоспособными, если $\mu \geq 0{,}6\sqrt{R_{11}^{(1)}}$. В том случае, когда в выражениях (6) и (7) используются слагаемые (13) и (14), сглаживающие цепи ОКО остаются работоспособными при значительных вариациях параметров γ, ε и μ. При моделировании исследовались функции ψ типа Хьюбера, Хэмпела и бивес Тьюки

$$\psi_1(\theta) = \begin{cases} -K, & \theta < -K; \\ \theta, & |\theta| \leq K, \; K > 0; \\ K, & \theta > K; \end{cases} \quad (16)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\psi_2(\theta)=\begin{cases}\theta, & 0\le\theta\le a, \quad a=\sqrt{R_{11}^{(\theta)}};\\ a, & a<\theta\le b, \quad b=3\sqrt{R_{11}^{(\theta)}};\\ \varepsilon\theta, & b<\theta\le c, c=4\sqrt{R_{11}^{(\theta)}};\\ 0, & \theta>c\end{cases} \qquad (17)$$

$$\psi_3(\theta)=\begin{cases}\theta(1-\theta^2)^2, & |\theta|\le 1;\\ 0, & |\theta|>1.\end{cases} \qquad (18)$$

Наилучший результат фильтрации при модели погрешностей (15) получается для функции контраста (17). При этом алгоритмы (6), (7) - (12) с модификацией (13) и (14) оказываются работоспособными, даже если $\varepsilon=0{,}4$; $\gamma=0{,}15$; $\mu=0{,}96$.

Выводы

1. Установлено, что функционирование ОКО в сложной помеховой обстановке, характеризующейся высоким уровнем неравномерных по спектру шумов и мешающего сигнала, оказывается малоэффективным из-за статистической неустойчивости алгоритмов фильтрации.

2. Алгоритмы фильтрации в ДССС в ОКО, работающих в сложных помеховых условиях и при пропадании сигнала, оказываются неработоспособными вообще.

3. Применение робастных процедур фильтрации позволяет значительно расширить область работоспособности алгоритмов ОКО.

4. Наибольший практический интерес представляют алгоритмы фильтрации с использованием функции контраста Хэмпела, которые позволяют повысить значительно статистическую устойчивость систем фильтрации в ДССС.

Литература

1. Сосулин Ю.Г. Теория обнаружения и оценивания стохастических сигналов. - М.: Сов. радио, 1978. - 320с.

2. Казаков И.Е., Артемьев В.М. Оптимизация динамических систем случайной структуры. - М.: Наука, 1980. - 384с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ОБ ОДНОМ КРИТЕРИИ ВЫБОРА ОПТИМАЛЬНЫХ ПРИЗНАКОВ, ОСНОВАННОМ НА ПОНЯТИИ ДИФФЕРЕНЦИАЛЬНОЙ ЭНТРОПИИ

В. К. Анциферов, И. А. Пестунов

Задача выбора информативных признаков является одной из наиболее важных в проблеме распознавания образов [1] . Суть этой задачи заключается в следующем: необходимо найти такое преобразование F (линейное или нелинейное), которое отображает вектор исходных признаков x размерности m в вектор признаков y размерности k $(k < m)$, минимизируя или максимизируя при этом некоторый критерий, характеризующий эффективность вектора признаков y . Применительно к задаче распознавания в статистической постановке наилучшим критерием эффективности признаков с теоретической точки зрения является байесова вероятность ошибки классификации (байесова ВОК). В практических задачах, как правило, из-за недостатка априорных сведений о статистических характеристиках классов вычисление байесовой ВОК осуществляется с помощью непараметрических процедур, что связано со значительными вычислительными затратами [2 - 4] .

В данной работе предлагается основанный на понятии дифференциальной энтропии критерий информативности признаков, который, как показали результаты численного моделирования, непосредственно связан с байесовой ВОК, а его численная реализация в условиях непараметрической априорной неопределенности менее трудоемка.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Дифференциальная энтропия и ее непараметрическая оценка

Для непрерывной K -мерной случайной величины ξ , имеющей плотность распределения $p_\xi(x)$, $x \in \Omega \subset R^K$, функционал дифференциальной энтропии определяется формулой

$$h(\xi) = -\int_\Omega p_\xi(x) \ln p_\xi(x) dx . \qquad (1)$$

Пусть $X = \{x_1, x_2, \dots, x_N\}$ – выборка объема N из генеральной совокупности с плотностью распределения $p_\xi(x) \in L_2$.

В качестве прямой [5] непараметрической оценки функционала (1) примем статистику

$$\hat{h}_N = -\int_\Omega \hat{p}_N(x) \ln \hat{p}_N(x) dx , \qquad (2)$$

где $\hat{p}_N(x)$ – непараметрическая оценка плотности распределения $p_\xi(x)$, равная

$$\hat{p}_N(x) = (N \cdot C^K(N))^{-1} \cdot \sum_{i=1}^{N} \Phi\left(\frac{x - x_i}{C(N)}\right);$$

здесь интегрируемая с квадратом функция $\Phi(\cdot)$ и параметр сглаживания $C(N)$ удовлетворяют определенным условиям сходимости [6] . Тогда справедливо следующее утверждение.

ТЕОРЕМА. Пусть $p_\xi(x)$ положительна, ограничена и непрерывна на области Ω . Тогда, если интегрируемая с квадратом функция $\Phi(\cdot)$ и параметр $C(N)$ удовлетворяют вышеупомянутым условиям сходимости, то

$$\lim_{N \to \infty} M\left\{ (\hat{h}_N - h)^2 \right\} \to 0 .$$

Для нахождения наилучшего $C(N)$ при оценивании (2) в качестве меры уклонения $\hat{h}_N$ от h возьмем $(\hat{h}_N - h)^2$ и будем выбирать $C(N)$ так, чтобы

$$(\hat{h}_N - h)^2 = \min_{C(N)} . \qquad (3)$$

Не трудно показать, что минимум (3) по $C(N)$ будет достигаться при том же значении $C_{опт}(N)$, что и минимум по $C(N)$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

интеграла

$$J(C(N)) = \int_{\Omega} (\hat{P}_N(x) - P_\xi(x))^2 dx. \qquad (4)$$

Таким образом, оптимизация $\hat{h}_N$ по $C(N)$ сводится к оптимизации $\hat{P}_N$ [2].

Поведение оценки $\hat{h}_N$ в зависимости от объема выборки N исследовалось методом статистического моделирования. С помощью датчика псевдослучайных чисел генерировались выборки объемов N = 25,50,75,100 одномерных реализаций из нормальных распределений $N(0,1)$, $N(0,3)$ и по каждой из этих выборок вычислялась величина $\Delta h = (|h - \hat{h}_N|/h)$. 100%. Значения h вычислялись по формуле $h = \ln(\sqrt{2\pi e}\cdot\sigma)$. Колоколообразная функция $\Phi(\cdot)$ имела вид

$$\Phi\left(\frac{x-x_i}{C(N)}\right) = \begin{cases} 0.5, & \text{если } |x - x_i| \le C(N); \\ 0, & \text{если } |x - x_i| > C(N). \end{cases}$$

Непараметрическая оценка (2) соответственно имела вид

$$\hat{h}_N = -\frac{1}{N}\sum_{i=1}^{N} \ln \sum_{j=1}^{N} \frac{1(x_i, x_j)}{2C(N)},$$

где

$$1(x_i, x_j) = \begin{cases} 1, & \text{если } |x_i - x_j| \le C(N); \\ 0, & \text{если } |x_i - x_j| > C(N). \end{cases}$$

Оптимальные коэффициенты сглаживания $C_{опт}(N)$ определялись из условия

$$(\hat{h}_N - h)^2 = \min_{C(N)}.$$

Усреднение проводилось по 5 расчетам с использованием независимых выборок. Усредненные результаты расчетов приведены в табл. I.

Таблица I

σ^2 \ N	25		50		75		100	
	Δh	$C_{опт}$	Δh	$C_{опт}$	Δh	$C_{опт}$	Δh	$C_{опт}$
I	0.249	0.6	0.119	0.45	0.114	0.3	0.041	0.3
3	0.34	1.05	0.108	0.9	0.062	0.55	0.023	0.55

Критерий информативности признаков

В случае двух классов ω_1 и ω_2 с априорными вероятностями

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

появления $p(\omega_1)$ и $p(\omega_2)$ и условными плотностями распределения $p(x/\omega_1)$ и $p(x/\omega_2)$ соответственно предлагаемый критерий можно записать следующим образом:

$$I(\omega_1,\omega_2)=\Big[\sum_{i=1}^{2} p(\omega_i)\big(h_{\omega_i}-\ln(p(\omega_i)/c)\big)- h_{\omega_1\omega_2}-\ln c\Big]/c, \qquad (5)$$

где

$$h_{\omega_i}=-\int p(x/\omega_i)\ln p(x/\omega_i)\,dx,$$

$$h_{\omega_1\omega_2}=-\int p(x)\ln p(x)\,dx,$$

$$p(x)=p(\omega_1)p(x/\omega_1)+p(\omega_2)p(x/\omega_2),$$

$$c=e\cdot\max_{x}\{p(x)\};\quad e=2.71\ldots.$$

Заметим, что в случае непересечения классов $I(\omega_1,\omega_2)=0$.

Если $p(\omega_i)$, $p(x/\omega_i)$, $i=1,2$, неизвестны, а имеются лишь выборки $X_1=\{x_1^1,\ldots,x_{N1}^1\}$ объема $N1$ и $X_2=\{x_1^2,\ldots,x_{N2}^2\}$ объема $N2$ ($N=N1+N2$) наблюдений из классов ω_1, ω_2 соответственно, то, заменяя в (5) $p(\omega_i)$ на N_i/N, $h_{\omega_i}, h_{\omega_1\omega_2}$ их оценками $\hat{h}_{\omega_i}$, $\hat{h}_{\omega_1\omega_2}$ вида (2), а c – на $\hat{c}=e\cdot\max_{x\in X_1\cup X_2}\{\hat{p}_N(x)\}$, получим оценку критерия $I(\omega_1,\omega_2)$ (обозначим ее $\hat{I}(\omega_1,\omega_2)$) для случая непараметрической неопределенности.

Связь критерия $\hat{I}$ с байесовой (теоретической) ошибкой классификации $p_б$ исследовалась методом статистического моделирования. Генерировались 7 пар выборок ($N1=N2=100$) одномерных реализаций из нормальных распределений со следующими параметрами: $\sigma_1^2=\sigma_2^2=1$, а разность математических ожиданий в парах $|M_1-M_2|=0,1,2,3,4,5,20$.

По каждой из этих пар выборок вычислялось значение критерия $\hat{I}$. Соответствующие каждому из семи вышеуказанных случаев значения $p_б$ вычислялись по формуле $p_б=\Phi(-|M_1-M_2|/(2\sigma))$, где $\Phi(\cdot)$ – функция распределения стандартного нормального распределения.

Усреднение проводилось по 5 расчетам с использованием независимых выборок. Усредненные результаты расчетов приведены в табл.2.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Таблица 2

$\lvert M_1 - M_2 \rvert$	20	5	4	3	2	I	0
$P_б$ (%)	0	0.5	2.2	6.5	I5.7	30.7	50
$\hat{I}$	0	0.3I	0.43	0.75	I.35	2.22	3.I3

Заметим, что при оптимизации по параметру $C(N)$ оценок $\hat{h}_{\omega_i}$ и $\hat{h}_{\omega_1\omega_2}$, используемых в критерии $\hat{I}$, целесообразно вместо (4) минимизировать интеграл

$$J(C(N)) = \int_{\Omega} q(x)\left(\hat{P}_N(x) - p_{\xi}(x)\right)^2 dx,$$

где $q(x)$ – неотрицательная и ограниченная на Ω весовая функция (наибольший вес придается области пересечения классов).

Замечание относится и к оптимизации непараметрических алгоритмов распознавания образов, рассмотренных в [2].

Литература

I. Фукунага К. Введение в статистическую теорию распознавания.- М.:Наука, I979.- 367 с.
2. Лапко А.В., Медведев А.В. К анализу непараметрических алгоритмов распознавания образов.- В кн.: Статистические проблемы управления. Вильнюс: Пяргале, I976, вып. I4, с. I05-II6.
3. Абусев Р.А. К вопросу о непараметрических оценках вероятности ошибок классификации.- В кн.: Адаптивные системы и их приложения. Новосибирск: Наука, I978, с. I03-I08.
4. Галфаян С., Чилингарян А. Вычисление байесовского риска при помощи КБС оценок плотности вероятности.- В кн.: Статистические проблемы управления. Вильнюс, I984, вып. 66, с. 66-78.
5. Тарасенко Ф.П. Непараметрическая статистика.-Томск: Изд-во ТГУ, I976.- 292 с.
6. Медведев А.В. Непараметрические системы адаптации.- Новосибирск: Наука, I983. - I74 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ЭФФЕКТИВНАЯ МОДИФИКАЦИЯ АЛГОРИТМА k - БЛИЖАЙШИХ СОСЕДЕЙ

Э.К. Атаянц, В.П. Рынин

Функционирование устройства распознавания на основе решающего правила k - ближайших соседей /I/ делится на два этапа. На первом происходит обучение устройства распознавания, т.е. формирование эталонной выборки. Второй этап связан непосредственно с распознаванием.

Исходным материалом для отбора эталонов служит экспериментально полученная обучающая выборка объемом N. При решении двухальтернативной задачи распознавания она состоит из $N^{(1)}$ элементов первого и $N^{(2)}$ элементов второго классов, $N=N^{(1)}+N^{(2)}$.

Распознавание неизвестного элемента $X=\{x_1,\dots,x_d\}^T$ размерностью d с помощью решающего правила "k - ближайших соседей" проводится на основе анализа соотношения /I/:

$$k_1 \gtrless k_2 \rightarrow X \in \begin{cases} \omega_1 \\ \omega_2 \end{cases},$$

где k_1 и k_2 - количество ближайших эталонов соответственно первого и второго классов, $k=k_1+k_2$. При $k=1$ решающее правило переходит в правило "ближайшего соседа" (*nearest neighbor - NN*) /I/. Мерой близости между i-м эталоном $A_i=\{a_{i1},\dots,a_{id}\}^T$ произвольного класса и распознаваемым элементом X является величина расстояния /I/.

В процессе принятия решения вычисляется расстояние от распознаваемого элемента до каждого эталона каждого класса. При использовании правила "ближайшего соседа" решение принимается в пользу класса, содержащего эталон, наиболее близкий (минимальная величина D) к распознаваемому элементу.

Обучение устройства распознавания на основе метода *NN* состоит в том, что вся исходная экспериментально полученная обучающая выборка принимается за эталонную.

Улучшение характеристик устройства распознавания (снижение величины вероятности ошибочного распознавания) достигается увеличе-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

нием объема экспериментальной выборки, что приводит как к росту объема памяти устройства, требуемого для хранения эталонов, так и времени принятия решения.

Существующие модификации метода *NN*, основной целью которых является сокращение объема эталонной выборки, отличаются введением на этапе обучения различных алгоритмов отбора эталонов. Уменьшение числа эталонов достигается либо формированием выборки усредненных эталонов с контролем качества распознавания на каждом этапе усреднения /2/, либо включением в эталонную выборку только объектов, расположенных на границе пересечения функций плотности распределения вероятностей (ФПРВ) классов /3-5/, т.к. именно они в основном участвуют при вынесении решения и влияют на величину ошибки.

Авторами был исследован метод *NN* и ряд его модификаций, а именно (названия соотвествуют приводимым в литературе, в случае отсутствия давались авторами условно): *NN* /1/, *CNN* /3/, *RNN* /4/, *TNN1* /5/, *TNN2A* /5/ (без последующего применения *CNN*), *TNN2C* /5/ (с последующим применением *CNN*), *WNN* /6/ и *CNN* - 3 /3/ (для случая $k=3$). Сравнение различных алгоритмов обучения было проведено в ходе решения нескольких задач распознавания. Одна из задач состояла в распознавании объектов 2-х классов, двумерные ФПРВ которых описываются выражениями:

$$p_{\bar{1}}(X)=\{2\,exp[-(x_1^2/2+x_2^2/2)]+exp[-((x_1-14)^2/4+x_2^2/4)]\}/8\pi\,;$$

$$p_{\bar{2}}(X)=\{exp[-(x_1^2/8+(x_2-3,5)^2/8)]+2\,exp[-((x_1-3)^2/4+x_2^2/4)]\}/16\pi\,.$$

ФПРВ, соотвествующие выражениям для $p_{\bar{1}}(X)$ и $p_{\bar{2}}(X)$, показаны на рис. I. Вид ФПРВ и параметры распределений выбраны эвристически. Их выбор обеспечил достаточно сложное распределение объектов распознаваемых классов (неодномодальные ФПРВ, пересечение классов со сложной границей, наличие отдельно расположенных групп объектов) при приемлемой простоте генерации и позволил сравнить алгоритмы в различных условиях (например, зависимость характеристик от степени пересечения классов, объема обучающей выборки, ее искажений и т. д.). Результаты обучения представлены в таблице.

Основными недостатками перечисленных выше алгоритмов являются:

I) практически линейный и быстрый рост объема памяти (и времени принятия решения) при увеличении объема исходной обучающей выборки *N*.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Результаты обучения

Количество объектов памяти		NN	CNN	RNN	TNN1	TNN2A	TNN2C	CNN-3	WNN
	Память I-го класса	I00	I6	I4	I4	24	I6	I9	92
	Память 2-го класса	I00	I7	I4	I2	22	I3	22	68
	Всего (в %'от исх. обуч. выборки)	I00	I6,5	I4	I3	23	I4,5	20,5	90

Объясняется это тем, что граница классов описывается не точками (концами векторов) поверхности в многомерном пространстве, а совокупностью точек в объеме пространства вблизи решающей гиперповерхности;

2) случайное появление в исходной выборке элементов, соответствующих малым значениям ФПРВ , приводит к существенному росту вероятности принятия ошибочного решения.

Значимость указанных недостатков растет по мере увеличения пересечения многомерных ФПРВ классов.

Предлагаемый алгоритм обучения позволил избавиться от перечисленных недостатков, что привело к уменьшению вероятности принятия ошибочного решения и сокращению объема эталонной памяти (а значит, и времени принятия решения). В основе алгоритма лежит предварительная обработка обучающей выборки с целью исключения элементов выборки, расположенных в области пересечения ФПРВ классов, с последующим отбором элементов, находящихся вблизи указанной области.

Для оценки качества предложенного метода обработки обучающей выборки был проведен статистический эксперимент на ЭВМ. В ходе эксперимента решалась задача распознавания, упомянутая выше. Для сравнения был взят один из алгоритмов (*CNN*), отбирающий в качестве эталонов объекты, расположенные на границе пересечения ФПРВ.

На рис. 2-6 показаны характеристики, полученные в ходе статистического эксперимента. Рис. 2 показывает зависимость объема эта-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

лонной выборки $n = n^{(1)} + n^{(2)}$ (где $n^{(1)}$ и $n^{(2)}$ – количество эталонов соответственно первого и второго классов) для предложенного (I) и *CNN* (2) алгоритмов обучения от объема исходной обучающей выборки ($N = N^{(1)} + N^{(2)}$). Пунктиром показаны соответствующие 90%-ные доверительные интервалы.

На рис. 3 показаны зависимости величины вероятности принятия ошибочного решения от N, для указанных алгоритмов обработки с соответствующими 90%-ными доверительными интервалами.

Чтобы оценить, насколько существенны улучшения характеристик распознавания на основе предложенного алгоритма, было проведено сравнение его с байесовским решающим правилом. Устройство распознавания на основе байесовского решающего правила тоже было оценено с помощью статистического эксперимента. Характеристика этого устройства с соответствующим ей 90% – ным доверительным интервалом показана на рис. 3 (прямая 3).

Рис. 4 показывает зависимость относительного количества эталонов (n/N) от величины пересечения функций плотности распределения вероятности классов. В качестве величины, характеризующей степень пересечения, взята полная ошибка распознавания, даваемая байесовским решающим правилом ($\mathcal{E}_б$). Рис. 5 и 6 характеризуют устойчивость алгоритмов при искажениях обучающей выборки, когда часть объектов обучающей выборки распознаваемых классов $\mathcal{E}_и = N_и^{(1)}/N^{(1)} = N_и^{(2)}/N^{(2)}$ заменялась объектами противоположного класса ($N_и^{(1)}$ и $N_и^{(2)}$ – количество искаженных объектов соответственно первого и второго классов).

На рис. 5 приведены зависимости вероятности принятия ошибочного решения при различных искажениях обучающей выборки для предложенного (I) и *CNN*- (2) алгоритмов с учетом 90%-го доверительного интервала.

Рис.6 иллюстрирует изменение относительного объема эталонной памяти в зависимости от величины $\mathcal{E}_и$. Характеристики на рис. 2-6 во-первых, показывают, что использование предложенного алгоритма обучения приводит к существенному сокращению объема памяти, необходимого для хранения эталонов; во-вторых, предложенный алгоритм показал бо́льшую, чем *CNN*, эффективность распознавания (в среднем величина вероятности ошибки распознавания меньше) при одинаковом объеме обучающего материала и, в-третьих, характеристики устройств распознавания на основе предложенного алгоритма отличаются большей устойчивостью к искажениям обучающей выборки.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

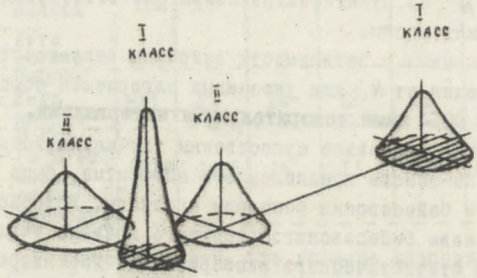

Рис. 1

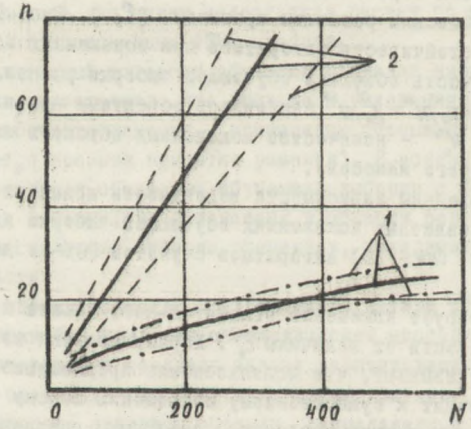

Рис. 2

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

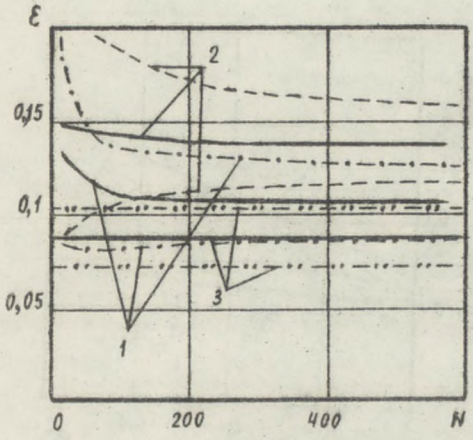

Рис. 3

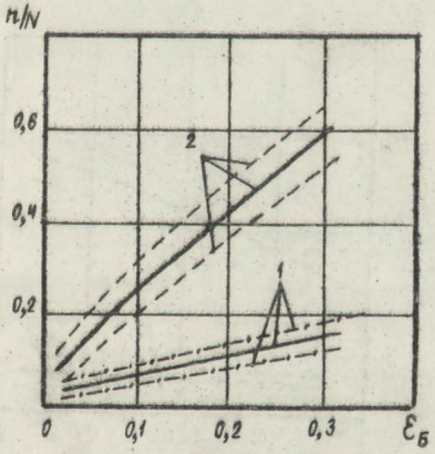

Рис. 4

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

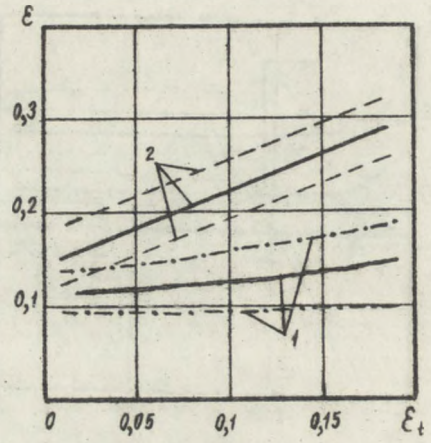

Рис. 5

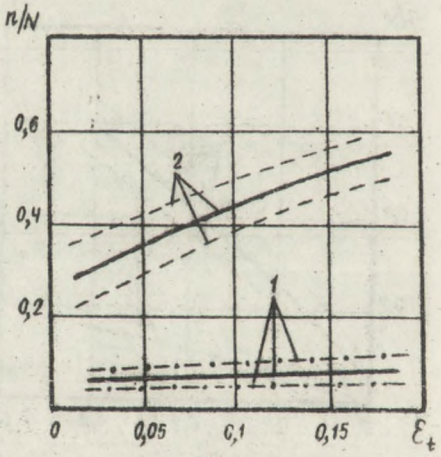

Рис. 6

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ЛИТЕРАТУРА

1. Фукунага К. Введение в статистическую теорию распознавания образов. - М.: Наука, 1979. - 367с.
2. Chin-Liang Chang. Finding prototypes for nearest neighbor classifiers. IEEE Trans. Comput., IT-23, №11, 1974.
3. Hart P.E. The condenced nearest neighbor rule. IEEE Trans. Information Theory, IT-14, 1968.
4. Gates G.W. The reduced nearest neighbor rule. IEEE Trans. Information Theory, IT-18, №3, 1972.
5. Tomek I. Two modification of CNN. IEEE Trans. Syst., Man, and Cybern., IT-6, №11, 1976.
6. Wilson D.L. Asymptotic properties of nearest neighbor rules using edited date. IEEE Trans. Syst., Man, and Cybern., v. SMC-2, July, 1972.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

НЕПАРАМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ НЕЛИНЕЙНЫХ СТОХАСТИЧЕСКИХ ПРОЦЕССОВ С ТРЕНДОМ

Й.Й.Балтрунас, Р.С.Вальтерите

Рассмотрим задачу идентификации нелинейных стохастических процессов (НСП) с наличием детерминированной составляющей (тренда), которые описываются следующего типа разностными уравнениями:

$$\left.\begin{aligned} X_t &= F(X_{t-1}, X_{t-2}, \dots, X_{t-n}) + \xi_t; \\ y_t &= X_t + \psi_t . \end{aligned}\right\} \qquad (I)$$

Для того, чтобы обеспечить устойчивость процессов, а также возможность однозначной идентифицируемости по результатам наблюдений $\{y_t\}$ $(t=\overline{1,N})$, введем совокупность предположений.

I. Последовательность $\{\xi_t\}$ состоит из независимых, одинаково распределенных случайных величин с нулевым математическим ожиданием и конечной дисперсией, т.е.

$$E(\xi_t) = 0, \quad D(\xi_t) = \sigma_\xi^2, \quad \forall t, \quad 0 < \sigma_\xi^2 < \infty. \qquad (2)$$

2. Условием устойчивости процесса (I) является

$$\forall x, y \in R^m \quad |F(x+y) - F(y)| \le \alpha_1 |x_1| + \dots + \alpha_n |x_n|, \qquad (3)$$

$$F(0) = 0, \quad \alpha\alpha_1 + \dots + \alpha_n = \alpha < 1$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Если функция $F(\cdot)$ является дифференцируемой, то указанное условие означает, что

$$\sum_j |\partial F(\cdot)/\partial x_j| < 1 . \qquad (4)$$

F_N^* 3. Регулярность, или перемешиваемость, процесса (I). Через F_N^* обозначим σ -алгебру, сгенерированную случайными величинами $y_N, y_{N+1}, \ldots, y_t$. Тогда

$$\rho(\tau) = \sup_{A^* \in F_{-\infty}^{\circ},\ B^* \in F_{\tau}^{\infty},} |P(AB) - P(A)P(B)| \le Ce^{-\delta\tau}, \qquad (5)$$

где C и δ – константы, зависящие от свойств функции $F(\cdot)$.

Пусть наблюдаем

$$\left.\begin{aligned} x_t &= F(x_{t-1}) + \xi_t \ ; \\ y_t &= x_t + \psi_t \qquad (t = \overline{1,N}) . \end{aligned}\right\} \qquad (6)$$

Идентификация НСП с трендом типа (6) является многоэтапной задачей. Сначала непараметрическим методом оценивается тренд ψ_t, потом он выделяется из соответствующего стохастического процесса. Остаточный нелинейный процесс авторегрессии x_t идентифицируется непараметрическим методом.

Тренд ψ_t в (6) рассмотрим как детерминированную функцию со случайным аргументом, т.е.

$$\psi_t = \varphi(\eta_t) , \quad (t = \overline{1,N}) , \qquad (7)$$

где $\varphi(\cdot)$ – детерминированная функция; $\eta_t = \nu\eta_t^*$, а $\{\eta_t^*\} \in \{N(0,1), R[-1,1]\}$ или случайные величины с некоторым другим законом распределения.

Согласно (7) оценку $\hat{\varphi}(x)$ тренда ψ_t понимаем как значение $\hat{\varphi}(x)$ в точке x, где x – некоторая фиксированная величине из множества $\{\eta_t , t = \overline{1,N}\}$.

Для определения оценки $\hat{\varphi}(x)$ функции тренда ψ_t будем использовать непараметрический робастный подход, так называемые робастные M -оценки, имеющие вид

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$Q(x, h_N, c) = \sum_{t=1}^{N} F(y_t - c)\, K\left(\frac{y_t - x}{h_N}\right), \qquad (8)$$

где $F(u)$ – выпуклая функция, которая стремится к $+\infty$ при $|u| \to \infty$; $K(\cdot)$ – ядро типа Парзена; h_N – шаг; c – параметр.

Оценка $\hat{\psi}(x)$ определяется как минимум эмпирической функции потерь (8), т.е.

$$\hat{\psi}_N(x) = \arg\left\{\min_c \sum_{t=1}^{N} F(y_t - c)\, K\left(\frac{y_t - x}{h_N}\right)\right\}. \qquad (9)$$

Если считать, что дисперсия ненаблюдаемого сигнала ξ_t конечна, т.е. $\sigma_\xi^2 < \infty$, тогда в качестве функции $F(\cdot)$ используется квадратическая функция потерь $F(x) = x^2$. Это оценка типа Надарая–Ватсона /I/:

$$\hat{\psi}_N(x) \triangleq \hat{c} = \sum_{t=1}^{N} y_t\, K\left(\frac{y_t - x}{h_N}\right) \Big/ \sum_{t=1}^{N} K\left(\frac{y_t - x}{h_N}\right). \qquad (10)$$

В случае, когда σ_ξ^2 сигнала ξ_t бесконечна или не существует (например, ξ_t распределена по закону Коши), в качестве оценки $\psi(x)$ принимается оценка $\hat{\psi}_N(x)$, которая минимизирует эмпирическую модульную функцию потерь $F(x) = |x|$, т.е.

$$\hat{\psi}_N(x) = \arg\left\{\min_c \sum_{t=1}^{N} |y_t - c|\, K\left(\frac{y_t - x}{h_N}\right)\right\}. \qquad (11)$$

Минимум (II) имеет вид

$$\hat{\psi}_N(x) \triangleq \hat{c} = med(x) = \begin{cases} y_{n/2}, & \text{когда } n \text{ – нечетное;} \\ \left(y_{n/2} + y_{n/2+1}\right)/2, & \text{когда } n \text{ – четное,} \end{cases} \qquad (12)$$

где n – число y_t, попавших в область, ограниченную ядром $K(\cdot)$.

Из выражения (8) следует, что оценка $\hat{\psi}_N(x)$ зависит от типа ядра $K(\cdot)$, величины шага h_N и длины реализации

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$\{y_t\}$ $(t = \overline{1,N})$.

В общем случае ядро $K(\cdot)$ и шаг h_N должны удовлетворять некоторым условиям:

$$\left.\begin{aligned} &K(x) \geq 0;\\ &K(x) = K(-x);\\ &K(x_2) \geq K(x_1), \text{ когда } |x_2| \leq |x_1|;\\ &\int K(x)\, x^i dx < \infty, \quad i \geq 2;\\ &h_N \geq 0,\ h_N \to 0, \text{ когда } N \to \infty. \end{aligned}\right\} \qquad (13)$$

На основе математического моделирования по множеству реализаций $\{y_t^{(\nu)}\}$ $(t = \overline{1,N},\ \nu = \overline{1,M})$, представляющих собой исследуемый стохастический процесс (I), были проведены исследования с целью определения оптимального шага h_N^*, оценки тренда $\hat{\psi}_N(X_k)$, выделения его из процесса y_t, оценки остаточного процесса нелинейной авторегрессии $\hat{F}_N(X_k)$ и определения статистических свойств оценок $\hat{\psi}_N(X_k)$ и $\hat{F}_N(X_k)$.

Оценки $\hat{\psi}(X_k)$ вычисляются в интервале

$$X_k \in \left[-3\hat{\sigma}_y + \hat{m}_y,\ \hat{m}_y + 3\hat{\sigma}_y\right] \triangleq I_3, \quad k = \overline{1,L},$$

$\hat{m}_y$, $\hat{\sigma}_y$ – оценки среднего и среднеквадратического отклонения для первой реализации $\{y_t^{(1)}\}$ $(t = \overline{1,N})$. По данным $\{y_t^{(1)}\}$ вычисляется статистика

$$\mathcal{I}_k^{(1)}(h) = \frac{1}{L}\sum_{k=1}^{L}\left[y_k^{(1)} - \hat{\psi}_N^{(1)}(X_k)\right]^2 \qquad (14)$$

и из выражения

$$h_{опт.} \triangleq h^* = \arg\left\{\min_h \mathcal{I}_k^{(1)}(h)\right\} \qquad (15)$$

методом Пауэлла определяется оптимальный шаг $h_{опт}$ /3/.

Далее по наблюдениям $\{y_t^{(\nu)}\}$ с шагом $h = h^*$ вычисляются

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

оценки $\hat{\psi}_N^{(\nu)}(X_\kappa)$ $(\kappa = \overline{1,L}, \nu = \overline{1,M})$, определяются средние значения

$$\overline{\psi}_N(X_\kappa) = \frac{1}{M}\sum_{\nu=1}^{M}\hat{\psi}_N^{(\nu)}(X_\kappa) \quad (\kappa = \overline{1,L}). \tag{16}$$

Ряд $\{\overline{\psi}_N(X_\kappa), \kappa = \overline{1,L}\}$ аппроксимируется некоторой функцией (функциями) $\varphi_i(X_\kappa, \beta_i)$:

$$\hat{\beta}_i : Q(\hat{\beta}_i) = \arg\left\{\min_{\beta}\left[\overline{\psi}_N(X_\kappa) - \varphi_i(X_\kappa, \beta_i)\right]^2\right\}. \tag{17}$$

Функционал (17) минимизируется методом Маркуардта (подпрограмма *NLIN*).

Оценку $\hat{\psi}(X_t)$ $(t = \overline{1,N})$ тренда $\psi(X_\kappa)$, где точка аргумента $X_t \triangleq \eta_t$, вычисляем по выражению

$$\hat{\psi}_t = \varphi_i(X_t, \hat{\beta}_i). \tag{18}$$

Далее согласно (6) выделяется тренд

$$\hat{X}_t^{(\nu)} = y_t^{(\nu)} - \hat{\psi}_t. \tag{19}$$

Точность оценивания тренда $\psi(X_\kappa)$ определяется статистикой

$$\rho_\psi = \sqrt{\sum_{\kappa=1}^{L}\left(\psi(X_\kappa) - \hat{\psi}(X_\kappa)\right)^2 \Big/ \sum_{\kappa=1}^{L}\psi^2(X_\kappa)}, \tag{20}$$

а статистики ρ_1, ρ_2 и ρ_3 характеризуют качество оценивания тренда ψ_t и процесса X_t во всех точках реализации $\{y_t^{(\nu)}\}$ $(t = \overline{1,N})$. Они вычисляются следующим образом:

$$\rho_1 = \sqrt{\sum_{t=1}^{N}\left\{\frac{1}{M}\sum_{\nu=1}^{M}\left[\psi^{(\nu)}(X_t) - \frac{1}{M}\sum_{\nu=1}^{M}\hat{\psi}^{(\nu)}(X_t)\right]^2\right\} \Big/ \sum_{t=1}^{N}\left[\frac{1}{M}\sum_{\nu=1}^{M}\psi^{(\nu)}(X_t)\right]^2}; \tag{21}$$

$$\rho_2 = \sqrt{\sum_{t=1}^{N}\left[\frac{1}{M}\sum_{\nu=1}^{M}\left(X_t^{(\nu)} - \frac{1}{M}\sum_{\nu=1}^{M}\hat{X}_t^{(\nu)}\right)\right]^2 \Big/ \sum_{t=1}^{N}\left(\frac{1}{M}\sum_{\nu=1}^{M}X_t^{(\nu)}\right)^2}; \tag{22}$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\rho_3 = \sqrt{\sum_{t=1}^{N}\left[\frac{1}{M}\sum_{\nu=1}^{M} X_t^{(\nu)} \cdot \frac{1}{M}\sum_{\nu=1}^{M}\Delta X_t^{(\nu)}\right]^2 \Big/ \sum_{t=1}^{N}\left(\frac{1}{M}\sum_{\nu=1}^{M} X_t^{(\nu)}\right)^2}, \qquad (23)$$

где $\Delta X_t^{(\nu)} = X_t^{(\nu)} - \hat{X}_t^{(\nu)}$.

Для идентификации процесса нелинейной авторегрессии первого порядка $X_t = F(X_{t-1}) + \xi_t$ в (6), полученный после выделения тренда, ряд $\{\hat{X}_t^{(\nu)}\}$ оценивается с помощью непараметрического метода типа Надараи–Ватсона. Для этой цели формируются два ряда:

$$\begin{aligned} X_t &:= \hat{X}_t \qquad (t = \overline{1, N-1}); \\ Z_t &:= \hat{X}_{t+1}, \end{aligned} \qquad (24)$$

и оценка функции авторегрессии вычисляется по формуле

$$\hat{F}(X_\kappa) = \sum_{t=1}^{N-1} Z_t K\left(\frac{X_t - X_\kappa}{h_N}\right) \Big/ \sum_{t=1}^{N-1} K\left(\frac{X_t - X_\kappa}{h_N}\right). \qquad (25)$$

Здесь ядро $K(\cdot)$ и шаг h_N удовлетворяют условиям (I3). Точность оценивания функции $F(X_\kappa)$ и характеристика качества оценивания вычисляются аналогично статистикам (20) и (2I).

В докладе приводятся результаты математического моделирования с процессом нелинейной авторегрессии первого порядка и трендами различного типа, когда ξ_t в (6) распределен по законам $\{N(0, \sigma_\xi^2), L(m_\xi, \sigma_\xi^2), C(\xi_{\text{нач.}})\}$, где $N(\cdot,\cdot)$ – закон Гаусса, $L(\cdot,\cdot)$ – закон Лапласа, $C(\cdot)$ – закон Коши.

Оценки $\{\hat{\psi}(X), \hat{F}(X)\} \overset{\Delta}{=} f(X)$ обладают некоторыми статистическими свойствами при $N \to \infty$. Для оценок типа (I0) и (25) величины

$$(Nh_N)^{1/2}\left\{\hat{f}_N(X_\kappa) - f(X_\kappa)\right\} \xrightarrow{D} N(0, \sigma^2(x)), \qquad (26)$$

$\xrightarrow{D}$ – сходятся по распределению. Дисперсия $\sigma^2(X)$ имеет следующий вид:

$$\sigma^2(x) = \frac{1}{P(x)} \int_{R^m} K^2(x)\, dx\, D[y_t / y_t = x], \qquad (27)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

где $p(x)$ – плотность распределения случайного вектора $y_t \in R^1$ в точке $y_t = x$;

$$D[y_t / y_t = x] = \sigma_j^2 < \infty . \qquad (28)$$

Для оценок $\hat{f}_N(x)$ типа (II) дисперсия $\sigma^2(x)$ в выражении (26) имеет вид

$$\sigma^2(x) = \frac{1}{p(x)} \frac{K}{[p(f(x)/x)]^2} , \qquad (29)$$

где $p(f(x)/x) = P_j(0)$, $P_j(0) > 0$, а $p(x/y) = P_j(x - f(y))$ /4, 5/.

Литература

1. Надарая Э.А. Непараметрическое оценивание плотности вероятностей и кривой регрессии. – Тбилиси: изд-во Тбилис у-та, 1983.-191с.
2. Кошкин Г.М. Об одном подходе к исследованию функционалов от условных распределений при статистической неопределенности.– Автоматика и телемеханика, 1978, № 8, с.53-65.
3. Химмельблау Д. Прикладное нелинейное программирование. – М.: Мир, 1975. – 534 с.
4. Roussas G.G. Nonparametric Estimation of the Transition Distribution Function of a Markov Process. – The Annals of Mathematical Statistics, 1969, t. 40, Nr. 4, c. 1386 – 1400.
5. Robinson P.M. Robust Nonparametric Autoregression. – Lecture Notes in statistics, 1974, t. 26, c. 247-255.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

НЕПАРАМЕТРИЧЕСКИЕ ОЦЕНКИ НЕЛИНЕЙНОЙ ФУНКЦИИ АВТОРЕГРЕССИИ

Л.Й.Балтрунас, В.Ю.Рудзкене

Пусть процесс авторегрессии определяется нелинейным разностным уравнением

$$X_t = f(X_{t-1}, \dots, X_{t-n}) + \varepsilon_t , \qquad (1)$$

где ε_t – последовательность независимых, одинаково распределенных гауссовских случайных величин; f – неизвестная функция, удовлетворяющая условию устойчивости процесса (I).

Рассматривается задача построения состоятельной оценки $\hat{f}(x)$, $x \in [-1,1]^n$ по выборке $X_1, X_2, \dots, X_N$ процесса (I). Обозначим $V_t = (X_{t-1}, \dots, X_{t-n})$. В непараметрическом оценивании важное место занимают так называемые ядерные оценки. В рассматриваемом случае это оценки вида

$$\hat{f}(x) = \sum_{t=n+1}^{N} X_t K\left(\frac{\|V_t - x\|}{h}\right) \Big/ \sum_{t=n+1}^{N} K\left(\frac{\|V_t - x\|}{h}\right) , \qquad (2)$$

где $K(\cdot)$ – некоторая функция, не зависящая от N, $h = h(N)$.

Известно (см., например [1]), что при достаточной гладкости функции f и при определенных ограничениях на $K(\cdot)$ и h для процессов вида (I) имеет место соотношение

$$E[f(x) - \hat{f}(x)]^2 = O\left(N^{-\frac{4}{4+n}}\right) , \quad N \to \infty . \qquad (3)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

С другой стороны, если f является полиномом, то можно построить оценку, для которой среднеквадратическая ошибка имела бы порядок $O(N^{-1})$. Поэтому тот факт, что для сколь угодно гладкой функции f (за исключением случая $f = const$) ядерные оценки дают среднеквадратическую ошибку только порядка $O(N^{-4/(4+n)})$, обосновывает важность изучения так называемых проекционных оценок, которые являются весьма популярными в математической статистике.

Пусть $\{e_k(x)\}$ – полная система ортонормированных функций в пространстве $L_2(-1,1)^n$. Тогда

$$f(x) = \sum_{k=0}^{\infty} c_k e_k(x), \quad x \in [-1,1]^n, \tag{4}$$

где $c_k = \int_D f(x) e_k(x) dx$, $D \in [-1,1]^n$. Определим оценку $\hat{f}$ равенством

$$\hat{f}(x) = \sum_{k=0}^{m} \hat{c}_k e_k(x), \tag{5}$$

где $m = m(N)$, $\hat{c}_k = \int_D e_k(z) \cdot x_{t(z)} dz$, $t(z) = \arg\left\{\min_{s=n+1,\dots,N} \|v_s - z\|\right\}$

В работе [2] показано, что при $n \leq 4$ в зависимости от степени гладкости функции f величина среднеквадратического отклонения оценки $\hat{f}$ вида (5) может иметь порядок сколь угодно близкий к $O(N^{-1})$. Например, если $|e_k(x)| \leq C$, $k = 1, 2, \dots$, и $|c_k| \leq c_1 e^{-c_2 k}$, то

$$E \int_D [f(x) - \hat{f}(x)]^2 dx = O(\ln N / N)$$

при выборе $M = \ln N$.

Оценка (5) легко вычисляется при $n = 1$. Однако при $n > 1$ алгоритм вычисления оценки (5) усложняется. С вычислительной точки зрения более удобными являются проекционные оценки, предложенные Н.Н.Ченцовым. Известно [3], что если случайный вектор (x_t, v_t) имеет непрерывную плотность распределения, то функ-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

цию f можно записать в виде

$$f(x) = \sum_{i=1}^{\infty} a_i e_i(x) \Big/ \sum_{j=1}^{\infty} b_j e_j(x) \,, \qquad (6)$$

где $a_i = E\, x_t e_i(V_t)$, $b_j = E\, e_j(V_t)$.

Пусть

$$\hat{f}(x) = \sum_{i=1}^{m_1} \hat{a}_i e_i(x) \Big/ \sum_{j=1}^{m_2} \hat{b}_j e_j(x) \,, \qquad (7)$$

где $\hat{a}_i = \frac{1}{N-n} \sum_{t=n+1}^{N} x_t e_i(V_t)$, $\hat{b}_j = \frac{1}{N-n} \sum_{t=n+1}^{N} e_j(V_t)$,

$m_k \to \infty$ при $N \to \infty$.

В докладе приводятся результаты исследования оценок вида (2), (5), (7) при помощи моделирования последовательности $\{x_t\}$ на ЭВМ.

Литература

I. Collomb G. Estimation non parametrique de la regression: revue bibliographique.- International Statistical Review, 1981, v. 49 , p. 75-93.

2. Балтрунас И.И., Рудзкене В.Ю. Непараметрическое оценивание нелинейных процессов авторегрессии. - Тр./АН ЛитССР. Серия Б, 1987, т. 3(160), с. 85-98.

3. Ченцов Н.Н. Оценка неизвестной плотности распределения по наблюдениям.- ДАН СССР, 1962, т.147, №1, с.45-48.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

НЕПАРАМЕТРИЧЕСКИЕ ОЦЕНКИ ПО СЛУЧАЙНО ЦЕНЗУРИРОВАННЫМ СПРАВА ВЫБОРКАМ

Ю.Н.Благовещенский

Пусть $\mathcal{F}$ - семейство непрерывных распределений F в R^1 и статистическое пространство $(\Omega, \mathcal{A}, \mathcal{P})$ задано семейством

$$\mathcal{P} = \{P_F, F \in \mathcal{F}\}$$

вероятностных мер P_F на $\mathcal{A}$. Модель Каплана-Мейера представляет случайно цензурированную выборку $\tau^{(n)} = \{\tau_1, \dots, \tau_n\}$, где τ_j - пара величин $\xi_j = \min(\xi_j, \eta_j)$, $w_j = I\{\xi_j \le \eta_j\}$, $j = 1, \dots, n$, определяемых по элементам двух независимых между собой полных выборок

$$\xi^{(n)} = \{\xi_1, \dots, \xi_n\} \quad \text{и} \quad \eta^{(n)} = \{\eta_1, \dots, \eta_n\},$$

причем $P_F\{\xi_i < x\} = F(x)$, $P_F\{\eta_i < x\} = G_F(x)$ и $G_F(x)$ выступает в качестве мешающего параметра при получении тех или иных выводов об $F(x)$.

Впервые задача оценивания $F(x)$ в такой ситуации рассмотрена в работе [1], где была предложена оценка

$$F_n(x) = 1 - \prod_{j=1}^{n} \left(1 - w_j I\{\xi_j \le x\} (n - R_n(\xi_j))^{-1}\right) \qquad (1)$$

с $R_n(u) = \sum_{i=1}^{n} I\{\xi_i < u\}$ и установлена ее сходимость к $F(x)$ в среднем квадратическом, если

$$x < \theta_F = \sup\{u : (1 - G_F(u))(1 - F(u)) > 0\}.$$

Оценки типа (1), получившие название PL - оценок, в последние 5-10 лет широко изучались в самых разных аспектах (см., например, [2] и [3]) методами теории мартингалов.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Основные достижения связаны с асимптотическим анализом линейных функционалов от $F_n(x)$, и с обобщением результатов на случай зависимых ξ_j - х (марковская зависимость, $\eta^{(n)}$ - условная независимость ξ_j - х и т.п.).

Наряду с развитием этих теоретических результатов, в прикладных областях знаний (прежде всего - эксплуатационная надежность) развивались свои, эвристические приемы обработки таких данных.

Интересное "переоткрытие" оценки (1) было дано в работе [4], где оценивались ранги членов вариационного ряда $x_1 \le x_2 \le \ldots \le x_\nu$ тех из ξ_j, для которых $w_j = 1$, то есть $\xi_j = \xi_j$ ($\nu = w_1 + \ldots + w_n$) в неизвестном нам вариационном ряду $\xi^{(n)}$. Если оценки этих рангов обозначить $Q_1, \ldots, Q_\nu$, то Джонсон предложил для них следующую рекуррентную процедуру:

$$Q_K = Q_{K-1} + \frac{(n+1) - Q_{K-1}}{(n+1) - R_n(x_K)}, \quad K = 1, \ldots, \nu, \qquad (2)$$

полагая $Q_0 = 0$. До последнего времени оценки Джонсона исследовались лишь посредством машинного моделирования и рассматривались как некая альтернатива к PL - оценкам (см., например, [5]).

По-видимому, в статье [6] впервые было установлено, несмотря на всю тривиальность этого, что замена в (2) величины $(n+1)$ на n приводит в точности к PL-оценке (1) в точках $x = x_1, x_2, \ldots, x_\nu$, так что все асимптотические результаты для этих двух оценок совпадают, хотя численное различие при конечных n и разные алгоритмы счета мешали это заметить.

Наконец, в прекрасной работе [7] прикладного характера была предложена оценка $H_n(x)$ для крайне важной в теории надежности функции $H(x) = -\ln(1-F(x))$, а именно:

$$H_n(x_K) = \sum_{\ell=1}^{K} \left(n - R_n(x_\ell)\right)^{-1}, \quad K = 1, \ldots, \nu. \qquad (3)$$

Асимптотическое исследование (3) на состоятельность и асимптотическую нормальность было впервые проведено автором и повторено другим методом в [8] и [2], соответственно.

Эти оценки, как и PL-оценки, используются на практике и для графического (графоаналитического) оценивания параметров сдвига и масштаба при заданном типе распределения ξ_j - х или их логарифмов. Эта работа посвящена анализу этих оценок, их обобщению, $1/\sqrt{n}$ - асимптотике методом аппроксимации U- статистиками

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

и обоснованию графоаналитического метода оценивания. На основании этих исследований предлагается состоятельный критерий проверки гипотез о типе распределения.

Пусть $C_n \geq n$ и функция $\gamma(v) \geq v$, определенная для $v \geq 1$, таковы, что

$$\gamma(v) = 2 \quad \text{для} \quad v \in [1;2];$$
$$\gamma(v) \leq v + C_0;$$
$$C_n \leq n + C_1 \qquad (4)$$

с некоторыми константами C_0 и C_1. Определим

$$F_n(x/C_n,\gamma) = 1 - \prod_{j=1}^{n}\left(1 - \frac{w_j\, I\{\xi_j \leq x\}}{\gamma(C_n - R_n(\xi_j))}\right). \qquad (5)$$

Очевидно, что $F_n(x/C_n,\gamma_1) \neq F_n(x/C_n,\gamma_2)$ при $\gamma_1(v) = \gamma_2(v)$ для $v \geq v_0$ лишь тогда, когда при каком-нибудь j одновременно $\xi_j \leq x$ и $C_n - R_n(\xi_j) \leq v_0$, так что это менее вероятно, чем неравенство $n - R_n(\xi_j) \leq v_0$ (из-за неубывания $R_n(x)$ по x); однако $P_F\{R_n(x) \geq n - v_0\} \leq Cn^{-s}$ с любым $s < \infty$ и некоторым $C < \infty$, если только $-\infty < x \leq x_0$ и $x_0 < \theta_F$. Из отмеченного факта следует, что при конструировании оценок по (5) можно "забыть" интервал $v \in [1;2]$ как пренебрежимый в их исследовании, и не различать функции $\gamma(v)$; совпадающие при $v \geq 2$. С учетом этого замечания мы можем считать, что PL - оценка из (1) суть $F_n(x/C_n,\gamma_0)$ с $C_n = n$ и $\gamma_0(v) = v$.

ЛЕММА I. Оценки Джонсона (2) определяются равенством

$$Q_k = (n+1)\, F_n(x_k/C_n,\gamma_D) \qquad (6)$$

при $C_n = n+1$ и $\gamma_D(v) = v$, $k = 1,\dots,\nu$,

а оценки $H_n(x_k)$ Нельсона (3) определяются равенством

$$H_n(x_k) = -\ln\left[1 - F_n(x_k/n,\gamma_H)\right],$$

где $\gamma_H(v) = \left(1 - \exp\left(-\frac{1}{v}\right)\right)^{-1}$. (7)

Доказательство начнем со второй части леммы, заметив, что оценка

$$H_n(x) = \sum_{j=1}^{n} \frac{w_j\, I\{\xi_j \leq x\}}{n - R_n(\xi_j)} \qquad (8)$$

в точках $x = x_k$ совпадает с $H_n(x_k)$, так что для обоснования (7) достаточно рассмотреть

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$F_n^*(x) = 1 - \exp(-H_n(x)),$$

но

$$\exp(-H_n(x)) = \prod_{j=1}^{n} \exp\left(-\frac{w_j I\{\xi_j \le x\}}{n - R_n(\xi_j)}\right) =$$

$$= \prod_{j=1}^{n} \left(1 - w_j I\{\xi_j \le x\}\left(1 - \exp\left(-\frac{1}{n - R_n(\xi_j)}\right)\right)\right) =$$

$$= \prod_{j=1}^{n} \left(1 - \frac{w_j I\{\xi_j \le x\}}{\gamma_H(n - R_n(\xi_j))}\right) = 1 - F_n(x/n, \gamma_H).$$

Для доказательства первой части заметим, что при любой $\gamma(v)$ и любых C_n

$$\frac{1 - F_n(x_k/C_n, \gamma)}{1 - F_n(x_{k-1}/C_n, \gamma)} = 1 - \frac{1}{\gamma(C_n - R_n(x_k))}, \qquad (9)$$

как это видно из определения X_k и представления (5) для всех $k = 1, \dots, \nu$, если положить $x_0 = -\infty$ (то есть, $F_n(x_0/C_n, \gamma) = 0$). Умножая числитель и знаменатель левой части (9) на C_n и вводя

$$Q_k(C_n, \gamma) \equiv C_n F_n(x_k/C_n, \gamma), \qquad (10)$$

получим сразу же соотношения

$$Q_k(C_n, \gamma) = Q_{k-1}(C_n, \gamma) + \frac{C_n - Q_{k-1}(C_n, \gamma)}{\gamma(C_n - R_n(x_k))} \qquad (11)$$

для $k = 1, 2, \dots, \nu$ с $Q_0(C_n, \gamma) = 0$.

Равенство (11) доказывает при $\gamma(v) = v$ и $C_n = n+1$ лемму окончательно, а дополнительно вскрывает общую "природу" всех введенных оценок.

Отметим лишь, что

$$v\left[\gamma_H(v)\right]^{-1} = 1 - \frac{1}{2!v} + \frac{1}{3!v^2} - \frac{1}{4!v^3} + \dots, \quad v \ge 1,$$

так что $\gamma_H(v)$ удовлетворяет требованиям (4) при $v \ge 2$.

А теперь с каждой $\gamma(v)$ из (4) свяжем функцию $\gamma^*(v)$ равенством

$$\gamma^*(v) = -\frac{1}{\ln(1 - [\gamma(v)]^{-1})}. \qquad (12)$$

Отметим следующее: так как $\gamma(v) \ge 2$, то $\gamma^*(v) \ge \frac{1}{\ln 2}$,

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

и так как

$$v \le \gamma(v) \le v + C_0,$$

то легко оценить, что

$$|\gamma^*(v) - v| \le C_0^*. \qquad (13)$$

Воспользовавшись (12), для любой оценки из (5) мы можем теперь получить

$$H_n(x|C_n, \gamma^*) \equiv \ln \frac{1}{1 - F_n(x|C_n, \gamma)} = \sum_{j=1}^{n} \frac{w_j I\{\xi_j \le x\}}{\gamma^*(C_n - R_n(\xi_j))}. \qquad (14)$$

ТЕОРЕМА I. Если $x \le x_0$, $x_0 < \theta_F$ и $H_n(x)$ – оценка Нельсона, то

$$n M \sup_{x \le x_0} |H_n(x) - H_n(x|C_n, \gamma^*)| \le C_3 < \infty. \qquad (15)$$

Из (15) сразу же следует, что $1/\sqrt{n}$-асимптотика всего класса оценок из (5) или (14) одна и та же, независимо от выбора C_n и $\gamma(v)$, удовлетворяющих (4) при некоторых $n \ge n_0$ и $v \ge v_0$.

Доказательство. Из (13) и того, что $\gamma^*(v) \ge \frac{1}{\ln 2}$, легко получить, что

$$\left| \frac{1}{\gamma^*(C_n - R_n(\xi_j))} - \frac{1}{n - R_n(\xi_j)} \right| \le \frac{C_1^*}{(n - R_n(\xi_j))^2} \qquad (16)$$

при некотором достаточно большом $C_1^* < \infty$ (с учетом $n \le C_n \le n + C_1$, конечно). Таким образом,

$$n |H_n(x) - H_n(x|C_n, \gamma^*)| \le C_1^* n \sum_{j=1}^{n} \frac{w_j I\{\xi_j \le x\}}{(n - R_n(\xi_j))^2}. \qquad (17)$$

Далее, так как все элементы суммы в правой части (17) одинаково распределены (хоть и зависимы), то левая часть из (15) оценивается величиной

$$a_n = C_1^* n^2 M \frac{w_1 I\{\xi_1 \le x_0\}}{(n - R_n(\xi_1))^2},$$

но

$$\frac{w_1 I\{\xi_1 \le x\}}{(n - R_n(\xi_1))^2} \le \frac{w_1 I\{\xi_1 \le x_0\}}{(n - \sum_{i=2}^{n} I\{\xi_i < x_0\})^2} \le$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\leq \frac{1}{(1+\sum_{i=2}^{n} I\{\xi_i \geq x_0\})^2},$$

так что

$$a_n \leq C_1^* n^2 \sum_{k=0}^{n} (1+k)^{-2} C_{n-1}^{k} (p(x_0))^k (1-p(x_0))^{n-k-1},$$

где $p(x_0) = (1-F(x_0))(1-G_F(x_0)) > 0$. Далее, так как

$$\frac{n^2 C_{n-1}^{k}}{(k+1)^2} \leq 2 C_{n+1}^{k+2}, \quad k = 0, 1, \ldots, n-1,$$

то

$$a_n \leq 2C_1^* \sum_{k=0}^{n-1} C_{n+1}^{k+2} \frac{(p(x_0))^{k+2} (1-p(x_0))^{n+1-(k+2)}}{(p(x_0))^2} \leq$$

$$\leq \frac{2C_1^*}{(p(x_0))^2} < \infty,$$

что и завершает доказательство.

Рассмотрим статистики

$$T_n(x) = \frac{1}{n} \sum_{j=1}^{n} \frac{w_j a(\xi_j) I\{\xi_j \leq x\}}{(1-\psi(\xi_j))^2} \left[1 + \psi_n(\xi_j) - 2\psi(\xi_j)\right], \quad (18)$$

где $\psi_n(u) = n^{-1} R_n(u)$, $\psi(u) \equiv 1-(1-F(u))(1-G_F(u))$.

Статистики (18) являются аппроксимационными для изучаемых ниже статистик типа "Нельсона". Установим $\frac{1}{\sqrt{n}}$ - асимптотическую нормальность $T_n(x)$ и найдем асимптотические среднее и дисперсию.

Пусть γ_{nj} - j- й элемент суммы в (18). Легко видеть, что набор $(\gamma_{nj_1}, \ldots, \gamma_{nj_s})$ при любых различных $j_1, \ldots, j_s$ имеет то же распределение, что и $(\gamma_{n1}, \ldots, \gamma_{ns})$. Отсюда сразу получим

$$M T_n(x) = M \gamma_{n1}$$

и

$$M T_n^2(x) = \frac{1}{n} M \gamma_{n1}^2 + \frac{n-1}{n} M \gamma_{n1} \gamma_{n2}.$$

Займемся их вычислением покомпонентно, установив следующие важные вещи:

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\psi_n(\zeta_1) = \frac{1}{n}\sum_{j=1}^{n} I\{\zeta_j < \zeta_1\} = \frac{1}{n}\sum_{j=2}^{n} I\{\zeta_j < \zeta_1\}, \qquad (19)$$

поскольку $I\{\zeta_1 < \zeta_1\} = 0$;

$$\begin{cases} \psi_n(\zeta_1) = \frac{1}{n} I\{\zeta_2 < \zeta_1\} + \frac{1}{n}\sum_{j=3}^{n} I\{\zeta_j < \zeta_1\}; \\ \psi_n(\zeta_2) = \frac{1}{n} I\{\zeta_1 < \zeta_2\} + \frac{1}{n}\sum_{j=3}^{n} I\{\zeta_j < \zeta_2\} \end{cases} \qquad (20)$$

по тем же соображениям. Представления (19) и (20) нам будут полезны, поскольку в них при фиксации (w_1, ζ_1) (и фиксации "пары" (w_1, ζ_1) и (w_2, ζ_2)) мы выделяем из $\psi_n(\cdot)$ условно независимые компоненты.

Наконец, будем снабжать математическое ожидание "штрихом" либо "двумя штрихами", когда оно берется условно относительно σ-алгебр, порожденных либо (w_1, ζ_1), либо $(w_1, \zeta_1; w_2, \zeta_2)$ соответственно. Используя (19), получим, что

$$M'\psi_n(\zeta_1) = \frac{n-1}{n}\psi(\zeta_1),$$

так что

$$MT_n(x) = M(M'T_n(x)) = M\frac{w_1 a(\zeta_1) I\{\zeta_1 \le x\}}{(1-\psi(\zeta_1))^2}\left(1 - \frac{n+1}{n}\psi(\zeta_1)\right) =$$

$$= M\frac{w_1 a(\zeta_1) I\{\zeta_1 \le x\}}{1-\psi(\zeta_1)} - \frac{1}{n} M\frac{w_1 a(\zeta_1) I\{\zeta_1 \le x\}\psi(\zeta_1)}{(1-\psi(\zeta_1))^2}. \qquad (21)$$

Аналогично, так как

$$M'\psi_n^2(\zeta_1) = \left(\frac{n-1}{n}\right)^2\left(\psi^2(\zeta_1) + \frac{\psi(\zeta_1)(1-\psi(\zeta_1))}{n-1}\right),$$

то

$$M'(1+\psi_n(\zeta_1) - 2\psi(\zeta_1))^2 =$$

$$= \left(1 - \frac{n+1}{n}\psi(\zeta_1)\right)^2 + \frac{n-1}{n^2}\psi(\zeta_1)(1-\psi(\zeta_1)),$$

, значит,

$$\frac{1}{n} M\gamma_{n1}^2 = \frac{1}{n} M\frac{w_1 a^2(\zeta_1) I\{\zeta_1 \le x\}}{(1-\psi(\zeta_1))^2} -$$

$$- \frac{1}{n^3} M\frac{w_1 a^2(\zeta_1) I\{\zeta_1 \le x\}}{(1-\psi(\zeta_1))^4}\left[(n+1)\psi(\zeta_1)(1-\psi(\zeta_1)) - \psi^2(\zeta_1)\right]. \qquad (22)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Более подробные выкладки для (22) слишком просты, чтобы их приводить. А теперь займемся величиной

$$q_n(\xi_1,\xi_2)\equiv M''\big(1+\psi_n(\xi_1)-2\psi(\xi_1)\big)\big(1+\psi_n(\xi_2)-2\psi(\xi_2)\big). \quad (23)$$

В этом случае "работает" представление (20), которое после элементарных расчетов дает

$$q_n(\xi_1,\xi_2)=(1-\psi(\xi_1))(1-\psi(\xi_2))+$$
$$+\frac{1}{n}\left[1+3\psi(\xi_1)\psi(\xi_2)-2\psi(\xi_1)-2\psi(\xi_2)\right]+$$
$$+\frac{1}{n^2}\left[6\psi(\xi_1)\psi(\xi_2)-4\psi(\min(\xi_1,\xi_2))\right]. \quad (24)$$

При расчетах использовалось дополнительно лишь равенство

$$\psi(\xi_1)I\{\xi_1<\xi_2\}+\psi(\xi_2)I\{\xi_2<\xi_1\}=\psi(\min(\xi_1,\xi_2)),$$

пренебрегая событием $\xi_1=\xi_2$ вероятности нуль для непрерывных F.

Подстановка (24) в формулу

$$M\gamma_{n1}\gamma_{n2}=M\frac{w_1w_2a(\xi_1)a(\xi_2)I\{\xi_1\le x;\xi_2\le x\}}{(1-\psi(\xi_1))^2(1-\psi(\xi_2))^2}q_n(\xi_1,\xi_2) \quad (25)$$

позволяет выписать $MT_n^2(x)$ с учетом (22). Для нас, конечно же, нужен лишь главный член дисперсии $T_n(x)$. Константы в $DT_n(x)$ уничтожаются, поскольку (из (25))

$$M\frac{w_1w_2a(\xi_1)a(\xi_2)I\{\xi_1\le x;\xi_2\le x\}(1-\psi(\xi_1))(1-\psi(\xi_2))}{(1-\psi(\xi_1))^2(1-\psi(\xi_2))^2}=$$
$$=\left(M\frac{w_1a(\xi_1)}{1-\psi(\xi_1)}I\{\xi_1\le x\}\right)^2.$$

Выделяя лишь члены порядка $1/n$, получим, что

$$nDT_n(x)=Mw_1I\{\xi_1\le x\}a^2(\xi_1)(1-\psi(\xi_1))^{-2}+\frac{C_n(x)}{n}, \quad (26)$$

где

$$|C_n(x)|\le C_0(x)+3M\frac{w_1a^2(\xi_1)}{(1-\psi(\xi_1))^2}I\{\xi_1\le x\},$$

уже при $n>3$, что легко оценить при подробном разборе компонент, составляющих $DT_n(x)$.

Очевидно, что проделанные выкладки справедливы лишь при су-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ществовании соответствующих моментов, однако из полученных равенств (21) и (26) мгновенно доказывается следующее утверждение.

ТЕОРЕМА 2. Если $\alpha < \theta_F$ и

$$0 < \int_{-\infty}^{\infty} a^2(u)\,dF(u) < \infty, \qquad (27)$$

то существуют и конечны величины

$$A(x) = M\,\frac{w_1 a(\xi_1) I\{\xi_1 \le x\}}{1-\psi(\xi_1)} \equiv \int_{-\infty}^{x} \frac{a(u)\,dF(u)}{1-F(u)};$$

$$B(x) = M\,\frac{w_1 a^2(\xi_1) I\{\xi_1 \le x\}}{(1-\psi(\xi_1))^2},$$

причем $B(x) > 0$ и

$$\lim_{n\to\infty} P_F\left\{\frac{T_n(x) - A(x)}{\sqrt{B(x)}}\sqrt{n} < t\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} \exp\left(-\frac{u^2}{2}\right)du. \quad (2$$

Доказательство. Прежде всего, заметим, что с любой $q(t)$

$$w_1 I\{\xi_1 \le x\} q(\xi_1) = I\{\xi_1 \le \min(\eta_1, x)\}\, q(\xi_1)$$

(вспомним, что $\xi_1 = \min(\xi_1, \eta_1)$ и $w_1 = I\{\xi_1 \le \eta_1\}$); отсюда

$$B(x) = M\int_{-\infty}^{\min(\eta_1, x)} (1-\psi(u))^{-2} a^2(u)\,dF(u) \le \int_{-\infty}^{x} \frac{a^2(u)\,dF(u)}{(1-\psi(x))^2} < \infty,$$

поскольку $\psi(x) < 1$ и выполнено (27). Далее, снова по (27):

$$B(x) \ge \int_{-\infty}^{x} a^2(u)\,dF(u)\cdot P_F\{\eta_1 \ge x\} > 0,$$

так как $P_F(\eta_1 \ge x) = 1 - G_F(x) \ge 1 - \psi(x) > 0$.

А так как $A(x) \le \sqrt{B(x)}$, то остается лишь доказать (28). Но $T_n(x)$, как легко обнаружить, можно переписать в виде

$$T_n(x) = \frac{1}{n^2}\sum_{i,j} Q(\xi_i, w_i; \xi_j, w_j), \qquad (29)$$

причем функция

$$Q(z_1, \varepsilon_1; z_2, \varepsilon_2) = \frac{1}{2}\left[\varepsilon_1 q(z_1)\left(1 + I\{z_2 < z_1\} - 2\psi(z_1)\right) + \right.$$
$$\left. + \varepsilon_2 q(z_2)\left(1 + I\{z_1 < z_2\} - 2\psi(z_2)\right)\right]$$

с $q(z) = \dfrac{a(z) I\{z \le x\}}{(1-\psi(z))^2}$ симметрична относительно пар (z_1, ε_1), (z_2, ε_2). Представление (29) вместе с уже доказанным позволяет применить к $T_n(x)$ теоремы об асимптотической нормальности U-статистик (см. [9]), что и завершает доказательство.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Оценками типа Нельсона будем называть статистики

$$S_n(x) = \frac{1}{n} \sum_{j=1}^{n} \frac{a(\xi_j) w_j \cdot I\{\xi_j \le x\}}{1 - \psi_n(\xi_j)}, \qquad (30)$$

где $\psi_n(u) = \frac{1}{n} R_n(u) = \frac{1}{n} \sum_{i=1}^{n} I\{\xi_i < u\}$ является обычной оценкой функции $\psi(x) = P_F\{\xi_j < x\} = 1 - (1-F(x))(1-G_F(x))$. Будем предполагать, что $x < \theta_F$, где $\theta_F = \sup\{u : \psi(u) < 1\}$,

и для таких x:

$$\int_{-\infty}^{x} |a(u)|^k dF(u) < \infty \qquad (31)$$

при некотором $k \ge 2$.

Из тождества

$$\frac{1}{\beta} = \frac{2\alpha - \beta}{\alpha^2} + \left(\frac{\beta - \alpha}{\alpha}\right)^2 \cdot \frac{1}{\beta}$$

при $\alpha = 1 - \psi(\xi_j)$ и $\beta = 1 - \psi_n(\xi_j)$ получим, что $S_n(x) = T_n(x) + \Delta_n(x)$, где

$$\Delta_n(x) = \frac{1}{n} \sum_{j=1}^{n} \gamma_j(x) \frac{(\psi_n(\xi_j) - \psi(\xi_j))^2}{1 - \psi_n(\xi_j)} \qquad (32)$$

и

$$\gamma_j(x) = \frac{w_j \cdot a(\xi_j) I\{\xi_j \le x\}}{(1 - \psi(\xi_j))^2}, \quad j = 1, 2, \ldots, n. \qquad (33)$$

Отметим, что $a(x) \equiv 1$ приводит к оценке $H_n(x)$ Нельсона (см. (8)), а нижеследующая теорема утверждает $1/n$-асимптотическую эквивалентность (30) и $T_n(x)$ из (18).

ТЕОРЕМА 3. При сделанных предположениях

$$M(n|\Delta_n(x)|)^k \le C_k(x) < \infty. \qquad (34)$$

Доказательство. Учитывая, что все компоненты суммы в $\Delta_n(x)$ одинаково распределены, получим

$$M|\Delta_n(x)|^k \le M|\gamma_1(x)|^k \cdot \frac{(\psi_n(\xi_1) - \psi(\xi_1))^{2k}}{|1 - \psi_n(\xi_1)|^k}. \qquad (35)$$

Далее, если условное математическое ожидание относительно (ξ_1, w_1) обозначать знаком "штрих", то правая часть (35) перепишется в виде

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$M\left[|\gamma_1(x)|^k M'\frac{|\psi_n(\xi_1)-\psi(\xi_1)|^{2k}}{|1-\psi_n(\xi_1)|^k}\right].$$

По (19)

$$|\psi_n(\xi_1)-\psi(\xi_1)| \le \frac{1}{n}\psi(\xi_1)+\frac{\sqrt{n-1}}{n}K_{n-1}, \qquad (36)$$

где

$$K_{n-1}=\sqrt{n-1}\sup_u\left|\frac{1}{n-1}\sum_{i=2}^{n} I\{\xi_i<u\}-\psi(u)\right|$$

представляет статистику Колмогорова для выборки $\xi_2,\dots,\xi_n$, не зависящую от (w_1,ξ_1). Следовательно,

$$(\sqrt{n}\,|\psi_n(\xi_1)-\psi(\xi_1)|)^{2k} \le (1+K_{n-1})^{2k},$$

а так как $n(1-\psi_n(\xi_1))=1+\sum_{i=2}^{n} I\{\xi_i\ge\xi_1\}$, то легко получаем неравенство

$$M|\Delta_n(x)|^k \le \left(M(1+K_{n-1})^{2k}\right)^{1/2} M\left[|\gamma_1(x)|^k\left(M'\frac{1}{\left(1+\sum_{i=2}^{n} I\{\xi_i\ge\xi_1\}\right)^{2k}}\right)^{\frac{1}{2}}\right]. \qquad (37)$$

Пусть $p_1=1-\psi(\xi_1)$, тогда при целом z:

$$M'\frac{1}{\left(1+\sum_{i=2}^{n} I\{\xi_i\ge\xi_1\}\right)^z}=\sum_{m=0}^{n-1} C_{n-1}^{m}\frac{p_1^m(1-p_1)^{n-m-1}}{m^z}\le$$

$$\le\frac{2^z}{p_1^z n^z}\cdot\sum_{m=0}^{n-1} C_{n+z-1}^{m+z}\,p_1^{m+z}(1-p_1)^{n-m-1}\le\frac{2^z}{p_1^z n^z}.$$

Отсюда, очевидно, следует, что

$$n^k M|\Delta_n(x)|^k \le B_k M\left|\frac{\gamma_1(x)}{1-\psi(\xi_1)}\right|^k=$$

$$=B_k M\,\frac{w_1 I\{\xi_1\le x\}\,|a(\xi_1)|^k}{|1-\psi(\xi_1)|^{k+2}}\le$$

$$\le B_k\int_{-\infty}^{x}\frac{|a(u)|^k(1-G_F(u))\,dF(u)}{|1-\psi(u)|^{k+2}}\le$$

$$\le\frac{B_k}{(1-\psi(x))^{k+2}}\int_{-\infty}^{x}|a(u)|^k\,dF(u)<\infty,$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

что и требовалось.

Из этой теоремы следует, что теорема 2 остается верной, если в ней заменить $T_n(x)$ на $S_n(x)$.

А теперь вернемся к

$$H_n(x) = \frac{1}{n}\sum_{j=1}^{n} \frac{w_j I\{\xi_j \le x\}}{1-\psi_n(\xi_j)}$$

Легко подсчитать, что ($p_1 = 1-\psi(\xi_1)$):

$$M H_n(x) = M \frac{w_1 I\{\xi_1 \le x\}}{1-\psi_n(\xi_1)} =$$

$$= M w_1 I\{\xi_1 \le x\} \sum_{m=0}^{n-1} C_{n-1}^{m} \frac{n p_1^m (1-p_1)^{n-m-1}}{m+1} =$$

$$= M w_1 I\{\xi_1 \le x\} \sum_{m=0}^{n-1} C_n^{m+1} \frac{p_1^{m+1}(1-p_1)^{n-(m+1)}}{p_1} =$$

$$= M w_1 I\{\xi_1 \le x\} \frac{1-(1-p_1)^n}{p_1} =$$

$$= M \frac{w_1 I\{\xi_1 \le x\}}{1-\psi(\xi_1)} - M \frac{w_1 I\{\xi_1 \le x\} (\psi(\xi_1))^n}{1-\psi(\xi_1)}.$$

Отсюда при $x \le z$

$$|M H_n(x) - H(x)| \le \frac{[\psi(z)]^n}{1-\psi(z)} \qquad (38)$$

Из (38) следует, что при учете любых конечных моментов в асимптотике по $n \to \infty$ смещенностью $H_n(x)$ относительно $H(x)$ можно пренебречь.

ТЕОРЕМА 4. Пусть $\psi(z) < 1$ и $\varepsilon > 0$ – любое; существует $B(z,\varepsilon) < \infty$ такое, что

$$P_F\{\sup_{x \le z} \sqrt{n}\,|H_n(x) - H(x)| < B(z,\varepsilon)\} \ge 1-\varepsilon \qquad (39)$$

для всех n одновременно.

Доказательство. Введем вспомогательную статистику

$$H_n^*(z) = \frac{1}{n}\sum_{j=1}^{n} w_j I\{\xi_j \le x\} \frac{1}{1-\psi(\xi_j)}.$$

Пусть $\delta_n(z)$ означает случайную величину под знаком вероятности в (39), а

$$\delta_n^*(z) = \sup_{x \le z} \sqrt{n}\,|H_n^*(x) - H(x)| \qquad (40)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Легко видеть, что

$$\delta_n(z) \le \delta_n^*(z) + \sup_{x\le z} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \frac{w_j I\{\xi_j \le x\}}{1-\psi_n(\xi_j)} \cdot \frac{|\psi_n(\xi_j)-\psi(\xi_j)|}{1-\psi(\xi_j)} \le$$

$$\le \delta_n^*(z) + \frac{H_n(z)}{1-\psi(z)} \cdot K_n , \qquad (41)$$

где $K_n = \sup_x \sqrt{n}\,|\psi_n(x)-\psi(x)|$ является статистикой Колмогорова для $\xi_1,\dots,\xi_n$.

Поскольку $H_n(z)$ и K_n имеют равномерно по n ограниченные вторые моменты, то для (39) достаточно доказать соответствующее неравенство для $\delta_n^*(z)$.

Пусть $\mu_n(t) = \sqrt{n}\,\left(H_n^*(F^{-1}(t)) - H(F^{-1}(t))\right)$, так что

$$\delta_n^*(z) = \sup_{0\le t\le \lambda(z)} |\mu_n(t)|,$$

где $\lambda(z) = F(z) < 1$, т.к. $1-\psi(z) = (1-F(z))(1-G_F(z)) > 0$.

Отметим следующее: так как $\theta_1\mu_n(t_1)+\dots+\theta_s\mu_n(t_s)$ при любых действительных $\theta_1,\dots,\theta_s$ и любых $t_1,\dots,t_s$ из $(0,\lambda(z))$, очевидно, имеет гауссово предельное распределение, то все конечномерные моменты $\mu_n(t)$ сходятся к моментам гауссовского процесса $\mu(t)$ с ковариационной функцией

$$\mathcal{E}(t,s) \equiv M\mu_n(t)\mu_n(s) = H_G(\min(s,t)) - \ln(1-t)\ln(1-s), \qquad (42)$$

где

$$H_G(v) = \int_0^v \frac{du}{(1-u)^2(1-G_F(F^{-1}(u)))} .$$

Доказательство (42) представляет прямой счет. Отсюда и из гауссовости $\mu(t)$ получим

$$M[\mu(t+\Delta)-\mu(t)]^4 = 3\left(H_G(t+\Delta)-H_G(t)-\ln^2\left(1-\frac{\Delta}{1-t}\right)\right)^2 \underset{\Delta\to 0}{\to} 0,$$

так что $\mu(t)$ имеет непрерывные с вероятностью I траектории. Следовательно, при некотором B_0 (при фиксации z и $\varepsilon>0$)

$$P_F\{\sup_{0\le t\le\lambda(z)} |\mu(t)| < B_0\} = 1-\varepsilon/2 .$$

Но по теореме 4 гл.IX §5 из книги [10]

$$\lim_{n\to\infty} P_F\{\delta_n^*(z) < B_0\} = 1-\varepsilon/2 ,$$

то есть, при $n \ge n_0(z,\varepsilon)$

$$P_F\{\delta_n^*(z) < B_0\} \ge 1-\varepsilon .$$

Но при $n \le n_0(z,\varepsilon)$ величина $\delta_n^*(z)$ равномерно ограничена неким

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

числом B_0' (например, $n_0^2(z,\varepsilon)(1-\psi(z))^{-1}$), так что (39) для $\delta_n^*(z)$ верно с $B(z,\varepsilon)=\max(B_0,B_0')$. Теорема доказана.

В прикладной статистике хорошо известны графические методы оценивания масштаба и сдвига. Если $\xi_j(n)$ – j-я порядковая статистика выборки с распределением из семейства

$$\mathcal{F}=\{F(ax+b);\ a>0,\ -\infty<b<\infty\}, \qquad (43)$$

то $\tau_j(n)=F(a\xi_j(n)+b)$ при истинных a и b является j-й порядковой статистикой равномерно на $[0,1]$ распределенной выборки размера n, так что (см. [11])

$$a\xi_j(n)+b=F^{-1}\left(\frac{j}{n+1}\right)+\frac{\sqrt{j(n-j+1)}\,\tau_j^*(n)}{(n+1)\sqrt{n+2}\,f\left(F^{-1}\left(\frac{j}{n+1}\right)\right)}, \qquad (44)$$

где $\tau_j^*(n)$ определяет выборочный процесс, сходящийся к броуновскому мосту;

$$S_F^{(n)}(j)=\frac{1}{f\left(F^{-1}\left(\frac{j}{n+1}\right)\right)}\cdot\sqrt{\frac{j(n-j+1)}{(n+1)^2(n+2)}},\quad f(u)=F'(u),$$

при широких условиях на $f(u)$ стремится к нулю для $\min(j,n-j)\to\infty$. Из (44) вытекает "метод": приближенное равенство $a\xi_j(n)+b\cong F^{-1}\left(\frac{j}{n+1}\right)$ определяет графические оценки путем проведения подходящей прямой через точки

$$Q_j=\left(\xi_j(n);\ F^{-1}\left(\frac{j}{n+1}\right)\right),\ j=1,\dots,n, \qquad (45)$$

на плоскости. Идея таких оценок особенно приятна тем, что способ оценивания фактически один и тот же при любых неслучайно цензурированных выборках, если только число точек Q_j, остающихся после цензурирования, не слишком мало.

Аналогичная идея для случайно цензурированных данных была предложена Джонсоном (см. [4]), часто используется практиками либо напрямую, либо с некоторыми модернизациями (например, в работе [7]). Ниже мы рассмотрим оценки такого типа.

Итак, рассматривается все та же случайно цензурированная выборка с

$$w_j=I\{\xi_j\le\eta_j\},\quad \xi_j=\min(\xi_j,\eta_j), \qquad (46)$$

причем распределение ξ_j имеет вид

$$F(x)=F_0(x^a\exp(b)), \qquad (47)$$

где $F_0(t)$ сосредоточено на $(0,\infty)$ и имеет там положительную

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

непрерывную производную.

Отметим, что множество точек Q_j в (45) совпадает с множеством точек $Q_j^* = (\xi_j, z_n(\xi_j))$, $j = 1, \dots, n$, где $z_n(u) = \frac{1}{n+1} \sum_{k=1}^{n} I\{\xi_k \le u\}$, так что, взяв точки

$$P_i = (\ln \xi_i, \ln F_0^{-1}(1 - \exp(-H_n(\xi_i)))) \qquad (48)$$

с $w_i = 1$, мы должны получить на плоскости нечто похожее на прямую $y = ax + b$, поскольку при $w_i = 1$ абсциссой P_i является $\ln \xi_i$, $1 - \exp(-H_n(\xi_i)) \approx z_n(\xi_i) = z_n'(\ln \xi_i)$, и мы имеем равенство

$$a \ln t + b = \ln F_0^{-1}(F(t)). \qquad (49)$$

Соответствующие оценки a_n и b_n построим, проводя через P_i с $w_i = 1$ прямую методом наименьших квадратов. Фактически предлагаются следующие оценки.

Пусть $0 < z_1 < z_2 < \infty$,

$$h_j = I\{z_1 \le \xi_j \le z_2\} \ln F_0^{-1}(1 - \exp(-H_n(\xi_j))),$$

$$t_j = I\{z_1 \le \xi_j \le z_2\} \ln \xi_j, \quad j = 1, \dots, n;$$

$$A_\ell = \frac{1}{n} \sum_{j=1}^{n} w_j t_j^\ell, \quad B_\ell = \frac{1}{n} \sum_{j=1}^{n} w_j t_j^\ell h_j, \quad \ell = 0, 1, 2.$$

Тогда

$$a_n = \frac{-A_1 B_0 + A_0 B_1}{A_0 A_2 - A_1^2};$$

$$b_n = \frac{-A_1 B_1 + A_2 B_0}{A_0 A_2 - A_1^2}. \qquad (50)$$

Рассмотрим вначале величину $\sigma_n^2 = A_0 A_2 - A_1^2$. Заметим, что

$$\sigma_n^2 = \frac{1}{2n^2} \sum_{i \ne j}' w_i w_j (t_i - t_j)^2. \qquad (51)$$

Действительно, так как

$$\left(\frac{1}{n} \sum_{j=1}^{n} w_j\right)\left(\frac{1}{n} \sum_{j=1}^{n}{}' w_j t_j^2\right) = \frac{1}{n^2} \sum_{i,j=1}^{n}{}' w_i w_j t_j^2 =$$

$$= \frac{1}{n^2} \sum_{i,j=1}^{n}{}' w_i w_j t_i^2 = \frac{1}{2n^2} \sum_{i,j=1}^{n}{}' w_i w_j (t_i^2 + t_j^2),$$

то

$$\sigma_n^2 = \left(\frac{1}{n} \sum_{j=1}^{n} w_j\right)\left(\frac{1}{n} \sum_{j=1}^{n}{}' w_j t_j^2\right) - \left(\frac{1}{n} \sum_{j=1}^{n} w_j t_j\right)^2 =$$

$$= \frac{1}{2n^2} \sum_{i,j=1}^{n}{}' w_i w_j (t_i^2 + t_j^2) - \frac{1}{2n^2} \sum_{i,j=1}^{n} 2 w_i w_j t_i t_j =$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$= \frac{1}{2n^2} \sum_{i,j=1}^{n}{}' w_i w_j (t_i^2 + t_j^2 - 2t_i t_j),$$

что и требуется, поскольку $t_i^2 + t_j^2 - 2t_i t_j = 0$ при $i = j$. Далее легко подсчитать, что

$$M w_i (t_i - \lambda)^2 = \int_{z_1}^{z_2} (x-\lambda)^2 (1 - G_L(x)) \, dF_L(x),$$

где F_L и G_L – распределения $\ln \xi_i$ и $\ln \eta_i$ соответственно. Таким образом, при $i \neq j$

$$M w_i w_j (t_i - t_j)^2 = M\left[w_j M\left(w_i (t_i - t_j)^2 / w_j, \xi_j\right)\right] =$$

$$= \int_{z_1}^{z_2} (1 - G_L(x)) M w_j (t_j - x)^2 \, dF_L(x) =$$

$$= \int_{z_1}^{z_2} \int_{z_1}^{z_2} (1 - G_L(x))(1 - G_L(y))(x-y)^2 \, dF_L(x) \, dF_L(y) \geq$$

$$\geq \delta^2(z_1) \int_{z_1}^{z_2} \int_{z_1}^{z_2} (x-y)^2 \, dF_L(x) \, dF_L(y) =$$

$$= a^{-2} \delta^2(z_1) \int_{d_1}^{d_2} \int_{d_1}^{d_2} \left(\ln \frac{u}{v}\right)^2 f_0(u) f_0(v) \, du \, dv > 0,$$

где

$$\delta(z_2) = 1 - G_F(\exp(a z_2 + b)) > 0, \; d_\ell = \exp(a z_\ell + b), \; \ell = 1, 2.$$

Поскольку σ_n^2 представляет U-статистику в силу (51) и $M w_i t_i^4 < \infty$ (очевидно, из-за выбора $z_2 < \infty$), то σ_n^2 асимптотически нормальна и сходится к положительной величине. Далее, если $H(x) = -\ln(1 - F(x))$, где $F(x)$ – истинное распределение ξ_j-тых, то при истинных a и b и при

$$h_j^* = I\{\xi_j \in [z_1; z_2]\} \ln F_0^{-1}(1 - \exp(-H(\xi_j))), \qquad (52)$$

$j = 1, \ldots, n$, мы с соответственно определенными B_0^* и B_1^* (заменой h_j на h_j^*) получим точные равенства:

$$a = \frac{-A_1 B_0^* + A_0 B_1^*}{A_0 A_2 - A_1^2}; \quad b = \frac{-A_1 B_1^* + A_2 B_0^*}{A_0 A_2 - A_1^2}. \qquad (53)$$

Отсюда

$$a_n - a = \frac{A_0 (B_1 - B_1^*) - A_1 (B_0 - B_0^*)}{A_0 A_2 - A_1^2} \qquad (54)$$

и аналогично для $b_n - b$. С учетом уже проведенного анализа очевидно что все трудности доказательства асимптотической нормальности вектора

$$\Delta_n = \sqrt{n} \begin{pmatrix} a_n - a \\ b_n - b \end{pmatrix}$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

сосредоточены фактически в доказательстве асимптотической нормальности величины

$$\beta_n = \sqrt{n}\,(B_0 - B_0^*), \qquad (55)$$

остальное же составляет непринципиальные аналитические выкладки, громоздкие, но в целом обычные. Для последнего замечания достаточно заметить, что структура величины

$$\sqrt{n}\,\sigma_n^{-2}\,[\theta_1(a_n - a) + \theta_2(b_n - b)]$$

с произвольными θ_1 и θ_2 аналогична структуре β_n; $\sigma_n^2 \to \sigma_0^2 > 0$.

ТЕОРЕМА 5. Если $F_0(x)$ имеет на $(0,\infty)$ положительную непрерывную производную $f_0(x)$ и существует $\alpha > 0$ такое, что

$$\lim_{x\to\infty} x^{\alpha}(1 - F_0(x)) = 0, \qquad (56)$$

то вектор Δ_n асимптотически нормален с нулевым средним и конечными дисперсиями.

Доказательство. Итак, как было обсуждено, достаточно обосновать асимптотическую нормальность β_n. Пусть

$$V(t) \equiv \ln F_0^{-1}(1 - \exp(-t)), \qquad (57)$$

тогда

$$\beta_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} w_j I\{\xi_j \in [z_1; z_2]\}\left(V(H_n(\xi_j)) - V(H(\xi_j))\right). \qquad (58)$$

Наряду с β_n рассмотрим аппроксимирующую статистику

$$\beta_n^* = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} w_j I\{\xi_j \in [z_1; z_2]\}\left(H_n(\xi_j) - H(\xi_j)\right) V'(H(\xi_j)). \qquad (59)$$

Поскольку имеет место легко проверяемое тождество

$$H_n(\xi_j) - H(\xi_j) = [T_n(\xi_j) - H(\xi_j)] + \frac{1}{n}\sum_{k=1}^{n} \frac{w_k I\{\xi_k \le \xi_j\}}{1 - \psi_n(\xi_k)}\left[\frac{\psi_n(\xi_k) - \psi(\xi_k)}{1 - \psi(\xi_k)}\right]^2, \qquad (60)$$

где $T_n(u)$ взято из (18) с $a(u) \equiv 1$, то асимптотическая нормальность β_n^* следует из двух моментов. Во-первых,

$$d_n = \sup_{z\in[z_1;z_2]} \sqrt{n}\left|\frac{\psi_n(z) - \psi(z)}{1 - \psi(z)}\right|$$

имеет, очевидно, все моменты, так что второй член в (60) оценивается по модулю величиной $n^{-1} d_n H_n(\xi_j)$ и тем самым асимптотически пренебрежим. Первый же член после замены им в β_n^*

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

величин $H_n(\xi_j)-H(\xi_j)$ дает U-статистику 3-го порядка, т.е. вида

$$\frac{1}{n^3}\sum_{i_1,i_2,i_3=1}^{n}\varphi(\nu_{i_1},\nu_{i_2},\nu_{i_3}),$$

от независимых случайных величин $\nu_i=(w_i,\xi_i)$ с конечными моментами у $\varphi(\dots)$ второго порядка, причем $M(\tau_n(\xi_j)-H(\xi_j))=O(\frac{1}{n})$, так что β_n^* асимптотически нормально с нулевым средним.

Рассмотрим теперь отдельный элемент в сумме $\beta_n-\beta_n^*=$ $=\frac{1}{n}\sum_{j=1}^{n}\tau_j^{(n)}$, где

$$\tau_j^{(n)}=\sqrt{n}\,w_j\,I\{\xi_j\in[z_1;z_2]\}\Big[V(H_n(\xi_j))-V(H(\xi_j))-V'(H(\xi_j))(H_n(\xi_j)-H(\xi_j))\Big]. \quad (62)$$

Поскольку любые наборы $(\tau_{j_1}^{(n)},\dots,\tau_{j_s}^{(n)})$ при различных $j_1,\dots,j_s$ и любом фиксированном s одинаково распределены, то утверждение

$$M|\beta_n-\beta_n^*|\to 0 \qquad \text{при} \quad n\to\infty$$

следует, очевидно, из того, что $M|\tau_1^{(n)}|\to 0$; для обоснования последнего достаточно показать, что $\tau_1^{(n)}\to 0$ по вероятности и что

$$M|\tau_1^{(n)}|^2\le C<\infty. \quad (63)$$

Отметим, что стремление $\tau_1^{(n)}$ к нулю по вероятности является очевидным следствием того, что $f_0(t)$ равномерно непрерывна на $[z_1,z_2]$ и что

$$\delta_n=\sup_{z_1\le z\le z_2}\sqrt{n}\,|H_n(z)-H(z)| \quad (64)$$

равномерно ограничено по вероятности в силу теоремы 4.

А теперь перейдем к доказательству утверждения (63). По условию теоремы

$$F_0(x)\ge 1-\frac{1}{x^\alpha}\ge 1-\exp\left(-(\ln x)^{1/2}\right)$$

при $x\ge x_\alpha$. Отсюда

$$x\ge F_0^{-1}\left(1-\exp\left(-(\ln x)^{1/2}\right)\right),$$

что после замены $t=\sqrt{\ln x}$ и логарифмирования приводит к неравенству

$$V(t)\le t^2 \quad (65)$$

при $t\ge t_\alpha$, $t_\alpha<\infty$.

Поскольку $n(1-\Psi_n(z_1))\ge 1+\sum_{i=2}^{n}I\{\xi_i\ge z_2\}$, то для $\theta<\frac{1}{2}(1-\Psi(z_2))$ получим

$$P\{H_n(\xi_j)\ge\tfrac{1}{\theta}\}\le P\{1+\sum_{i=2}^{n}I\{\xi_i\ge z_2\}\le\theta n\}\le\frac{C_s}{n^s} \quad (66)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

при любом $S<\infty$.

Далее $H_n(S_j)\le n$ всегда. Это вместе с (65) показывает, что вклад значений $H_n(S_1)\ge\frac{1}{\theta}$ во второй момент $\tau_1^{(n)}$ стремится к нулю при $n\to\infty$. Далее

$$H_n(S_1)=\sum_{k=1}^{n}\frac{w_k I\{S_k\le S_1\}}{1+\sum_{i\ne k} I\{S_i\ge S_k\}}\ge\sum_{k=1}^{n} w_k I\{S_k\le S_1\},$$

так что $H_n(S_1)\ge 1$, если только $S_1\in[z_1,z_2]$ и $w_1=1$, т.е. при $|\gamma_{n1}|>0$ обязательно $H_n(S_1)\ge 1$. Но на отрезке $[1;\frac{1}{\theta}]$ функции, входящие в γ_{n1} , равномерно непрерывны и тем самым ограничены, что и завершает доказательство.

В конце статьи приведем критерий проверки на тип распределения. В принятых выше обозначениях рассмотрим величину

$$q_n(z_1,z_2)=\frac{1}{n}\sum_{j=1}^{n} w_j(a_n t_j+b_n-h_j)^2, \qquad (67)$$

вычисленную в предположении, что формирующим логарифмический тип распределением является $F_0(x)$. Пусть

$$\tilde h_j=I\{S_j\in[z_1,z_2]\}\ln\tilde F_0^{-1}(1-\exp(-H_n(S_j)))$$

и $\tilde q_n(z_1,z_2)$ вычисляется по (67) с заменой h_j на $\tilde h_j$ и a_n, b_n, соответствующими $\tilde a_n$, $\tilde b_n$. Тогда отношение

$$\mathcal{K}_n=\frac{\tilde q_n(z_1,z_2)}{q_n(z_1,z_2)}$$

является состоятельным критерием для проверки гипотез о типе распределения.

$$\lim_{n\to\infty} P_{F_0}\{\mathcal{K}_n\ge B\}=1$$

и

$$\lim_{n\to\infty} P_{\tilde F_0}\{\mathcal{K}_n\le\frac{1}{B}\}=1$$

для любого $B<\infty$, если только имеется положительная вероятность того, что цензурирующая случайная величина лежит правее z_2 . Доказательство этого опускаем, оно, очевидно, в основных чертах повторяет доказательство теоремы 5 . Добавим лишь, что опыт практической работы показывает, что при объемах выборок в 50 ÷ 200 и величинах $\frac{1}{n}\sum_{j=1}^{n} w_j>0{,}1$ разумной является рекомендация:

при $\mathcal{K}_n\ge 2$ принимается F_0 ,

при $\mathcal{K}_n\le\frac{1}{2}$ принимается $\tilde F_0$,

а в остальных случаях решение не принимается до получения новой информации.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Литература

I. Kaplan E.L., Meier P. Nonparametric estimation from incomplete observations. - J. Amer. Stat. Ass.. I958, v.53, p. 457-48I.

2. Gill R.D. Censoring and stochastic integrals: Mathemam. centrum, I980. - I72 p.

3. Csërgö S. Estimation in, and testing for the proportional hazard model of random censorship. - In: Proc. 7-th International Summer School on Problem of Model Choice in regression analysis. - Berlin: Humboldt Universität, I986, p. I7-24.

4. Jonson L.G. The statistical treatment of fatigue experiments. - Amsterdam, Elsevier Publishing Company, I964. - I06 p.

5. Скрипник В.М., Назин А.Е. Оценка надежности технических систем по цензурированным выборкам. - Минск: Наука и Техн., I98I.

6. Агеев В.В., Благовещенский Ю.Н. Анализ оценок функции распределения по случайно цензурированной выборке. - В кн: Статист. мет. обработки результатов наблюдений при контроле качества и над)жн. машин и приборов. Л.: ЛДНТП, I979, с.26-3I.

7. Nelson W. Hazard plotting for incomlete failure data. - Jour. of Quality Technology, I969, v.I, N I, p. 524-53I.

8. Благовещенский Ю.Н. Об асимптотической нормальности одного класса статистик для случайно цензурированных справа выборок. - Теор. вероятн. и ее примен., I979, т.24, № 3, с.660-662.

9. Hoeffding W. A class of statistics with asymptotically normal distribution. - Ann. Math. Statist., I948, v. I9, p. 293-325.

I0. Гихман И.И., Скороход А.В. Введение в теорию случайных процессов. - М.: Наука, I965. - 655 с.

II. Благовещенский Ю.Н. О распределении крайних членов вариационного ряда. - ДАН СССР, т. I88, № 4, с. 735-738.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ОБ АСИМПТОТИЧЕСКОЙ ОПТИМАЛЬНОСТИ ДВУХЭТАПНОЙ ПРОЦЕДУРЫ ОЦЕНИВАНИЯ

В.З.Борисов

В данной заметке на примере оценивания среднего последовательности независимых случайных величин показана асимптотическая оптимальность двухэтапной последовательной процедуры оценивания параметров линейной регрессии с неизвестной дисперсией и произвольным распределением помех [I]. В случае гауссовского распределения помех аналогичный результат был получен в [2].

Итак, рассмотрим последовательность x_t, определяемую соотношениями

$$x_t = \lambda + \sigma \xi_t , \quad t = 1, 2, \ldots , \qquad (I)$$

где $\{\xi_t\}$ – последовательность независимых случайных величин с $M\xi_t = 0$, $M\xi_t^2 = 1$ и плотностью распределения $f(\cdot)$; параметры λ и σ неизвестны. Как известно [3], в данном случае при фиксированном объеме выборки не существует процедуры оценивания параметра λ, обеспечивающей априори заданные ограничения на ковариацию оценки. Последовательный алгоритм, предложенный в [I] позволяет в отличие от планов оценивания, основанных на фиксированной выборке, обоснованно определять объем наблюдений, необходимый для оценивания неизвестного параметра λ с ковариацией, не превышающей заданный порог точности.

Применительно к модели (I) последовательный алгоритм выглядит следующим образом. Выбирая произвольные постоянные t_1 и H (определяющие объем выборки на первом этапе оценивания и порог для момента остановки), вычислим момент прекращения наблюдений по формуле

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\tau = \inf\{t > t_1 : (t-t_1)/\hat{\sigma}^2(t_1) \geq H b^2(t_1)\}, \qquad (2)$$

где $\hat{\sigma}^2(t_1) = \sum_{s=1}^{t_1} (x_s - \bar{x})^2/(t_1-1)$, $\bar{x} = \sum_{s=1}^{t_1} x_s$,

$$b^2(t_1) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} 1/\left(\sum_{s=1}^{t_1-1} y_s^2\right) f(y_1) \dots f(y_{t_1-1})\, dy_1 \dots dy_{t_1-1} .$$

С учетом (2) последовательная оценка параметра λ определяется как:

$$\lambda_\tau^* = \frac{1}{\tau - t_1 + 1} \sum_{s=t_1+1}^{\tau} x_s . \qquad (3)$$

Двухэтапный последовательный план (τ, λ_τ^*) для всех λ, σ обладает следующими свойствами [I]:

$$M_{\lambda,\sigma}\, \lambda_\tau^* = \lambda ; \qquad (4)$$

$$M_{\lambda,\sigma}(\lambda_\tau^* - \lambda)^2 \leq 1/H ; \qquad (5)$$

$$P_\sigma(\tau < \infty) = 1 .$$

Определим теперь понятие оптимальности последовательной процедуры оценивания. Для этого отметим, прежде всего, что если дисперсия σ^2 последовательности случайных величин $\{x_t\}$ известна, то для выполнения условий(4),(5)метод наименьших квадратов, являющийся оптимальным в классе линейных несмещенных планов оценивания, требует $n_0(H,\sigma)$ наблюдений

$$n_0(H,\sigma) = \min\{n \geq 1 : n/\sigma^2 \geq H\} .$$

В том случае, когда σ^2 неизвестна, будем говорить, что последовательный план (τ, λ_τ^*) является асимптотически оптимальным, если

$$\lim_{H\to\infty} \frac{n_0(H,\sigma)}{M_{\lambda,\sigma}\,\tau} \geq 1 \qquad \text{для всех } \lambda, \sigma .$$

Оказывается, что если объем выборки t_1 на первом этапе оценивания сделать зависимым от порога H , то двухэтапная процедура может быть асимптотически оптимальной в смысле данного выше определения. Сформулируем этот результат в виде теоремы.

Т е о р е м а . Пусть объем выборки t_1 на первом этапе оценивания зависит от величины H и удовлетворяет условиям:

$$\lim_{H\to\infty} t_1(H) = \infty ; \qquad (6)$$

$$\lim_{H\to\infty} t_1(H)/H = 0 . \qquad (7)$$

Тогда, если функция плотности $f(\cdot)$ непрерывна или ограничена в окрестности точки 0, момент остановки τ , определяемый по формуле (2), не зависит от оцениваемого параметра λ и обладает следующими свойствами:

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$M_\sigma \tau \le H B^2(t_1)\sigma^2 + t_1 + 1 \qquad \text{для всех } \sigma, H;$$

$$\lim_{H\to\infty} \frac{\tau(H)}{n_0(H,\sigma)} = 1 \qquad \text{для всех } \sigma \ (P\text{-п.н.});$$

$$\lim_{H\to\infty} \frac{M_\sigma \tau(H)}{n_0(H,\sigma)} = 1 \qquad \text{для всех } \sigma.$$

Соотношения (6,7) определяют достаточные условия для асимптотической оптимальности двухэтапной процедуры оценивания, определяемой формулами (2),(3). Для того, чтобы эти условия были выполнены, достаточно положить, например, $t = \sqrt{H}$.

Литература

I. Борисов В.З. Последовательное оценивание параметров случайных процессов рекуррентного типа при неизвестной дисперсии помех. – В кн.: БНТИ. Методы и алгоритмы прикладной математики. Якутск: Изд-во ЯФ СО АН СССР, 1982, с. 7–13.

2. Борисов В.З. Об асимптотической эффективности последовательной процедуры Стейна. – В кн.: Методы и алгоритмы прикладной математики в задачах теплофизики и обработки эксперимента. Якутск: Изд-во ЯФ СО АН СССР, 1983, с. 100–105.

3. Закс Ш. Теория статистических выводов. – М.: Мир, 1975, 776 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ПОИСК СТАТИСТИЧЕСКИХ ЗАКОНОМЕРНОСТЕЙ В ОБУЧАЮЩИХ МАССИВАХ ДАННЫХ

Т.В.Вирядченко, О.Ю.Кульчицкий

I. Введение

В задачах статистической классификации, которые возникают в самых разнообразных приложениях, часто приходится сталкиваться с фактом неустойчивости результатов классификации на контрольных массивах данных: результаты, полученные на обучающих массивах, существенно отличаются от результатов, полученных на контрольных массивах. Авторам, в частности, известны примеры такой неустойчивости результатов классификации при решении задач медицинской диагностики по электрокардиографическим данным. Причиной неустойчивости результатов часто оказывается неверное вероятностное описание классов, сделанное на основании обучающих выборок. Обычно стараются сделать какие-то простые предположения о видах распределения вероятностей элементов обучающих выборок, однако такие простые описания могут оказаться неадекватными исследуемым массивам данных, возможно даже, что исследуемые данные вообще нельзя смоделировать как выборку из некоторой генеральной совокупности с вероятностной мерой[I]

Приведем пример. Пусть имеется массив электрокардиографических данных $\{x^k\}_1^N$, $x \in R_n$, снятых со здоровых людей. Как задать плотность распределения x в классе "здоровых"? Как показали исследования [2,3], простые гауссовские гипотезы приводят к неустойчивости результатов классификации, если в контрольных выборках среди ЭКГ здоровых окажутся ЭКГ спортсменов. Поэтому, целесообразно предположить, что выборка имеет более сложную структуру, именно формируется как смесь случайных величин с простыми распределениями. Распределение $x - p(x)$ складывается из нескольких исходных распределений $p_i(x)$:

$$p(x) = p_\alpha(x) = \sum_i \alpha^{(i)} p_i(x), \quad \sum_i \alpha^{(i)} = 1; \quad 0 \le \alpha^{(i)} \le 1. \qquad (I)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Гипотезу (I) можно принять или отвергнуть, исследуя массивы исходных данных. Если к тому же такое исследование позволит определить исходные распределения $p_i(x)$, то устойчивое решение задач статистической классификации будет в значительной степени обеспечено.

Задача настоящей работы - предложение способов проверки гипотезы (I) и оценки исходных распределений в условиях, когда никаких параметрических предположений о распределениях не делается.

2. Постановка задачи

Пусть $\hat{p}_{\alpha_i}(x)$, $i=\overline{1,M+1}$ - парзеновские оценки плотностей вероятностей [4]:

$$\hat{p}_{\alpha_i}(x)=\frac{1}{N_i\delta_i}\sum_{k=1}^{N_i}h\left(\frac{x-x_{\alpha_i}^{k}}{\delta_i}\right), \qquad (2)$$

построенные по заданным выборкам $\left\{x_{\alpha_i}^{k}\right\}_1^{N_i}$, $i=\overline{1,M+1}$, значений случайного вектора $x\in R_n$.

Требуется проверить гипотезу о представимости плотностей $p_{\alpha_i}(x)$ через $L\le M$ исходных, т.е. узнать, верно ли, что существуют парциальные плотности $p_j(x)$, $j\in K_L$, такие, что

$$p_{\alpha_i}(x)=\sum_{j\in K_L}\alpha_i^{(j)}p_j(x). \qquad (3)$$

Здесь K_L - множество индексов L из M парциальных плотностей.

Требуется также построить оценки исходных парциальных плотностей $p_i(x)$, $i\in K_L$.

3. Проверка гипотезы смеси

Если соотношение (3) справедливо, то существуют такие числа γ_i, $\sum\gamma_i=0$, $\sum\gamma_i^2=1$, что при любом значении x

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\sum \gamma_i p_{\alpha_i}(x) = 0 \qquad (4)$$

или в интегральной форме

$$J_l = \int \left(\sum \gamma_i p_{\alpha_i}(x)\right)^2 dx = 0 . \qquad (5)$$

Суммирование выполняется по каждому набору $l+1$ плотностей $p_{\alpha_i}(x)$. Всего таких наборов C_M^{l+1}.

Равенства (5) для функционала J_l примем в качестве проверяемой гипотезы, считая ее эквивалентной гипотезе (3). Точные значения плотностей $p_{\alpha_i}(x)$ в (5) неизвестны. Поэтому для проверки гипотезы (5) воспользуемся статистикой

$$\hat{J}_l = \int \left(\sum \gamma_i \hat{p}_{\alpha_i}(x)\right)^2 dx , \qquad (6)$$

где $\hat{p}_{\alpha_i}(x)$ вычисляются по формуле (2).

Используя свойства парзеновских оценок, можно показать, что при некоторых условиях на функции $h(\cdot)$ и $p(\cdot)$ при

$$N_i \to \infty , \quad \delta_i = \delta(N_i) \to 0 , \quad \frac{N_i (\delta(N_i))^n}{\log \log N_i} \to \infty \qquad (7)$$

статистика $\hat{J}_l \xrightarrow{\text{п.н.}} J_l$.

Условия (7) являются достаточными для поточечной сходимости парзеновских оценок плотностей [5].

Таким образом, при больших N_i, $\hat{J}_l \simeq J_l$ и для проверки гипотезы (5) можно использовать следующий критерий. Гипотеза (5) принимается при $l = l^*, 1 \le l^* \le M$, если $\frac{\max \hat{J}_{l^*}}{\max \hat{J}_{l^*-1}} \ll 1$. Знак $\max$ означает вычисление максимального значения из значений $\hat{J}_{l^*}$ при всех наборах l^*+1 плотностей $p_{\alpha_i}(x)$ в (6).

Значения коэффициентов γ_i для $\hat{J}_l$ определяются из условия минимума по γ_i $\hat{J}_l$ при ограничениях $\sum \gamma_i = 0$, $\sum \gamma_i^2 = 1$. Вопрос выбора конкретного порога для отношения $\hat{J}_{l^*} / \hat{J}_{l^*-1}$ связан, по-видимому, с объемом используемых выборок и требуемой точностью проверки гипотезы и должен

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

решаться экспериментально в каждой конкретной задаче. После определения γ_i значения $\hat{J}_L$ могут быть вычислены по формуле

$$\hat{J}_L = \int \Big(\sum_{i=1}^{N} k(x - x_i^i, i) \Big)^2 dx = \sum_{i=1}^{N} \sum_{j=1}^{N} k(x - x_i^i, i) k(x - x_j^j, j) dx \quad , \qquad (8)$$

где

$$k(x - x_i^i, i) = \begin{cases} \dfrac{\gamma_1}{N_1 \delta_1} h\left(\dfrac{x - x_{\alpha_1}^i}{\delta_1} \right), & i = \overline{1, N_1} \ ; \\ \dfrac{\gamma_2}{N_2 \delta_2} h\left(\dfrac{x - x_{\alpha_2}^{i - N_1}}{\delta_2} \right), & i = \overline{N_1 + 1, N_2 + N_1} \ ; \\ \vdots \\ \dfrac{\gamma_{L+1}}{N_{L+1} \delta_{L+1}} h\left(\dfrac{x - x_{\alpha_{L+1}}^{i - \sum_{k=1}^{L} N_k}}{\delta_{L+1}} \right), & i = \overline{\sum_{k=1}^{L} N_k + 1, N = \sum_{k=1}^{L+1} N_k} \ . \end{cases}$$

В (8) для определенности взят набор из первых $L+1$ оценок плотностей $\hat{p}_{\alpha_i}(x)$

4. Проверка гипотезы смеси на малых выборках

Изложенный в предыдущем разделе способ проверки гипотезы смеси основывался на предположении больших N_i . В практических задачах классификации это предположение часто оказывается неверным, и тогда выбор параметров δ_i в соответствии с формулой (7) не является эффективным. В данном разделе предложен способ квазиоптимального выбора параметров δ_i из соображений наибольшей возможной близости функционалов J_L и $\hat{J}_L$.

По неравенству треугольника имеем

$$\sqrt{J_L} - \sqrt{\hat{J}_L} \le \sum |\gamma_i| \sqrt{\int (p_{\alpha_i}(x) - \hat{p}_{\alpha_i}(x))^2 dx} \quad . \qquad (9)$$

Минимальное значение (9) обеспечивается, если при каждом α_i обеспечить минимум интегральных ошибок оценок плотностей $\Delta p_{\alpha_i}(x) = p_{\alpha_i}(x) - \hat{p}_{\alpha_i}(x)$. Таким образом, оптимальное значение δ_i^* для каждой оценки плотности выбираем из условия .

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\min_{\delta_i} M\int(p_{\alpha_i}(x)-\hat{p}_{\alpha_i}(x))^2dx=\min_{\delta_i} M_{\{x_{\alpha_i}^k\}_1^{N_i}}\int(p_{\alpha_i}(x)-\frac{1}{N_i\delta_i^n}\sum_{k=1}^{N_i}h\left(\frac{x-x_{\alpha_i}^k}{\delta_i}\right))^2dx .$$

Индексы α_i , i далее опустим

Производя вычисления, получим

$$M\int(p(x)-\hat{p}(x))^2dx=\int p^2(x)dx+\frac{c}{N\delta}-$$

$$-\frac{1}{\delta}\iint\left[2h\left(\frac{x-y}{\delta}\right)-\frac{N-1}{N\delta}k_1\left(\frac{x-y}{\delta}\right)\right]p(x)p(y)dxdy , \qquad (10)$$

где $c=\int h^2(x)dx$, $k_1\left(\frac{x-y}{\delta}\right)=\int h\left(\frac{u-x}{\delta}\right)h\left(\frac{u-y}{\delta}\right)du$.

Заменяя математические ожидания функций их несмещенными оценками и учитывая, что первое слагаемое в (10) не зависит от δ , получим следующее правило выбора значения δ^*:

$$\delta^*=\operatorname{Arg}\min_{\delta}\left\{\frac{c}{N\delta}-\frac{1}{\delta(N-1)N}\sum_{i=1}^{N}\sum_{\substack{j=1\\ i\neq j}}^{N}\left[2h\left(\frac{x^i-x^j}{\delta}\right)-\right.\right.$$

$$\left.\left.-\left(1-\frac{1}{N}\right)\cdot\frac{1}{\delta}k_1\left(\frac{x^i-x^j}{\delta}\right)\right]\right\} . \qquad (11)$$

Вычислительная процедура нахождения δ^* в соответствии с (11) хотя принципиально и не вызывает каких-либо трудностей, но является достаточно трудоемкой. Поэтому хорошее начальное приближение может оказаться чрезвычайно полезным при практическом нахождении δ^* . Оно может быть рассчитано следующим образом. Например, при

$$h\left(\frac{x-y}{\delta}\right)=\frac{1}{(\sqrt{2\pi})^n}e^{-\frac{\|x-y\|^2}{2\delta^2}} \qquad (12)$$

предположим, что $p_\alpha(x)$ является гауссовской плотностью. Тогда интегралы в (10) вычисляются аналитически. Ввиду громоздкости запишем результат для случая $x\in R_1$:

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$M\int(p(x)-\hat{p}(x))^2dx-\int p^2(x)dx=\frac{1}{2\sqrt{\pi}\sigma}\left(\frac{1}{N\Delta}-\frac{2\sqrt{2}}{\sqrt{2+\Delta^2}}+\frac{N-1}{N\sqrt{1+\Delta^2}}\right),$$

где $\Delta=\frac{\delta}{\sigma}$ и σ^2 – дисперсия x.

Обозначим

$$\Delta_N^*=\operatorname*{Argmin}_{\Delta}\left(\frac{1}{N\Delta}-\frac{2\sqrt{2}}{\sqrt{2+\Delta^2}}+\frac{N-1}{N\sqrt{1+\Delta^2}}\right). \qquad (13)$$

Величину

$$\delta_N^0=\Delta_N^*\hat{\sigma},$$

где Δ_N^* определяется соотношением (13), а $\hat{\sigma}$ – оценка дисперсии x, построенная по выборке $\{x^i\}_1^N$, можно принять в качестве начального приближения для δ^*

5. Оценка парциальных плотностей двухкомпонентных смесей

Пусть $M=2$, т.е. представление (I) имеет вид

$$p_\alpha(x)=\alpha p_1(x)+(1-\alpha)p_2(x).$$

Известны оценки $\hat{p}_{\alpha_i}(x)$, $i=\overline{1,3}$. Задача состоит в том, чтобы оценить плотности $p_1(x)$ и $p_2(x)$.

Очевидно, что парциальные плотности допускают представление вида

$$p_i(x)=\lambda_i p_{\alpha_1}(x)+(1-\lambda_i)p_{\alpha_2}(x),\quad i=1,2,$$

в котором значения λ_i не обязательно принадлежат промежутку $[0,1]$, в частности, возможно $\lambda_i<0$, но обязательно таково, что

$$\lambda_i p_{\alpha_1}(x)+(1-\lambda_i)p_{\alpha_2}(x)\geq 0.$$

Иначе

$$\lambda_i\left(p_{\alpha_2}(x)-p_{\alpha_1}(x)\right)\leq p_{\alpha_2}(x).$$

Следовательно,

$$\lambda_i\begin{cases}\leq \dfrac{p_{\alpha_2}(x)}{p_{\alpha_2}(x)-p_{\alpha_1}(x)}, & \text{если } p_{\alpha_2}(x)>p_{\alpha_1}(x);\\ \geq \dfrac{p_{\alpha_2}(x)}{p_{\alpha_2}(x)-p_{\alpha_1}(x)}, & \text{если } p_{\alpha_2}(x)<p_{\alpha_1}(x).\end{cases}$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Для определенности примем

$$\lambda_1 = \min \frac{p_{\alpha_2}(x)}{p_{\alpha_2}(x) - p_{\alpha_1}(x)} \quad \text{при} \quad p_{\alpha_2}(x) > p_{\alpha_1}(x);$$

$$\lambda_2 = \max \frac{p_{\alpha_2}(x)}{p_{\alpha_2}(x) - p_{\alpha_1}(x)} \quad \text{при} \quad p_{\alpha_2}(x) < p_{\alpha_1}(x).$$

Заменяя в (I4) плотности $p_{\alpha_1}(x), p_{\alpha_2}(x)$ на их оценки, получим оценки $\hat{\lambda}_1$, $\hat{\lambda}_2$ и оценки исходных парциальных плотностей

$$\hat{p}_i(x) = \hat{\lambda}_i \hat{p}_{\alpha_1}(x) - (1-\hat{\lambda}_i)\hat{p}_{\alpha_2}(x), \; i = 1,2. \qquad (I5)$$

В силу соотношения (3) результат (I4), (I5) не изменится, если исходное представление сделано через другие плотности ($p_{\alpha_1}(x), p_{\alpha_3}(x)$ или $p_{\alpha_2}(x), p_{\alpha_3}(x)$). При этом точность оценок (I5) определяется точностью заданных оценок плотностей $\hat{p}_{\alpha_i}(x)$.

6. Результаты математического моделирования

Парзеновские оценки плотностей $\hat{p}_{\alpha_i}(x)$, $i = \overline{1,3}$, с весовой функцией $h = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x - x^k_{\alpha_i})^2}{2(0.06)^2}}$ вычислялись по выборкам $\{x^k_{\alpha_i}\}_1^{100}$, $i = \overline{1,3}$, являющимся смесями двух случайных величин с гауссовскими распределениями вероятностей $p_1(x) \sim N(0.3, 0.15)$, $p_2(x) \sim N(-0.3, 0.15)$. Значения коэффициентов смеси в обучающих выборках были $\alpha_1 = 0,2$; $\alpha_2 = 0,7$; $\alpha_3 = 0,4$. Для иллюстрации предложенного способа проверки гипотезы (3) на рис.I приведены оценки плотностей распределения вероятностей и оценки функций $\hat{z}(x) = \sum_{i=1}^{3} \gamma_i \hat{p}_{\alpha_i}(x)$; $\gamma_1 = 0,487$; $\gamma_2 = 0,329$; $\gamma_3 = -0,811$; $\hat{z}(x) = \frac{1}{\sqrt{2}}(\hat{p}_{\alpha_2}(x) - \hat{p}_{\alpha_1}(x))$, построенные для проверки соответственно гипотезы двухкомпонентной смеси и гипотезы отсутствия смеси. Как видно, в случае гипотезы смеси, которая адекватна модели, функция $\hat{z}(x)$ близка к нулю, а в случае гипотезы "нет смеси" функ-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ция $\hat{\mathcal{R}}(x)$ имеет вид, существенно от нуля отличающийся, что приводит к соответственно большему значению $\hat{\lambda}_1$.

Расчет оптимального значения параметра сглаживания δ^* оценок плотностей. Расчет оптимального значения δ^* производился по формуле (II) в случае $x \in R_1$ и функций $h(\cdot)$ вида (I2). Тогда $k_1\left(\frac{x^i - x^j}{\delta}\right)$ в (II) есть

$$k_1\left(\frac{x^i - x^j}{\delta}\right) = \frac{\delta}{2\sqrt{\pi}}\, e^{-\left(\frac{x^i - x^j}{4\delta^2}\right)^2}$$

и формула (II) с точностью до множителя, не зависящего от δ, принимает вид

$$\delta^* = \operatorname{Argmin}\left\{\frac{1}{\delta}\left[1 - \frac{1}{N-1}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ i\neq j}}^{N}\left(2\sqrt{2}\, e^{-\frac{(x^i - x^j)^2}{2\delta^2}} - \left(1 - \frac{1}{N}\right) e^{-\frac{(x^i - x^j)^2}{4\delta^2}}\right)\right]\right\}.$$

Оптимизация проводилась перебором на сетке значений δ . Начальные значения определялись с помощью формулы (I3), по которой получено $\Delta_1^* = \sqrt{2}$; $\Delta_{10}^* = 0{,}75$; $\Delta_{50}^* = 0{,}52$; $\Delta_{100}^* = 0{,}45$; $\Delta_{200}^* = 0{,}39$; $\Delta_{500}^* = 0{,}33$; $\Delta_{1000}^* = 0{,}275$.

В результате расчетов для выборки $\{x_{\alpha_1}^i\}_1^{100}$ описанного выше примера получено $\delta^* = 0{,}0875$.

Расчет исходных плотностей. Оценки исходных плотностей рассчитывались для описанного выше примера. Зависимости $\hat{\lambda}_1(x) = \frac{\hat{P}_{\alpha_2}(x)}{\hat{P}_{\alpha_2}(x) - \hat{P}_{\alpha_1}(x)}$ при $\hat{P}_{\alpha_2}(x) > \hat{P}_{\alpha_1}(x)$ и $\lambda_2(x)$ соответственно при $\hat{P}_{\alpha_2}(x) < P_{\alpha_1}(x)$ приведены на рис.2. Полученные значения оценок коэффициентов в (I5) $\hat{\lambda}_1 = 1{,}21$; $\hat{\lambda}_2 = -0{,}56$ определили оценки парциальных плотностей, приведенные на рис.3. Кривые, обозначенные пунктиром, представляют собой точные парциальные плотности.

Таким образом, в рассмотренном примере удалось выделить исходные парциальные плотности и даже получить достаточно точные их оценки.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

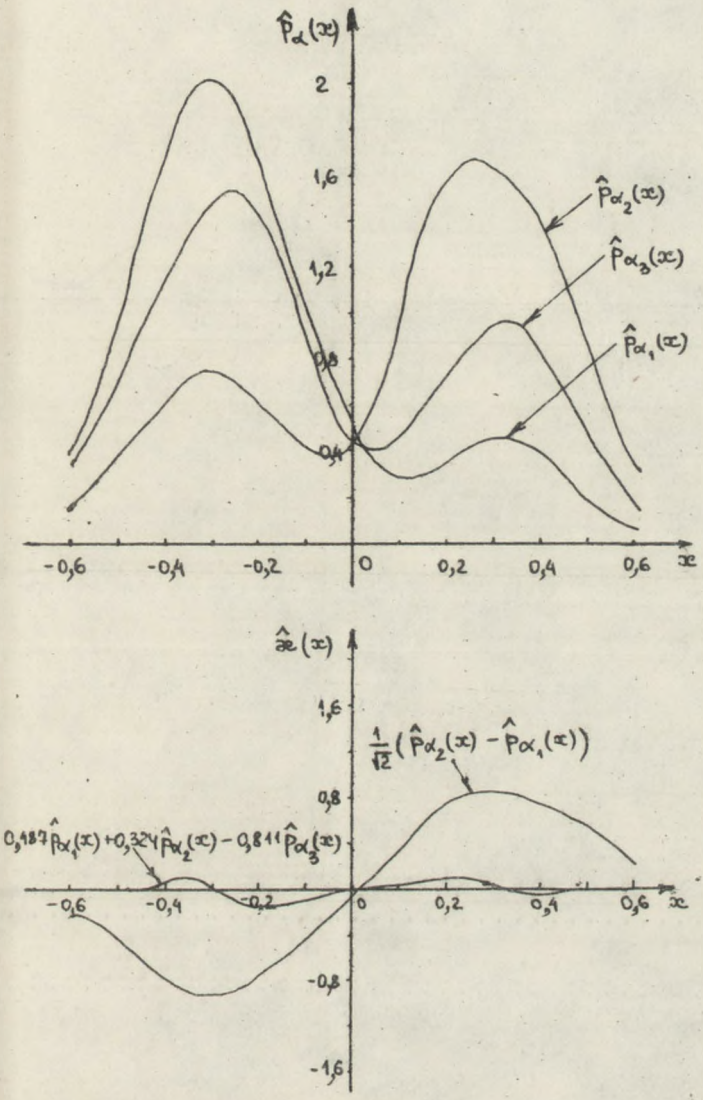

Рис. 1.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

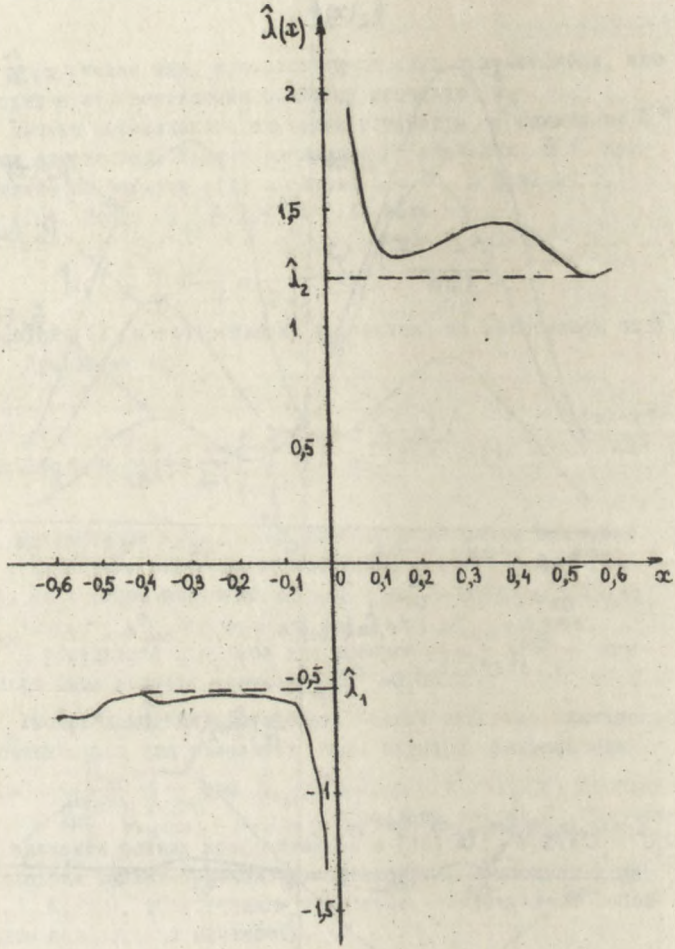

Рис. 2

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

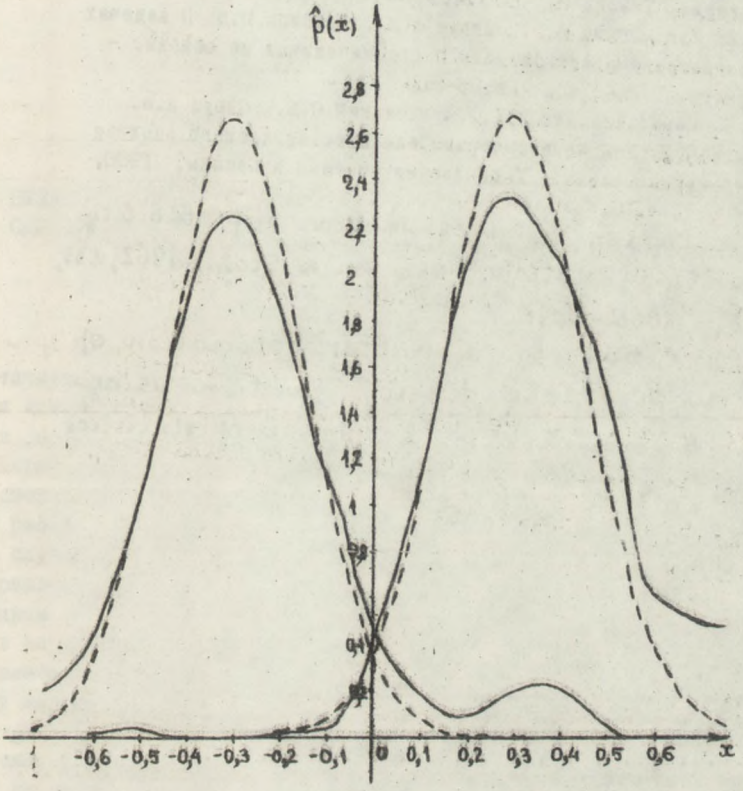

Рис. 3

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Литература

I. Алимов Ю.Н. Об оценивании устойчивости эмпирических распределений. - В кн.:Математическая статистика и ее приложения, Томск: Изд-во Том. ун-та, 1983, с.5-9.

2. Кульчицкий Ю.О., Слоущ А.В., Соколов Ю.В. О задачах статистической классификации с ограничениями на ошибки. - Кибернетика, 1981, №2, с.83-91.

3. Первозванский А.А., Кульчицкий О.Ю., Слоущ А.В. Соколов Ю.В. Комплекс программ для автоматического анализа электрокардиограмм. - Управляющие системы и машины, 1980, №2, с.107-110.

4. Parsen E. On estimation of probability density function.- Ann. Math. Stat., 1962, v.33, №3, p. 1065 - 1076.

5. V. Susarla, G. Walter. Estimation of a multivariate density function using delta sequences.- The Annals of statistics, 1981, v. 9, №2, c. 347-355.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

НЕПАРАМЕТРИЧЕСКАЯ ОЦЕНКА СПЕКТРАЛЬНОЙ ПЛОТНОСТИ СЛУЧАЙНОГО ПРОЦЕССА МЕТОДОМ НЕЛИНЕЙНОГО РЕЗОНАНСНОГО ФИЛЬТРА

Т.Б.Варядченко, О.Ю.Кульчицкий

Введение

Статистический спектральный анализ является в настоящее время важным инструментом исследования свойств случайных сигналов различной физической природы. Начиная с монографии А.А.Харкевича [I], заложившей физические основы статистической спектрометрии, в литературе известно значительное количество работ, посвященных специальным методам спектрального анализа случайных процессов, адекватным различным условиям съёма и обработки изменений. Перечислим основные из них.

Одним из первых предложен получивший широкое распространение в литературе метод вычисления спектральных плотностей через выборочные корреляционные функции [I-3]. Обзор работ по этому методу имеется, например, в [4]. Однако при всей своей структурной простоте метод выборочных корреляционных функций оказывается достаточно сложен при технической реализации даже на незначительных объемах обрабатываемых выборок. Для устранения этого недостатка в последнее время метод спектрального анализа выборочных корреляционных функций сочетается с методами быстрого преобразования Фурье [5], что оказывается эффективным при спектральном анализе реализаций случайных процессов большой длительности.

Другим широко распространенным методом выборочного спектрального анализа случайных процессов является метод дробно -

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

рациональной аппроксимации спектральных плотностей. При этом задача спектрального анализа сводится к задаче параметрического оценивания [6]. При всем удобстве алгоритмов этого метода он требует значительной априорной информации о виде оцениваемой спектральной плотности, в том числе о характере убывания спектра и порядке дробно-рациональной аппроксимации. Поскольку ошибки в априорных допущениях приводят здесь к существенным погрешностям результирующей аппроксимации, то применение этого метода оказывается не эффективным при отсутствии точных модельных представлений об исследуемом случайном процессе.

В настоящей работе исследованы алгоритмы непараметрического оценивания спектральной плотности случайных процессов, основанный на идеях обработки исследуемых случайных процессов с помощью нелинейных резонансных фильтров [2]. Полученные на этом пути алгоритмы оценки спектральной плотности являются технически просто реализуемыми и обеспечивают хорошую точность даже на сравнительно небольших выборках.

Постановка задачи

Рассмотрим центрированный стационарный случайный процесс $x(t)$ с непрерывным временем $t \in (-\infty, \infty)$, спектральную плотность которого $S_x(\omega)$ требуется оценить. Предположим далее, что

$$y(t) = \int_{-\infty}^{t} h(t-s)x(s)ds, \qquad (1)$$

где $\int_0^\infty |h(t)|dt < \infty$. Из теории случайных процессов известно, что $y(t)$ также является центрированным, стационарным процессом таким, что

$$M\{y^2(t)\} = \sigma_y^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |H(j\omega)|^2 S_x(\omega)d\omega, \qquad (2)$$

где $H(p) = \int_0^\infty e^{-pt}h(t)dt$,

$S_x(\omega)$ - спектральная плотность процесса $x(t)$.

Основная идея нелинейного резонансного фильтра для оценки плотности $S_x(\omega)$ состоит в том, чтобы выбрать в (2) передаточную функцию $H(p)$, обладающую свойством

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$|H(j\omega)|^2 \simeq 2\pi\,\delta(\omega-\omega_0).$$

При этом из (I) и (2) будем иметь

$$S_x(\omega_0) \simeq \sigma_y^2 = \lim_{K\to\infty}\frac{1}{K}\int_0^K y^2(t)\,dt,$$

что и обеспечит желаемую оценку.

Перейдем к более детальному описанию изложенной идеи. Для этого зададим функцию $H(p)$ в виде

$$H(p) = \frac{k}{Tp+1} \qquad (3)$$

при $\frac{k}{T} = \sqrt{2\beta}$, $\frac{1}{T} = \beta - j\omega_0$; $\beta > 0$,

так что для нее оказывается справедливым равенство

$$|H(j\omega)|^2 = |H(j(\omega-\omega_0))|^2 = \frac{2\beta}{(\omega-\omega_0)^2+\beta^2}\,, \qquad (4)$$

Подставим (4) в (2) и, перейдя к пределу при $\beta \to +0$, получим [7]:

$$\sigma_y^2(\omega_0) = \lim_{\beta\to+0}\sigma_y^2(\omega_0,\beta) =$$

$$= \lim_{\beta\to+0}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{2\beta}{(\omega-\omega_0)^2+\beta^2}\,S_x(\omega_0)\,d\omega = S_x(\omega_0), \qquad (5)$$

если функция $S_x(\omega)$ ограничена при всех $\omega \in (-\infty,\infty)$ и непрерывна в точке $\omega = \omega_0$.

Таким образом, согласно (5), дисперсия случайного процесса $y(t)$, полученного из (I), (3) при достаточно малом β , оказывается близкой к значению $S_x(\omega_0)$. Поэтому задача оценки спектральной плотности процесса $S_x(\omega_0)$ сводится к выборочной оценке дисперсии процесса $y(t)$.

На практике часто реализация исходного случайного процесса $x(t)$ задается в виде выборки на регулярной сетке аргумента с

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

шагом τ . По этой дискретной выборке и требуется оценить значения спектральной плотности $S_x(\omega_0)$ исходного случайного процесса $x(t)$ как процесса с непрерывным временем.

Покажем , каким образом в изложенной выше методике может быть учтен дискретный характер реализации процесса $x(t)$, доступной для статистической обработки. Для этого отметим, что из представления процесса $y(t)$ в виде (I), (3) следует

$$y_{k+1} = e^{-\frac{\tau}{T}} y_k + \frac{k}{T}\int_0^{\tau} e^{-\frac{(\tau-s)}{T}} x(k\tau+s)\,ds,$$
$$y_0 = \int_{-\infty}^{0} h(-s)\,x(s)\,ds, \qquad (6)$$

где $y_k \triangleq y(k\tau)$.

Если тактовый интервал выборки τ намного меньше времени корреляции процесса $x(t)$, то интеграл

$$f_k = \frac{k}{T}\int_0^{\tau} e^{-\frac{(\tau-s)}{T}} x(k\tau+s)\,ds,$$

входящий в правую часть разностного уравнения (6), может быть приближенно вычислен, например, по одной из формул

$$f_k \approx \frac{k}{T}\int_0^{\tau} e^{-\frac{(\tau-s)}{T}} ds \cdot x_k = k\left(1-e^{-\frac{\tau}{T}}\right)x_k \; ; \qquad (7)$$

$$f_k \approx \frac{k}{T}\int_0^{\tau} e^{-\frac{(\tau-s)}{T}} ds \cdot \frac{x_{k+1}+x_k}{2} = \frac{k}{2}\left(1-e^{-\frac{\tau}{T}}\right)(x_{k+1}+x_k); \qquad (8)$$

$$f_k \approx \frac{k}{T}\int_0^{\tau} e^{-\frac{(\tau-s)}{T}} \left(x_k + \frac{s}{\tau}(x_{k+1}-x_k)\right) = \qquad (9)$$

$$= k\left(1-\frac{T}{\tau}\left(1-e^{-\frac{\tau}{T}}\right)\right)x_{k+1} - k\left(1-\left(1+\frac{T}{\tau}\right)\left(1-e^{-\frac{\tau}{T}}\right)\right)x_k .$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Точность вычисления величин f_k по формулам (7) и (8) при $\tau \to 0$ имеет порядок τ, а по формуле (9) – τ^2. В дальнейшем изложении настоящей статьи будем пользоваться формулой (7) как наиболее простой из приведенных.

Подставляя (7) в (6), получим разностное уравнение, связывающее дискретные по времени выборки из случайных процессов $x(t)$ и $y(t)$:

$$y_{k+1} = q y_k + \varkappa x_k, \qquad y_0 = \hat{y}_0, \tag{10}$$

где

$$q = e^{-\frac{\tau}{T}} = e^{-(\beta + j\omega_0)\tau};$$

$$\varkappa = k\left(1 - e^{-\frac{\tau}{T}}\right) = \frac{\sqrt{2\beta}\,(\beta + j\omega_0)}{\beta^2 + \omega_0^2}\left(1 - e^{-(\beta+j\omega_0)\tau}\right).$$

Оценку значения спектральной плотности $S_x(\omega_0)$ предлагается находить по формуле

$$S_x(\omega_0) = \frac{1}{K}\sum_{k=0}^{K} |y_k|^2 \simeq \sigma_y^2(\omega_0, \beta), \tag{11}$$

где K – объем выборки процесса $x(t)$.

Формулы (10), (11) задают нелинейный резонансный фильтр для поточечной (непараметрической) оценки значений спектральной плотности процесса $x(t)$ по его дискретной выборке $x(0)$, $x(\tau)$, $x(2\tau)$, ..., $x(K\tau)$.

Отметим также, что начальное условие $\hat{y}_0$ линейной части резонансного фильтра (10) является неизвестной случайной величиной, так как выражается через неизвестные значения процесса $x(t)$, и потому подлежит специальному выбору из условия обеспечения наибольшей точности оценивания.

Исследование точности оценки спектральной плотности

Вычислим среднее значение и дисперсию оценки (11) спектральной плотности $\hat{S}_x(\omega_0)$:

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$M\{\hat{S}_x(\omega_0)\} = \frac{1}{K+1}\sum_{k=0}^{K} M\{|y_k|^2\}\ ; \qquad (12)$$

$$D\{\hat{S}_x(\omega_0)\} = M\{\hat{S}_x^2(\omega_0)\} - M\{\hat{S}_x(\omega_0)\}^2\ ; \qquad (13)$$

$$M\{\hat{S}_x^2(\omega_0)\} = \frac{1}{(K+1)^2}\sum_{k=0}^{K}\sum_{z=0}^{K} M\{|y_k|^2\cdot|y_z|^2\} =$$

$$= \frac{1}{(K+1)^2}\Big(\sum_{k=0}^{K} M\{|y_k|^4 + 2\sum_{k=0}^{K}\sum_{z=0}^{k-1} M\{|y_k|^2|y_z|^2\}\ . \qquad (14)$$

Таким образом, для вычисления точностных характеристик оценки (11) необходимо вычислить величины $M\{|y_k|^2\}$ и $M\{|y_k|^2\cdot|y_z|^2\}$. Сделаем это, использовав уравнение (10), из которого следует, что

$$y_k = \tilde{y}_k + q^k \Delta y_0\ , \qquad (15)$$

где $\tilde{y}_k = \sum_{-\infty}^{k} q^{k-z} x_z$, $\Delta y_0 = \hat{y}_0 - \tilde{y}_0$

Поэтому

$$M\{|y_k|^2\} = \sigma_{\tilde{y}}^2 + 2\,\mathrm{Re}\,(q^k R(k)) + |q|^{2k}\sigma_{\Delta y_0}^2 \qquad (16)$$

при $R(k) \triangleq M\{\Delta y_0 \tilde{y}_k^*\}$.

Знак * здесь и далее означает сопряженное значение.

Подстановка (16) в (12) дает

$$M\{\hat{S}_x(\omega_0)\} = \sigma_{\tilde{y}}^2 + \frac{1-|q|^{2(K+1)}}{(K+1)(1-|q|^2)}\sigma_{\Delta y_0}^2 + \qquad (17)$$

$$+ \frac{2}{K+1}\sum_{k=0}^{K}\mathrm{Re}\,(q^k R(k))\ .$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Для третьего слагаемого в (I7) может быть сделана следующая оценка сверху:

$$\left| \frac{2}{K+1} \sum_{k=0}^{K} \mathrm{Re}\left(q^k R(k)\right) \right| \leq \frac{2\left(1-|q|^{K+1}\right)}{(K+1)(1-|q|)} \sigma_{\Delta y_0} \sigma_{\tilde{y}} . \quad (I8)$$

Поэтому оценка (II) спектральной плотности $S_x(\omega_0)$ является смещенной, причем величина смещения имеет порядок $\frac{1}{K}$ при $K \to \infty$ и пропорциональна величине $\sigma_{\Delta y_0}$.

Для дисперсии окончательные формулы можно записать в предположении гауссовости процесса y_k; воспользовавшись значением интеграла

$$I = M\{w^2 v^2\} = \int_{-\infty}^{\infty}\int w^2 v^2 \frac{1}{2\pi (\det R)^{1/2}} e^{-\frac{1}{2}(w,v)R^{-1}\binom{w}{v}} dw\, dv =$$

$$= \sigma_w^2 \cdot \sigma_v^2 + 2R_{wv}^2 ,$$

где

$$R = \begin{pmatrix} \sigma_w^2 & R_{wv} \\ R_{wv} & \sigma_v^2 \end{pmatrix},$$

получим

$$M\{|y_k|^2 |y_z|^2\} = \sigma_{\tilde{y}}^4 + |R_y(k,z)|^2 + |\tilde{R}_y(k,z)|^2, \quad (I9)$$

где

$$R_y(k,z) = M\{y_k y_z^*\} = R_{\tilde{y}}(z-k) + q^z R^*(k) +$$

$$+ q^k R(z) + |q|^{2z} q^{k-z} \sigma_{\Delta y_0}^2 ;$$

$$\tilde{R}_y(k,z) = M\{y_k y_z\} = \tilde{R}_{\tilde{y}}(z-k) + q^z \tilde{R}(k) + \quad (20)$$

$$+ q^k \tilde{R}(z) + q^{k+z} \sigma_{\Delta y_0}^2 ;$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$R_{\tilde{y}}(\tau-\kappa) \triangleq M\{\tilde{y}_\kappa \tilde{y}_\tau^*\} \; ; \qquad R(\kappa) \triangleq M\{\tilde{y}_\kappa^* \Delta y_0\} \; ;$$

$$\tilde{R}_{\tilde{y}}(\tau-\kappa) \triangleq M\{\tilde{y}_\kappa \tilde{y}_\tau\} \; ; \qquad \tilde{R}(\kappa) \triangleq M\{\tilde{y}_\kappa \Delta y_0\} .$$

Подстановка (20) в (I9) и в (I4) дает соотношение, позволяющее сделать оценки сверху, аналогичные (I8).

Исследование точности оценок при нулевых начальных условиях Положив в (I5) $\hat{y}_0 = 0$, получим

$$M\{\hat{S}_x(\omega_0)\} = \sigma_{\tilde{y}}^2 \left(1 + \frac{1-|q|^{2(K+1)}}{(K+1)(1-|q|^2)}\right) - \frac{2}{K+1} \sum_{\kappa=0}^{K} \mathrm{Re}\left(q^\kappa R_{\tilde{y}}(\kappa)\right) , \tag{21}$$

где

$$R_{\tilde{y}}(\kappa) \triangleq M\{\tilde{y}_0 \tilde{y}_\kappa^*\} .$$

Для точного вычисления смещения оценки следует определить значения $R_{\tilde{y}}(\kappa)$ через корреляционную функцию входа. Сделаем это, учитывая, что

$$\tilde{y}_{\kappa+1} = q\tilde{y}_\kappa + \varkappa x_\kappa \; , \qquad \tilde{y}_0 = \sum_{\kappa=-\infty}^{0} \varkappa q^{-\kappa} x_\kappa \; ,$$

$$\tilde{y}^*_{\kappa+1} = q^*\tilde{y}^*_\kappa + \varkappa^* x_\kappa .$$

Умножая последнее соотношение на $\tilde{y}_0$ и осредняя, получим разностное уравнение для $R_{\tilde{y}}(\kappa)$, решение которого при нулевых начальных условиях имеет вид

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$R_{\tilde{y}}(\kappa) = \frac{|\varkappa|^2}{1-|q|^2} q^{*s} \Big(\sum_{p=-\infty}^{s-1} q^{*-p} R_x(p) + |q|^{-2s} \sum_{p=s}^{\infty} q^p R_x(p) \Big) . \qquad (22)$$

Подстановка (22) в (21) дает следующее соотношение, позволяющее вычислять математическое ожидание оценки спектральной плотности (21) при заданных корреляционных функциях входа

$$\sum_{k=0}^{K} q^k R_{\tilde{y}}(k) = \frac{|\varkappa|^2}{(1-|q|^2)^2} \left(1-|q|^{2(K+1)}\right) \cdot \Big(\sum_{p=-\infty}^{0} q^{*-p} R_x(p) + \sum_{p=1}^{K} q^p R_x(p) \Big) + \qquad (23)$$

$$+ \frac{|\varkappa|^2}{1-|q|^2} \Big((K+1) \sum_{p=K+1}^{\infty} q^p R_x(p) + \sum_{p=1}^{K+1} p q^p R_x(p) \Big) .$$

Так, при гармоническом входном воздействии $x(t) = a\sin\omega t + b\cos\omega t$, $M\{a\} = M\{b\} = 0$, $M\{a^2\} = M\{b^2\} = \sigma_x^2$, $R_x(p) = \sigma_x^2 \cos\omega p$, на частоте резонанса $\omega = \omega_0$ при предельном значении параметра $p \to 0$ величина относительного смещения оценки получается равной

$$\gamma = \frac{M\{\hat{S}_x(\omega_0)\} - \sigma_{\tilde{y}}^2}{\sigma_{\tilde{y}}^2} \cdot 100\% = 50\% .$$

Это - смещение оценки, которое дает установившийся процесс на выходе фильтра.

Численные расчеты, которые проводились при малых значениях K (K =50), дали при предельном значении параметра величину смещения, равную 100%, и среднеквадратическое отклонение оценки, составляющее 20% от математического ожидания. Результаты вычислений для различных значений параметра p приведены в таблицах. По результатам видно, как на частотах

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

б) с увеличением значения β происходит увеличение смещения оценок и уменьшение их дисперсии. Например, при расстройке в 5 Гц и $\beta = 10^{-5}$ смещение составляет 1,5%, среднеквадратическое отклонение-18%, а при значении $\beta = 10$ эти величины равны соответственно 42 и 16%. Из этого факта, казалось бы, можно сделать вывод о том, что существует оптимальное (компромиссное) значение β, обеспечивающее одинаковый вклад систематической и случайной погрешностей. Однако это свойство не наблюдается на резонансной частоте: на резонансной частоте с уменьшением β растут и смещение, и дисперсия. Поэтому сумма систематической и случайной погрешностей не может быть использована для выбора оптимального значения

Следует отметить, что если спектральная плотность $S_x(\omega)$ оценивается с целью определения основной частоты процесса $x(t)$, то оценка этой частоты оказывается не очень критичной к выбору β.

Из таблиц можно установить величину разрешающей способности фильтра, т.е. интервал по частоте, при котором доверительные коридоры на оценку не перекрываются - этот интервал в широком диапазоне изменений β составляет 2,5 Гц.

Такая разрешающая способность фильтра, полученная при числе наблюдений $K = 50$, может оказаться достаточной в практических задачах. В частности, авторам известны задачи определения сетевой наводки в шумах электрокардиограмм, в которых число отсчетов на диастоле $K = 100 \div 200$, так что указанная величина 2,5 Гц дает нижнюю границу "разрешающей способности" фильтра.

Увеличение числа наблюдений K приводит к улучшению точностных характеристик, как это было уже продемонстрировано.

Выводы

Проведенные исследования позволили установить точностные характеристики предлагаемого нелинейного резонансного фильтра для оценки спектральной плотности случайного процесса. Уменьшение смещения оценок возможно за счет введения ненулевых начальных условий.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Таблицы значений точностных характеристик оценок спектральной плотности при различных значениях параметра β, гармоническом входном воздействии частотой 50 Гц и количестве наблюдений K =50 (частота настройки фильтра $f_0 \in [45\text{-}50]$Гц, σ_x^2 =I, τ =0,004)

I. β =I0,0

f_0, Гц	$\sigma_{\tilde{y}}^2$	$M\{\hat{S}_x(\omega_0)\}$	Абсолютное смещение	Относ. смещение	Среднекв. отклонение	Относит. среднеквадр. откл. x)
45	$0,I65\cdot I0^{-I}$	$0,957\cdot I0^{-2}$	$-0,700\cdot I0^{-2}$	42,4	$0,I52\cdot I0^{-2}$	I5,9
46	$0,244\cdot I0^{-I}$	$0,I40\cdot I0^{-I}$	$-0,I00\cdot I0^{-I}$	4I,0	$0,220\cdot I0^{-2}$	I5,7
47	$0,39I\cdot I0^{-I}$	$0,20I\cdot I0^{-I}$	$-0,I90\cdot I0^{-I}$	48,6	$0,3I9\cdot I0^{-2}$	I5,9
48	$0,686\cdot I0^{-I}$	$0,266\cdot I0^{-I}$	$-0,420\cdot I0^{-I}$	6I,2	$0,438\cdot I0^{-2}$	I6,5
49	0,I26	$0,3I6\cdot I0^{-I}$	$-0,946\cdot I0^{-I}$	75,I	$0,536\cdot I0^{-2}$	I7,0
50	0,I75	$0,334\cdot I0^{-I}$	-0,I42	8I,I	$0,572\cdot I0^{-2}$	I7,I

2. β =I

45	$0,I82\cdot I0^{-2}$	$0,I63\cdot I0^{-2}$	$-0,I86\cdot I0^{-3}$	I0,2	$0,29I\cdot I0^{-3}$	I7,8
46	$0,283\cdot I0^{-2}$	$0,296\cdot I0^{-2}$	$0,I36\cdot I0^{-3}$	4,8	$0,488\cdot I0^{-3}$	I6,5
47	$0,499\cdot I0^{-2}$	$0,505\cdot I0^{-2}$	$0,6I0\cdot I0^{-4}$	I,2	$0,858\cdot I0^{-3}$	I7,0
48	$0,III\cdot I0^{-I}$	$0,748\cdot I0^{-2}$	$-0,365\cdot I0^{-2}$	32,9	$0,I36\cdot I0^{-2}$	I8,2
49	$0,435\cdot I0^{-I}$	$0,944\cdot I0^{-2}$	$-0,340\cdot I0^{-I}$	78,2	$0,I8I\cdot I0^{-2}$	I9,2
50	$0,I75\cdot I0$	$0,I02\cdot I0^{-I}$	$-0,I74\cdot I0$	99,4	$0,I98\cdot I0^{-2}$	I9,4

x) Абсолютное смещение $\Delta = M\{\hat{S}_x(\omega_0)\} - \sigma_{\tilde{y}}^2$.

Относительное смещение $\gamma = |\Delta| \cdot 100\% / \sigma_{\tilde{y}}^2$.

Среднеквадратическое отклонение $d = \sqrt{D\{\hat{S}_x(\omega_0)\}}$.

Относительное среднеквадратическое отклонение $\varepsilon = d / M\{\hat{S}_x(\omega_0)\}$.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

3. β =0,I

f_0, Гц	$\sigma_{\tilde{y}}^2$	$M\{\hat{S}_x(\omega_0)\}$	Абсолютное смещение	Относ. смеще-ние	Среднекв. откло-нение	Относит. среднеквадр. откл.
45	$0,I82\cdot I0^{-3}$	$0,I78\cdot I0^{-3}$	$0,450\cdot I0^{-5}$	2,5	$0,3I9\cdot I0^{-4}$	I7,9
46	$0,283\cdot I0^{-3}$	$0,329\cdot I0^{-3}$	$0,456\cdot I0^{-4}$	I6,I	$0,545\cdot I0^{-4}$	I6,6
47	$0,50I\cdot I0^{-3}$	$0,569\cdot I0^{-3}$	$0,685\cdot I0^{-4}$	I3,7	$0,978\cdot I0^{-4}$	I7,2
48	$0,II2\cdot I0^{-2}$	$0,850\cdot I0^{-3}$	$-0,270\cdot I0^{-3}$	24,I	$0,I57\cdot I0^{-3}$	I8,5
49	$0,446\cdot I0^{-2}$	$0,I08\cdot I0^{-2}$	$-0,338\cdot I0^{-2}$	75,8	$0,2I0\cdot I0^{-3}$	I9,4
50	$0,I75\cdot I0^{2}$	$0,II6\cdot I0^{-2}$	$-0,I75\cdot I0^{2}$	I00,0	$0,23I\cdot I0^{-3}$	I9,9

4. β =0,0I

45	$0,I82\cdot I0^{-4}$	$0,I79\cdot I0^{-4}$	$-0,200\cdot I0^{-6}$	I,I	$0,320\cdot I0^{-5}$	I7,9
46	$0,283\cdot I0^{-4}$	$0,352\cdot I0^{-4}$	$0,490\cdot I0^{-5}$	I7,3	$0,550\cdot I0^{-5}$	I6,6
47	$0,500\cdot I0^{-4}$	$0,576\cdot I0^{-4}$	$0.750\cdot I0^{-5}$	I5,0	$0,990\cdot I0^{-5}$	I7,2
48	$0,II2\cdot I0^{-3}$	$0,86I\cdot I0^{-4}$	$-0,259\cdot I0^{-4}$	23,I	$0,I59\cdot I0^{-4}$	I8,5
49	$0,446\cdot I0^{-3}$	$0,I09\cdot I0^{-3}$	$-0,336\cdot I0^{-3}$	75,3	$0,2I3\cdot I0^{-4}$	I9,5
50	$0,I75\cdot I0^{3}$	$0,II8\cdot I0^{-3}$	$-0,I75\cdot I0^{3}$	I00,0	$0,234\cdot I0^{-4}$	I9,8

5. β =0,00001

45	$0,I82\cdot I0^{-7}$	$0,I79\cdot I0^{-7}$	$-0,276\cdot I0^{-9}$	I,5	$0,322\cdot I0^{-8}$	I7,9
46	$0,283\cdot I0^{-7}$	$0,333\cdot I0^{-7}$	$0,496\cdot I0^{-8}$	I7,5	$0,552\cdot I0^{-8}$	I6,5
47	$0,50I\cdot I0^{-7}$	$0,577\cdot I0^{-7}$	$0,763\cdot I0^{-8}$	I5,2	$0,993\cdot I0^{-8}$	I7,2
48	$0,I00\cdot I0^{-6}$	$0,862\cdot I0^{-7}$	$-0,258\cdot I0^{-7}$	25,8	$0,I60\cdot I0^{-7}$	I8,5
49	$0,400\cdot I0^{-6}$	$0,I00\cdot I0^{-6}$	$-0,300\cdot I0^{-6}$	75,0	$0,2I4\cdot I0^{-7}$	2I,4
50	$0,I65\cdot I0^{6}$	$0,I00\cdot I0^{-6}$	$-0,I65\cdot I0^{6}$	I00,0	$0,235\cdot I0^{-7}$	23,5

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Литература

1. Харкевич А.А. Спектры и анализ. - М.: Физматгиз, 1962.-98с.

2. Дженкинс Г., Ватте Д. Спектральный анализ и его приложение.- М.: Мир, 1971.-265с.

3. Серебренников М.Г., Первозванский А.А. Выявление скрытых периодичностей.-М.: Наука, 1965.-187с.

4. Цветков Э.И. Основы теории статистических измерений.- Л.: Энергоатомиздат, 1986.-378с.

5. Бендат Дж., Пирсол А. Измерение и анализ случайных процессов.-М.: Мир, 1974.-464 с.

6. Эйкхофф П. Основы идентификации систем управления.- М.: Мир, 1975.-256с.

7. Корн Г. и Корн Т. Справочник по математике для научных работников и инженеров.-М.: Наука, 1973.-831 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

АКСИОМАТИЧЕСКИЙ ПОДХОД К ВЫБОРУ КРИТЕРИЕВ ОЦЕНИВАНИЯ

Н.О.Вильчевский, Г.Л.Шевляков

Анализ результатов измерений на практике чаще всего сводится к определению характеристик средних значений, таких как параметр сдвига (центр распределений), линии регрессии и т.п. При независимых помехах $\{\varepsilon_i\}$ в измерениях $x_i = \theta + \varepsilon_i$ $(i = 1, \dots n)$ важный класс оценок $\hat{\theta}$ определяется решением следующей оптимизационной задачи (M – оценки):

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{n} \rho(x_i - \theta). \qquad (1)$$

Здесь свойства оценок $\hat{\theta}$ существенно зависят от выбора функций потерь ρ и вида распределения помехи $\{\varepsilon_i\}$.

В известном смысле произвол выбора функции потерь в критерии (1) был снят теорией робастного оценивания [1] . Основным условием правомерности использования этого статистического подхода является априорная уверенность в стабильности условий эксперимента и его воспроизводимости. Убедиться в этом можно лишь в случае реализации требования многократного воспроизведения прогнозируемого экспериментального результата при проведении серий экспериментов [2] .

Однако в последнее время наблюдается "вульгарное" применение методов робастного оценивания в условиях, когда обрабатывается лишь одна выборка. С одной стороны, это дискредитирует статистическую теорию робастого оценивания, а с другой – свидетельствует о недостаточном развитии методов анализа данных при отсутствии верификации стабильности и воспроизводимости условий эксперимента.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Следует иметь в виду, что, как и большинство статистических методов, классические варианты робастного оценивания дают хорошие в "среднем" на классе распределений результаты, а на отдельных выборках могут давать недостоверные результаты.

В настоящей работе сделана попытка ограничить выбор функции потерь ρ в критерии (I) на основе заранее заданных требований к свойствам оценок $\hat{\theta}$, причем эти требования должны выполняться для каждой отдельно взятой выборки.

В интересной работе [3] рассмотрена задача выбора такой функции потерь ρ, чтобы полученная оценка параметра сдвига $\hat{\theta}$ была инвариантна к изменению масштаба данных: т.е. если результаты измерений $x_1, \ldots, x_n$ заменяются на $\lambda x_1, \ldots, \lambda x_n$, то и оценка сдвига $\hat{\theta}$ заменяется на $\lambda\hat{\theta}$. Метрологически это означает независимость алгоритмического обеспечения обработки от типов шкал используемых приборов. Метод решения этой задачи, предложенный в [3], не допускает обобщения на многомерные постановки задачи. В связи с этим приведем здесь иное доказательство.

Необходимое условие решения задачи (I) при произвольном выборе масштаба λ имеет вид

$$\sum_{i=1}^{n} \rho'[\lambda x_i - \lambda\hat{\theta}(\lambda)] = 0. \qquad (2)$$

Требование инвариантности оценки θ к масштабированию равносильно условию

$$d\hat{\theta}/d\lambda = 0,$$

что при дифференцировании уравнения (2) дает

$$\sum_{i=1}^{n} (\lambda x_i - \lambda\theta)\rho''(\lambda x_i - \lambda\theta) = 0 \qquad (3)$$

Так как условия (2) и (3) должны выполняться для любых данных $\{x_i\}$ и любых объемов выборки n, то возьмем выборки, состоящие из A величин x и B величин y:

$$\begin{aligned} A\rho'(X) + B\rho'(Y) &= 0, \\ AX\rho''(X) + BY\rho''(Y) &= 0 \end{aligned} \qquad (4)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Здесь введены очевидные обозначения: $X=\lambda(x-\hat{\theta})$, $Y=\lambda(y-\hat{\theta})$.

В связи с произвольностью величин A и B получаем, что решение исходной задачи эквивалентно определению функции ρ , для которой выполнено условие равенства нулю определителя системы (4)

$$\begin{vmatrix} \rho'(X) & \rho'(Y) \\ X\rho''(X) & Y\rho''(Y) \end{vmatrix} = 0 ,$$

откуда имеем уравнение для функции ρ

$$\frac{X\rho''(X)}{\rho'(X)} = K , \qquad (5)$$

где K - произвольная постоянная. Интегрируя это уравнение, получаем, что справедлива следующая

ТЕОРЕМА I. Инвариантная к выбору масштаба функция потерь в задаче (I) имеет вид

$$\rho(X) = C|X|^{\rho}, \quad C>0, \ \rho=K+1, \ \rho\geq 1. \qquad (6)$$

Последнее условие ($\rho \geq 1$) следует из требования единственности решения задачи (I). Таким образом, инвариантность к масштабированию сужает класс возможных критериев в задаче (I) до стандартного класса L_ρ - оценок теории приближений. Величина ρ в выражении (6) должна быть выбрана либо из априорной статистической информации об условиях эксперимента, либо из других (помимо инвариантности к масштабу) требований к оценкам.

Задача (I) в случае, когда измерения представляют собой векторы $x_i = (x_{i1}, \dots, x_{im})^T (i = 1, \dots, n)$, имеет вид

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{n} \rho(x_{i1}-\theta_1, \dots, x_{im}-\theta_m), \qquad (7)$$
$$\theta = (\theta_1, \dots, \theta_m)^T.$$

В зависимости от требований практики возникают различные требования по инвариантности к масштабу. Наиболее часто накладывается одно из двух условий:

а) инвариантность оценки к различному и независимому изменению масштаба компонент вектора измерений;

б) инвариантность оценки к одинаковому изменению масштаба ком-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

понент вектора измерений.

Обобщение метода доказательства теоремы I позволяет получить следующие результаты.

ТЕОРЕМА 2. Инвариантность оценок параметра сдвига в задаче (7) в смысле условий а) реализуется при выборе функции потерь вида

$$\rho(X_1,\dots,X_m)=\sum_{j=1}^{m} c_j |X_j|^{\rho_j},\quad c_j>0,\ \rho_j\geq 1. \tag{8}$$

ТЕОРЕМА 3. Инвариантность оценок параметра сдвига в задаче (7) в смысле условия б) реализуется при выборе функции потерь вида

$$\rho(X_1,\dots,X_m)=f(|X_1|,\dots,|X_m|), \tag{9}$$

где f – произвольная выпуклая однородная порядка ρ функция.

При доказательстве этих теорем вместо уравнения (5) получаются системы уравнений в частных производных, граничные условия для которых определяются индуктивно через решения $(m-1)$-мерной задачи оценивания.

Полученные результаты позволяют перейти от задач робастного оценивания на классе функций ρ к задачам параметрического робастного оценивания в условиях теорем I и 2. В условиях теоремы 3 сохраняется необходимость постановки задачи робастного оценивания на классе функций потерь с учетом достаточно сильного требования однородности.

Помимо изученных требований к свойствам оценок могут быть рассмотрены и другие, например требования инвариантности решения задачи оценивания к невырожденным аффинным, ортогональным и т.п. преобразованиям. Эти условия могут быть наложены как дополнительные к рассмотренным выше, так и независимо от них. Изучение подобных свойств оценок представляет собой актуальную задачу. В частности, есть основания утверждать, что требование инвариантности к произвольному ортогональному преобразованию сужает класс функций потерь в задаче (7) до следующего:

$$\rho(X_1,\dots,X_m)=c\,d^{\rho}(X_1,\dots,X_m),\quad \rho\geq 1,$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

где d – евклидова норма

$$d(X_1,\ldots,X_m)=\left\{\sum_{j=1}^{m} X_j^2\right\}^{1/2}.$$

Литература

1. Хьюбер П. Робастность в статистике.-М.: Мир, 1984.- 300 с.
2. Алимов Ю.И. О практической ценности теории оценок.- Автоматика, 1981, № 2, с. 84-94.
3. Крейнович В.Я. Оптимизация в условиях неопределенных критериев: Тез. докл. IУ Всес. конф. по приложениям мат.логики, Таллин, 1986, с. 126-128.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

О ПОСЛЕДОВАТЕЛЬНОМ МЕТОДЕ ОБНАРУЖЕНИЯ СКАЧКООБРАЗНЫХ ИЗМЕНЕНИЙ ПАРАМЕТРОВ СЛУЧАЙНЫХ ПРОЦЕССОВ РЕКУРРЕНТНОГО ТИПА

С.Э.Воробейчиков, В.В.Конев

I. Введение

Проблема обнаружения моментов изменения свойств случайных процессов (разладок) актуальна во многих приложениях: в технической и медицинской диагностике, в задачах управления, при обработке временных рядов и т.п. В последние годы был предложен ряд процедур обнаружения разладок случайных процессов при различных предположениях относительно их структуры [1,2] . Для описания процесса наблюдений часто используются модели авторегрессионного типа, которые дают возможность отразить изменения в момент разладки среднего значения процесса, а также его спектральных и дисперсионных свойств. При анализе качества решающих процедур обнаружения разладки в таких моделях наиболее трудной является проблема нахождения их основных характеристик, определяющих частоту ложных тревог и длительность запаздывания в обнаружении разладки. В [3] был предложен метод обнаружения разладок случайных процессов, описываемых стохастическими разностными уравнениями, который дает возможность контролировать указанные характеристики одновременно. Разработанная в [3] решающая процедура является непараметрической, т.е. не требует знания распределения наблюдаемого процесса, и относится к классу последовательных [4-6] , в которых наблюдения проводятся циклами случайной длины, завершающимися решениями о наличии либо отсутствии разладки. Эта процедура прежде всего ориентирована на обнаружение изменений среднего значения и спектральных свойств

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

процесса. Если же в момент разладки среднее и спектральные свойства процесса изменяются мало, решающая процедура может приводить к большим запаздываниям в обнаружении разладки даже при существенных изменениях дисперсионных свойств процесса. В данной работе, являющейся развитием [3], предлагается общая процедура обнаружения любых изменений указанных свойств наблюдаемого процесса. Получены также уточненные формулы для основных статистических характеристик. Приведены результаты численного моделирования, показывающие эффективность общей процедуры.

2. Постановка задачи

Пусть наблюдаемый многомерный случайный процесс $x(t)=\{x_1(t), \ldots, x_n(t)\}$ описывается системой стохастических разностных уравнений

$$x(t+1)=A_1(t,x)+B_1(t,x)\xi(t+1),\ t=0,1,\ldots,\theta-1, \qquad (1)$$

до момента разладки θ и системой уравнений

$$x(t+1)=A_2(t,x)+B_2(t,x)\xi(t+1),\ t=\theta,\theta+1,\ldots, \qquad (2)$$

начиная с момента θ. Здесь $\{\xi(t)\}$ (не зависящая от $x(0)$) – последовательность независимых n -мерных одинаково распределенных случайных векторов с $M\xi(t)=0$, $M\xi(t)\xi^T(t)=E$ (верхний индекс "т" обозначает транспонирование, E - единичная матрица) и $M\|\xi(t)\|^4<\infty$; $\|\xi\|$ - евклидова норма вектора ξ. Распределение $\xi(t)$ неизвестно. Элементы матричных функций $A_i(t,x)$ и $B_i(t,x)$ (размера $n\times 1$ и $n\times n$) могут зависеть от реализации процесса x до момента t включительно, причем матрицы B_i невырождены. Момент разладки θ может быть детерминированным или случайным. В последнем случае предполагается, что θ не зависит от $\{\xi(t)\}$ и $x(0)$ и распределение θ неизвестно. Требуется по реализации процесса x построить решающую процедуру обнаружения разладки.

3. Построение решающей процедуры. Основной результат

Процедуру обнаружения разладки будем строить, используя изменение после разладки средних значений приращений функционала обобщенного метода наименьших квадратов

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$J(m)=\sum_{t=0}^{m-1}\left[x(t+1)-A_1(t,x)\right]^T\left(B_1B_1^T\right)^{-1}(t,x)\left[x(t+1)-A_1(t,x)\right]+$$

$$+\sum_{t=m}^{N}\left[x(t+1)-A_2(t,x)\right]^T\left(B_2B_2^T\right)^{-1}(t,x)\left[x(t+1)-A_2(t,x)\right],$$

N - длина реализации. Приращения этого функционала $\Delta J(m)=$ $=J(m)-J(m-1)$ согласно (1), (2) имеют вид

$$\Delta J(m)=\begin{cases}-r_1(m-1)-\zeta_1(m)+\eta_{1,1}(m), & \text{если } m<\theta;\\ r_2(m-1)+\zeta_2(m)+\eta_{2,1}(m), & \text{если } m\geq\theta,\end{cases}$$

где

$$r_k(m-1)=\left(\left(A_2-A_1\right)^T\left(B_jB_j^T\right)^{-1}\left(A_2-A_1\right)\right)(m-1,x);$$

$$\zeta_k(m)=\xi^T(m)C_k(m-1)\xi(m),\quad C_k(m-1)=\left(B_k^T\left(B_jB_j^T\right)^{-1}B_k\right)-E;$$

$$\eta_{k,1}(m)=2\xi^T(m)\left(B_k\left(B_jB_j^T\right)^{-1}\left(A_2-A_1\right)\right)(m-1,x),\ j=3-k,\ k=1,2. \tag{3}$$

Вычислив условное среднее для $\Delta J(m)$, получим

$$M\left\{\Delta J(m)\mid\mathcal{F}_{m-1}\right\}=\begin{cases}-r_1(m-1)-tr\,C_1(m-1), & \text{если } m<\theta;\\ r_2(m-1)+tr\,C_2(m-1), & \text{если } m\geq\theta,\end{cases}$$

где $\mathcal{F}_{m-1}$ - σ - алгебра, порожденная векторами $x(0),\xi(1),\dots,\xi(m-1)$. Перейдем для удобства к симметризованным величинам $\Delta\bar{J}(m)=$ $=\Delta J(m)+d(m-1)$, где

$$\Delta J(m)=\left[x(m)-A_1(m-1,x)\right]^T\left(B_1B_1^T\right)^{-1}(m-1,x)\left[x(m)-A_1(m-1,x)\right]-$$

$$-\left[x(m)-A_2(m-1,x)\right]^T\left(B_2B_2^T\right)^{-1}(m-1,x)\left[x(m)-A_2(m-1,x)\right], \tag{4}$$

$$d(m-1)=\frac{1}{2}\left(r_1-r_2\right)(m-1)+\frac{1}{2}tr\left(C_1-C_2\right)(m-1).$$

Значение $d(m-1)$ выбрано так, чтобы обеспечить равенство абсолютных значений условных средних $\Delta\bar{J}(m)$ до и после разладки:

$$M\left[\Delta\bar{J}(m)\mid\mathcal{F}_{m-1}\right]=\begin{cases}-r(m-1) & \text{, если } m<\theta;\\ r(m-1) & \text{, если } m\geq\theta,\end{cases}$$

$$r(m)=\frac{1}{2}\left(r_1+r_2\right)(m)+\frac{1}{2}tr\left(C_1+C_2\right)(m). \tag{5}$$

Заметим, что неотрицательные случайные величины $r(m)$ определяют меру различия моделей процесса в момент времени $t=m$, причем первое слагаемое характеризует различие средних значе-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ний и спектральных свойств процесса до и после разладки, а второе - различие дисперсионных матриц $B_1(m,x)$ и $B_2(m,x)$. В случае, когда шумы $\xi(t)$ являются гауссовскими, величина $r(m)$ совпадает с условной дивергенцией по Кульбаку между распределениями вектора $x(m+1)$, отвечающим моделям (1) и (2). Это легко проверить, вычислив условную дивергенцию:

$$I_m(x) = I(1:2; X_0^m) + I(2:1; X_0^m),$$

$$I(i:j; X_0^m) = \int_{R^n} P_i(y; X_0^m) \log \frac{P_i(y; X_0^m)}{P_j(y; X_0^m)} dy_1 \cdots dy_n ; y=(y_1,\dots,y_n),$$

где $P_i(y; X_0^m)$ - условная плотность распределения $x(m+1)$ при фиксированных значениях $x(0), \dots, x(m)$:

$$P_i(y; X_0^m) = (2\pi)^{-n/2} \cdot (\det B_i B_i^T)^{-1} \times$$

$$\times \exp\left\{-\frac{1}{2}(y - A_i(m,x))^T (B_i B_i^T)^{-1} (y - A_i(m,x))\right\}.$$

Последовательность статистик $\{\Delta \bar{J}(m)\}$ и $\{r(m)\}$ положим в основу построения решающей процедуры. Поскольку распределение момента разладки θ неизвестно, процедура должна иметь циклический характер. Это означает, что в определенные моменты времени выносятся решения о наличии или отсутствии разладки и накопленная к этим моментам информация в дальнейшем не используется. Для формирования решений будем использовать статистики $\Delta \bar{J}(m)$, а с помощью $r(m)$ - определять моменты принятия решений. Решения о разладке будем принимать в те моменты τ_i, когда сумма расхождений $r(m)$ между моделями достигнет заданного положительного порога H.

Положим

$$\tau_i = \tau_i(H) = \inf\left\{k > \tau_{i-1} : \sum_{t=\tau_{i-1}+1}^{k} r(t) \geq H\right\}, \; i \geq 1, \; \tau_0 \equiv -1. \quad (6)$$

Решение о разладке в i-м цикле наблюдений будем выносить на основе статистики

$$y_i(H) = \frac{1}{H} \sum_{t=\tau_{i-1}+1}^{\tau_i} \alpha_i(t) \Delta \bar{J}(t+1); \quad (7)$$

$$\alpha_i(t) = \begin{cases} \left[H - \sum_{s=\tau_{i-1}+1}^{\tau_i - 1} r(s)\right] / r(\tau_i), & \text{если } \tau_{i-1} < t < \tau_i; \\ 1, & \text{если } t = \tau_i; \\ 0, & \text{если } t \leq \tau_{i-1} \text{ или } t > \tau_i. \end{cases}$$

Благодаря специальному выбору моментов τ_i, из соотношения (5) имеем

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$y_i(H)=\begin{cases} -1+\varepsilon_{1i}(H) & \text{, если } \tau_i<\theta; \\ +1+\varepsilon_{2i}(H) & \text{, если } \tau_{i-1}\geq\theta; \end{cases}$$

$$\varepsilon_{ki}(H)=\frac{1}{H}\sum_{t=\tau_{i-1}+1}^{\tau_i}\alpha_i(t)\eta_k(t+1),\quad \eta_k(t)=\eta_{k1}(t)+\eta_{k2}(t), \qquad (8)$$

$$\eta_{k2}(t)=(-1)^k\left(\zeta_k(t)-M\zeta_k(t)\right),$$

где $\eta_{k1}(t)$ и ζ_k определяются формулами (3). Учитывая (8), решение о наличии разладки будем принимать в момент τ_i, в который статистика $y_i(H)$ впервые превысит порог δ ($|\delta|<1$). Перейдем к анализу качества построенной решающей процедуры и выбору ее параметров H и δ. Введем функцию

$$R(a,b)=M\sum_{t=a+1}^{b}\alpha(t)r(t),\quad a<b,$$

определяющую среднее суммарное расхождение между моделями процесса, накапливаемое на интервале наблюдений $(a,b]$. Здесь $\alpha(t)$ обозначает весовую функцию, $\alpha(\tau_i)=\alpha_i(\tau_i)$ и $\alpha(t)=1$ для остальных t. Оптимальные значения параметров H и δ будем определять, используя следующий критерий: минимизировать среднее суммарное расхождение R_1 моделей процесса, требуемое для обнаружения разладки после ее появления при заданном уровне R_0^* среднего суммарного расхождения R_0 между соседними моментами ложных тревог t_0, t_1, т.е.

$$R_1=R(\tau_\theta,\tau_\theta+t_3)\to\min \quad \text{при } R_0=R(t_0,t_1)\geq R_0^*. \qquad (9)$$

Здесь τ_θ – начало цикла наблюдений, первого после разладки процесса, а t_3 – время запаздывания в обнаружении разладки, $M\{\cdot\}$ – математическое ожидание величины $\{\cdot\}$. Критерий (9) в случае, когда расхождение моделей процесса постоянно, совпадает с известным критерием: минимизировать среднее время запаздывания в обнаружении разладки при заданном среднем времени между ложными тревогами. Точное решение задачи оптимизации параметров H и δ построенной процедуры потребовало бы конкретизации вида структурных функций A_i и B_i, а также распределения $\zeta(t)$ в уравнениях (1), (2). Поэтому эти параметры будем определять, используя оценки величин R_0 и R_1. При получении таких оценок будем предполагать, что заданы допустимые вероятности ложной тревоги α и пропуска сигнала β в каждом цикле.

ТЕОРЕМА I. Пусть: I) матрицы $B_1(t,x)$ и

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$B_2(t,x)$ в уравнениях (1), (2) при некоторых положительных постоянных ρ_1 и ρ_2 удовлетворяют соотношениям (в смысле квадратичных форм)

$$B_1^T(B_2B_2^T)^{-1}B_1 \le \rho_1 E,\quad B_2^T(B_1B_1^T)^{-1}B_2 \le \rho_2 E; \qquad (10)$$

2) с вероятностью единица

$$\sum_{t\ge 0} r(t) = +\infty.$$

Тогда для любой пары (α,β) $0<\alpha,\beta<1$, последовательная процедура (6), (7) с параметрами

$$H^* = \frac{1}{4}\left(\sqrt{\frac{d_1}{\alpha}}+\sqrt{\frac{d_2}{\beta}}\right)^2,\quad \delta^* = \frac{\sqrt{d_1\beta}-\sqrt{d_2\alpha}}{\sqrt{d_1\beta}+\sqrt{d_2\alpha}} \qquad (11)$$

(d_i - постоянные, задаваемые (П.4)) имеет вероятности ложной тревоги не больше, чем α, и вероятности пропуска сигнала не больше, чем β.

При этом справедливы неравенства

$$R_1 \le H^*/(1-\beta),\quad R_0 \ge H^*/\alpha. \qquad (12)$$

Доказательство теоремы I вынесено в приложение. Решение задачи оптимизации параметров процедуры дает

СЛЕДСТВИЕ 1. Значение α и β, минимизирующие $R_1^* = H^*/(1-\beta)$ при фиксированном $R_0^* = H^*/\alpha$, задаются параметрическими соотношениями

$$\alpha = \frac{u}{2\sqrt{R_0^*/d_1}},\quad \beta = \frac{1}{2}\left(1-\frac{1}{u}\right),\quad \frac{\sqrt{d_1R_0^*}}{d_2} = \frac{u^2}{(u-1)^3}.$$

СЛЕДСТВИЕ 2. Если в уравнениях (1), (2) $B_1 = B_2$, то, как следует из доказательства теоремы I, для ее справедливости достаточно конечности $M\|\xi(t)\|^2$. В этом случае процедура обнаружения разладки

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

(6), (7) совпадает с процедурой, предложенной в [3], причем

$$C_1 \equiv C_2 \equiv 0, \quad d_1 = d_2 = 4.$$

4. Обнаружение разладок процесса авторегрессии

Проиллюстрируем полученные результаты на примере скалярного процесса авторегрессии ρ -го порядка. Пусть наблюдаемый процесс описывается уравнением

$$x(t+1) = m_1 + \lambda_1^{(1)} x(t) + \ldots + \lambda_\rho^{(1)} x(t+1-\rho) + \sigma_1 \xi(t+1), \quad t < \theta \quad (13)$$

до момента θ и уравнением

$$x(t+1) = m_2 + \lambda_1^{(2)} x(t) + \ldots + \lambda_\rho^{(2)} x(t+1-\rho) + \sigma_2 \xi(t+1), \quad (14)$$

начиная с момента θ. Здесь $(m_\kappa, \lambda_1^{(\kappa)}, \ldots, \lambda_\rho^{(\kappa)}, \sigma_\kappa)$, $\kappa = 1, 2$, - постоянные вектора, не равные между собой, причем $\sigma_\kappa \neq 0$; $\xi(t)$ - последовательность независимых одинаково распределенных случайных величин с $M\xi^2(t) = 1$, $M\xi^3(t) = 0$ и $M(\xi^2(t) - 1)^2 = \mu$. В этом случае решения о наличии или отсутствии разладки принимаются в моменты τ_i, $i = 1, 2, \ldots$, определяемые по формулам (5), (6):

$$\tau_i = \min\left\{\kappa > \tau_{i-1} : \sum_{t=\tau_{i-1}+1}^{\kappa} r(t) \geq H\right\},$$

$$r(t) = \frac{1}{2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\left(a_2(t,x) - a_1(t,x)\right)^2 + \frac{1}{2}\left(\frac{\sigma_1^2}{\sigma_2^2} + \frac{\sigma_2^2}{\sigma_1^2} - 2\right), \quad (15)$$

$$a_\kappa(t,x) = m_\kappa + \lambda_1^{(\kappa)} x(t) + \ldots + \lambda_\rho^{(\kappa)} x(t+1-\rho), \quad \kappa = 1, 2.$$

Решение о разладке выносится тогда, когда статистика

$$y_i(H) = \frac{1}{H} \sum_{t=\tau_{i-1}+1}^{\tau_i} d_i(t)\left[\frac{\left(x(t+1) - a_1(t,x)\right)^2}{\sigma_1^2} - \frac{\left(x(t+1) - a_2(t,x)\right)^2}{\sigma_2^2}\right. \quad (16)$$

$$\left. + \frac{1}{2}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)\left(a_2(t,x) - a_1(t,x)\right)^2 + \frac{1}{2}\left(\frac{\sigma_1^2}{\sigma_2^2} - \frac{\sigma_2^2}{\sigma_1^2}\right)\right]$$

превысит порог δ. Для справедливости теоремы I достаточно расходимости с вероятностью единица ряда из величин (15), которую можно установить [8] с помощью леммы Бореля-Кантелли. Согласно теореме I процедура обнаружения разладки (15), (16) будет иметь вероятность ложной тревоги не больше α и вероятность пропуска сигнала не больше β, если ее параметры H и δ

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

определить по формулам (II), где

$$d_\kappa = 2\rho_\kappa \max\left(\frac{4\rho_\kappa}{1+\rho_\kappa}, \mu\right), \quad \kappa = 1,2;$$

$$\rho_1 = \frac{\sigma_1^2}{\sigma_2^2}, \quad \rho_2 = \frac{\sigma_2^2}{\sigma_1^2}.$$

Если дополнительно предположить, что процесс (I3), (I4) устойчив, т.е. корни характеристических полиномов $P_\kappa(z) = z^p - \lambda_1^{(\kappa)} z^{p-1} - \ldots - \lambda_p^{(\kappa)}$, $\kappa = 1,2$, лежат внутри единичного круга, то можно получить асимптотические формулы для статистических характеристик процедуры обнаружения разладки (I5), (I6).

ТЕОРЕМА 2. Пусть процесс (I3), (I4) устойчив. Тогда:

I) для вероятностей ложной тревоги α и пропуска сигнала β справедливы предельные соотношения

$$\alpha = \Phi\left(-(1+\delta)\sqrt{\frac{H}{D_1}}\right), \quad \beta = \Phi\left((\delta-1)\sqrt{\frac{H}{D_2}}\right), \qquad (17)$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$ - функция Лапласа, а постоянные D_i определяются формулой (П.20);

2) средние суммарные расхождения R_0 и R_1, накапливаемые соответственно между соседними моментами ложных тревог и с момента разладки до ее обнаружения, определяются равенствами

$$R_1 \sim \frac{H}{1-\beta}, \quad R_0 \sim \frac{H}{\alpha} \quad (H \to \infty). \qquad (18)$$

Доказательство теоремы 2 дано в приложении.

ЗАМЕЧАНИЕ I. В случае, когда $\sigma_1 = \sigma_2$, утверждение теоремы 2, как следует из доказательства, остается в силе при конечности лишь первых двух моментов случайных величин $\xi(t)$.

ЗАМЕЧАНИЕ 2. При доказательстве теорем I, 2 предполагалось, что мо-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

мент разладки θ не случаен. Полученные свойства решающей процедуры справедливы и для случайного момента θ с произвольным распределением, если θ не зависит от $\{\xi(t)\}$ и $x(0)$. Изменения в доказательстве теорем 1, 2, необходимые в этом случае, очевидны.

С помощью соотношений (17) по заданным вероятностям α и β можно определить параметры процедуры H и δ:

$$H = \frac{1}{4}\left(\sqrt{D_1}\,\Phi^{-1}(\alpha) + \sqrt{D_2}\,\Phi^{-1}(\beta)\right)^2, \quad \delta = \frac{\Phi^{-1}(\alpha)\sqrt{D_1} - \Phi^{-1}(\beta)\sqrt{D_2}}{\Phi^{-1}(\alpha)\sqrt{D_1} + \Phi^{-1}(\beta)\sqrt{D_2}}. \quad (19)$$

Формулы (18), (19) могут быть использованы для решения следующей задачи параметрической оптимизации: при заданном $R^* = R_0(\alpha, \beta)$ найти значения α, β, минимизирующие $R_1(\alpha, \beta)$. Пусть для простоты $\sigma_1 = \sigma_2$ (при этом условии $D_1 = D_2 = 4$) и требуется, чтобы среднее суммарное расхождение моделей процесса (13), (14), накапливаемое между двумя соседними моментами ложных тревог, было не менее, чем $R^* = 20000$. Минимизируя по α и β величину $R_1(\alpha, \beta)$, получим $\alpha^* = 0{,}0007$, $\beta^* = 0{,}291$, $R_1(\alpha^*, \beta^*) = 1976$, т.е. среднее суммарное расхождение, накапливаемое после разладки, не превышает 19,76, а вероятности ложной тревоги и пропуска сигнала в каждом цикле равны 0,0007 и 0,291 соответственно.

В заключение приведем результаты экспериментального исследования построенной процедуры обнаружения разладки. Моделировался процесс авторегрессии первого порядка

$$x(t+1) = m + \lambda\, x(t) + \sigma\, \xi(t), \quad t = 0, 1, \ldots,$$

параметры которого (m, λ, σ) изменялись в момент разладки θ со значений $(m_1, \lambda_1, \sigma_1)$ на значения $(m_2, \lambda_2, \sigma_2)$. Значения параметров указаны в таблице. Шумы $\xi(t)$ выбирались независимыми, гауссовскими с $M\xi(t) = 0$ и $M\xi^2(t) = 1$. Для каждого набора параметров было получено 50 реализаций процесса x. Момент разладки θ являлся равномерно распределенной случайной величиной на временном интервале (500, 600). (В процедуре априорная информация о θ не использовалась). В таблице даны среднее время запаздывания $t_з$ в обнаружении разладки и число ложных тревог $K_{лт}$ в 50 реализациях.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Из таблицы видно, что предлагаемая процедура позволяет обнаруживать скачкообразные изменения (в любом сочетании) среднего значения процесса, дисперсии шума и параметра λ, определяющего корреляционные свойства процесса.

Таблица

H	δ	m_1	m_2	λ_1	λ_2	σ_1	σ_2	K_{AT}	t_a
8,0	0,7	0,0	0,0	0,0	0,5	I,0	1,0	6	38,3
8,0	0,7	0,0	-0,6	0,0	-0,5	I,0	I,0	3	43,0
8,0	0,7	0,0	0,6	0,0	0,0	I,0	I,0	6	39,9
8,0	0,7	0,0	-0,6	0,0	0,0	I,0	I,0	I3	40,2
8,0	0,7	0,0	0,0	0,5	0,5	I,0	2,0	2	I5,5
I6,0	0,7	0,0	0,0	0,5	0,5	I,0	0,5	25	20,3
8,0	0,7	0,0	0,3	0,0	0,3	I,0	I,0	4	52,7
8,0	0,7	0,0	-0,3	0,0	-0,3	I,0	I,0	6	48,8
8,0	0,7	0,0	0,3	0,0	0,0	I,0	2,0	0	I2,2
I6,0	0,7	0,0	-0,3	0,0	0,0	I,0	0,5	39	I4,0
8,0	0,7	0,0	0,0	0,0	0,3	I,0	2,0	0	I2,2
I6,0	0,7	0,0	0,0	0,0	-0,3	I,0	0,5	29	I6,3
8,0	0,7	0,0	0,3	0,0	0,3	I,0	2,0	I	II,7
I6,0	0,7	0,0	0,0	0,0	0,3	I,0	0,6	8	29,4

Заключение

Построена общая процедура обнаружения разладок случайных процессов, описываемых стохастическими разностными уравнениями, которая дает возможность обнаруживать изменения в момент разладки среднего значения процесса, а также его спектральных и дисперсионных свойств. Результаты могут использоваться в динамических системах, подверженных действию помех.

Приложение

I. Доказательство теоремы I

Доказательство теоремы I проводится по схеме работы [3] и опирается на лемму I.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ЛЕММА I [3]. Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ заданы: а) неубывающая система σ-подалгебр $\{\mathcal{F}_n\}_{n\geq 0}$; б) σ-конечный марковский момент относительно $\{\mathcal{F}_n\}_{n\geq 0}$; в) $\{\eta_n\}_{n\geq 0}$ и $\{r_n\}_{n\geq 0}$ - согласованные с $\{\mathcal{F}_n\}_{n\geq 0}$ последовательности случайных величин, такие, что

$$M[\eta_{\sigma+k+1} \mid \mathcal{F}_{\sigma+k}] = 0, \quad M[\eta^2_{\sigma+k+1} \mid \mathcal{F}_{\sigma+k}] \leq r_{\sigma+k}, \quad k \geq 0, \qquad (\text{П.1})$$

причём $\sum_{k\geq 0} r(k) = +\infty$ с вероятностью единица. Тогда для момента τ_σ, определяемого по формуле

$$\tau_\sigma = \inf\{k \geq 0 : \sum_{j=0}^{k} r_{\sigma+j} \geq H\}$$

(полагаем $\inf\{\varnothing\} = +\infty$), при любом $H > 0$ имеют место соотношения

$$M\Big[\sum_{k=0}^{\tau_\sigma} \alpha(k,\sigma,\tau)\eta_{\sigma+k+1} \mid \mathcal{F}_\sigma\Big] = 0;$$
$$M\Big[\Big(\sum_{k=0}^{\tau_\sigma} \alpha(k,\sigma,\tau)\eta_{\sigma+k+1}\Big)^2 \mid \mathcal{F}_\sigma\Big] \leq H, \qquad (\text{П.2})$$

где

$$\alpha(k,\sigma,\tau) = \begin{cases} 1, & \text{если } 0 \leq k < \tau_\sigma; \\ \Big[H - \sum_{j=0}^{\tau_\sigma - 1} r_{\sigma+j}\Big] \Big/ r_{\sigma+\tau_\sigma}, & \text{если } k = \tau_\sigma; \\ 0, & \text{если } k > \tau_\sigma; \end{cases}$$

$\mathcal{F}_\sigma$ обозначает σ-алгебру событий, наблюдаемых до момента σ. С помощью леммы I ниже будет показано, что шумы ε_{ki}, определяемые (8), имеют следующие свойства:

$$M[\varepsilon_{ki}(H) \mid \mathcal{F}_{\sigma_i}] = 0, \quad M[\varepsilon^2_{ki}(H) \mid \mathcal{F}_{\sigma_i}] \leq d_k/H, \quad k = 1,2. \qquad (\text{П.3})$$

Здесь $\mathcal{F}_{\sigma_i}$ - σ-алгебра событий, наблюдаемых до момента $\sigma_i = \tau_{i-1}+1$ [4], т.е. совокупность всех событий A, для которых имеет место включение

$$A \cap [\sigma_i = \ell] \in \mathcal{F}_\ell,$$

где $\mathcal{F}_\ell$ - σ-алгебра, порожденная случайными векторами $x(0)$,

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$\xi(1),\dots\xi(\ell)$, d_1 и d_2 - постоянные, определяемые равенствами:

$$d_\kappa = \max\left(\frac{\delta(1+\varkappa_3)}{1+\rho_\kappa}\rho_\kappa^2,\ 2\rho_\kappa(\varkappa_1+4\varkappa_2)\right),\quad \kappa=1,2, \tag{П.4}$$

$$\varkappa_1 = M\|\xi(t)\xi^T(t)-E\|^2,\quad \varkappa_2=\frac{1}{2}\max_{1\le i,j\le n}\sum_{s=1}^{n}|M(\xi_i\xi_j\xi_s)(t)|,$$

$$\varkappa_3=\frac{1}{2}\max_{1\le s\le n}\sum_{i,j=1}^{n}|M(\xi_i\xi_j\xi_s)(t)|,$$

$\xi_i(t)$ - i-я координата вектора $\xi(t)$. Неравенства (П.3) позволяют с учетом (8) оценить вероятности ложной тревоги $P_0^i(H,\delta)$ и пропуска сигнала $P_1^i(H,\delta)$ в i-м цикле наблюдений

$$P_0^i(H,\delta)=P\{\varepsilon_{1i}(H)>1+\delta\}\le d_1/(1+\delta)^2H;$$
$$P_1^i(H,\delta)=P\{\varepsilon_{2i}(H)<\delta-1\,|\,\tau_{i-1}\ge\theta\}\le d_2/(1-\delta)^2H. \tag{П.5}$$

Неравенство для P_0^i получено в предположении отсутствия разладки. Соотношения (II) для параметров процедуры H и δ следуют из равенств:

$$\frac{d_1}{(1+\delta)^2H}=\alpha,\quad \frac{d_2}{(1-\delta)^2H}=\beta,$$

где α и β - допустимые вероятности ложной тревоги и пропуска сигнала. Для доказательства второго утверждения теоремы I заметим, что в силу (9)

$$R_1=N_1\cdot H,\qquad R_0=N_0\cdot H, \tag{П.6}$$

где N_0 - среднее число циклов между ложными тревогами, а N_1 - среднее число циклов запаздывания в обнаружении разладки. Используя (П.5), можно показать [3], что

$$N_0\ge 1/\alpha,\quad N_1\le 1/(1-\beta).$$

Отсюда и из (П.6) следуют неравенства (I2). Для завершения доказательства теоремы I остается проверить, что выполнены условия леммы I, обеспечивающие справедливость (П.3). Моменты τ_i, определяемые (6), являются марковскими относительно введенной системы σ-алгебр $\{\mathcal{F}_\ell\}$, а последовательности $\{r(\ell)\}$ и $\{\eta_\kappa(\ell)\}$, задаваемые (5) и (8), согласованы с этой системой. Используя (3), (5) и (8), получаем

$$M(\eta_\kappa^2(\sigma+\ell+1)\,|\,\mathcal{F}_{\sigma+\ell})=(C_\kappa(\sigma+\ell),(\xi\xi^T)(\sigma+\ell+1)-E)^2_+ \tag{П.7}$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$+ b_k(\sigma+\ell) + 4\,\|a_k(\sigma+\ell)\|^2,$$

$$b_k(\sigma+\ell) = 4\,M\left[\xi^T(\sigma+\ell+1)C_k(\sigma+\ell)(\xi\xi^T)(\sigma+\ell+1)a_k(\sigma+\ell)\,|\,\mathcal{F}_{\sigma+\ell}\right],$$

где $a_k(t) = (B_k(B_j B_j^T)^{-1}(A_2 - A_1))(t,x)$;

(U, V) обозначает скалярное произведение случайных матриц U и V:

$$(U, V) = M\left(tr\, U V^T \,|\, \mathcal{F}_{\sigma+\ell}\right).$$

Первое слагаемое в (П.7) оценивается по неравенству Шварца

$$\left(C_k(\sigma+\ell),\ (\xi\xi^T)(\sigma+\ell+1) - E\right)^2 \le \varkappa_1 \|C_k(\sigma+\ell)\|^2; \qquad (П.8)$$

$$\|C_k(\sigma+\ell)\|^2 = tr\, C_k^2(\sigma+\ell),\ \varkappa_1 = M\|(\xi\xi^T)(t) - E\|^2.$$

Второе слагаемое в (П.7) оценим с помощью элементарного неравенства

$$\sum_{i,j,s=1}^{n} \mu_{ijs}\, c_{ij}\, a_s \le \varkappa_2 \|c\|^2 + \varkappa_3 \|a\|^2,$$

где μ_{ijs}, c_{ij} и a_s – неотрицательные числа:

$$\varkappa_2 = \frac{1}{2}\max_{1\le i,j\le n}\sum_{s=1}^{n}\mu_{ijs};\ \varkappa_3 = \frac{1}{2}\max_{1\le s\le n}\sum_{i,j=1}^{n}\mu_{ijs};$$

$$\|c\|^2 = \sum_{i,j=1}^{n} c_{ij}^2;\quad \|a\|^2 = \sum_{s=1}^{n} a_s^2.$$

Записав это слагаемое в координатной форме и обозначив

$$\mu_{ijs} = |M(\xi_i\xi_j\xi_s)(t)|,\ c_{ij} = |\langle C_k(\sigma+\ell)\rangle_{ij}|,\ a_s = |\langle a_k(\sigma+\ell)\rangle_s|,$$

получаем

$$|b_k(\sigma+\ell)| \le 4\varkappa_2\|C_k(\sigma+\ell)\|^2 + 4\varkappa_3\|a_k(\sigma+\ell)\|^2; \qquad (П.9)$$

$\langle A\rangle_{ij}$ обозначает (i,j)-й элемент матрицы A. Из (П.7)-(П.9) следует, что

$$M\left[\eta_k^2(\sigma+\ell+1)\,|\,\mathcal{F}_{\sigma+\ell}\right] \le 4(1+\varkappa_3)\|a_k(\sigma+\ell)\|^2 + (\varkappa_1+4\varkappa_2)\|C_k(\sigma+\ell)\|^2. \qquad (П.10)$$

Оценим правую часть этого неравенства через $r(\sigma+\ell)$. Из условий (10) вытекают неравенства

$$(B_j B_j^T)^{-1} B_k B_k^T (B_j B_j^T)^{-1} \le \rho_k (B_j B_j^T)^{-1} \le$$

$$\le \frac{\rho_k^2}{1+\rho_k}\left[(B_1 B_1^T)^{-1} + (B_2 B_2^T)^{-1}\right],$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

которые приводят к соотношению

$$\|q_\kappa(\sigma+\ell)\|^2 \le \frac{\rho_\kappa^2}{1+\rho_\kappa} w(\sigma+\ell),$$

$$w(\sigma+\ell) = \left((A_2-A_1)^T\left[(B_1B_1^T)^{-1}+(B_2B_2^T)^{-1}\right](A_2-A_1)\right)(\sigma+\ell, x). \qquad (\text{П.II})$$

Оценим второе слагаемое в правой части (П.I0). Используя каноническое представление положительно определенных матриц $z_\kappa = B_\kappa^T(B_j B_j^T)^{-1}B_\kappa$:

$$z_\kappa = U_\kappa^T \Lambda_\kappa U_\kappa,$$

где U_κ - ортогональная, а Λ_κ - диагональная матрицы, получаем

$$c_\kappa^2(\sigma+\ell) = (z_\kappa - E)^2 = z_\kappa^{1/2}\left(z_\kappa - 2E + z_\kappa^{-1}\right) z_\kappa^{1/2} =$$

$$= U_\kappa^T \Lambda_\kappa^{1/2}\left(\Lambda_\kappa - 2E + \Lambda_\kappa^{-1}\right)\Lambda_\kappa^{1/2} U_\kappa. \qquad (\text{П.I2})$$

Из условия (I0) и диагональности матрицы Λ_κ следует неравенство

$$\Lambda_\kappa^{1/2}\left(\Lambda_\kappa - 2E + \Lambda_\kappa^{-1}\right)\Lambda_\kappa^{1/2} \le \rho_\kappa\left(\Lambda_\kappa - 2E + \Lambda_\kappa^{-1}\right).$$

Отсюда и из (П.I2) с учетом того, что $tr\, z_\kappa^{-1} = tr\, z_j$, находим

$$\|c_\kappa(\sigma+\ell)\|^2 \le \rho_\kappa\, tr\left(z_\kappa - 2E + z_\kappa^{-1}\right) = \rho_\kappa\, tr\left(z_1 + z_2 - 2E\right) \qquad (\text{П.I3})$$

Подставляя оценки (П.II) и (П.I3) в правую часть (П.I0), получаем

$$M\left[\eta_\kappa^2(\sigma+\ell+1)/\mathcal{F}_{\sigma+\ell}\right] \le \frac{4(1+\varkappa_3)\rho_\kappa^2}{1+\rho_\kappa} w(\sigma+\ell) +$$

$$+ (\varkappa_1+4\varkappa_2)\rho_\kappa\, tr\,(z_1+z_2-2E) \le \qquad (\text{П.I4})$$

$$\le \max\left(\frac{8(1+\varkappa_3)\rho_\kappa^2}{1+\rho_\kappa},\ 2\rho_\kappa(\varkappa_1+4\varkappa_2)\right) r(\sigma+\ell) =$$

$$= d_\kappa\, r(\sigma+\ell) = \widetilde{r}(\sigma+\ell).$$

В последнем неравенстве использовано определение $r(\sigma+\ell)$:

$$r(\sigma+\ell) = \frac{1}{2} w(\sigma+\ell) + \frac{1}{2} tr\,(z_1+z_2-2E).$$

Таким образом, для величин $\eta_\kappa(t)$, определяемых равенством (8), выполнены условия леммы I. Поэтому из (П.2) и (П.I4) следуют требуемые неравенства (П.3). Теорема I доказана.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

2. Доказательство теоремы 2

Для получения вероятностей ложной тревоги и пропуска сигнала в i -м цикле наблюдений $(\tau_{i-1}, \tau_i]$, учитывая (8), найдем предельное распределение нормированных шумов $\tilde{\varepsilon}_{\kappa i}(H)$, определяемых формулами:

$$\tilde{\varepsilon}_{\kappa i}(H_i) = \frac{1}{\sqrt{H_i}} \sum_{t=\tau_{i-1}+1}^{\tau_i} d_i(t)\, z(t+1), \tag{II.15}$$

$$z(t+1) = 2\left(a_2(t,x) - a_1(t,x)\right)\frac{\sigma_\kappa}{\sigma_j^2}\,\xi(t+1) + (-1)^\kappa\left(\frac{\sigma_\kappa^2}{\sigma_j^2} - 1\right)\left(\xi^2(t+1) - 1\right).$$

Для этого нам потребуется

ЛЕММА 2. Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ заданы случайные процессы $z(t)$ и $r(t)$, $t = 0, 1, \ldots$, согласованные с неубывающей системой σ-подалгебр $\{\mathcal{F}_t\}_{t\geq 0}$, для которых выполнены условия:

1^0. $M(z(t+1)/\mathcal{F}_t) = 0$, $M(z^2(t+1)/\mathcal{F}_t) \leq \gamma \cdot r(t)$, $0 < \gamma < \infty$.

2^0. Существует стационарный в узком смысле квадратично интегрируемый случайный процесс $\{\tilde{z}(t), t \geq 0\}$ такой, что разность $z(t) - \tilde{z}(t)$ сходится к нулю с вероятностью единица и в среднем квадратическом, причем $M\tilde{z}^2(t) = f > 0$.

3^0. Процесс $\tilde{z}(t)$ представляется в виде предела в среднем квадратическом m-зависимых стационарных в узком смысле случайных процессов $\{\tilde{z}_m(t), t \geq 0\}$, т. е. $\tilde{z}(t) = l.i.m.\ \tilde{z}_m(t)$ при $m \to \infty$, причем при каждом m процесс $\tilde{z}(t) - \tilde{z}_m(t)$ является стационарным в широком смысле и

$$\lim_{m\to\infty} \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\tilde{z}(t) - \tilde{z}_m(t)\right)^2 = 0 \quad \text{п. н.}$$

4^0. Для системы моментов остановки $\{\tau(H), H > 0\}$, определяемой по фор-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

муле

$$\tau(H) = \inf\left\{T \geq \ell : \sum_{t=\ell}^{T} r(t) \geq H\right\}$$

для любого $\ell = 0, 1, \ldots,$ справедливо соотношение

$$\lim_{H\to\infty} \frac{\tau(H)}{H} = h \quad \text{п.н.}$$

Тогда для любого $\ell = 0, 1, \ldots$ условное распределение случайной величины

$$\varepsilon(H,\ell) = \frac{1}{\sqrt{H}} \sum_{t=\ell}^{\tau(H)-1} z(t+1) + \frac{\alpha(H)}{\sqrt{H}} z(\tau(H)+1)$$

при условии σ-алгебры $\mathcal{F}_\ell$ является асимптотически нормальным (при $H \to \infty$) с параметрами $(0, f \cdot h)$; α находится из уравнения

$$\sum_{t=\ell}^{\tau(H)-1} r(t) + \alpha(H) r(\tau(H)) = H. \qquad (\text{П.I6})$$

Лемма 2 является видоизмененной теоремой из [9]. Проверим условия леммы 2 для процессов (I5) и (П.I5), предполагая, что наблюдаемый процесс (I3), (I4) и момента времени $t=\ell$ имеет параметры $(m_k, \lambda_1^{(k)}, \ldots, \lambda_p^{(k)}, \sigma_k)$, $k=1,2$. При этом символы вероятностей и математического ожидания будут использоваться с нижним индексом k.

Учитывая, что

$$M_k\left(z^2(t+1) \mid \mathcal{F}_t\right) = \frac{4\sigma_k^2}{\sigma_j^4}\left(a_2(t,x) - a_1(t,x)\right)^2 + \mu\left(\frac{\sigma_k^2}{\sigma_j^2} - 1\right)^2,$$

и используя (15), получаем

$$M_k\left(z^2(t+1) \mid \mathcal{F}_t\right) \leq \gamma_k \cdot r(t),$$

причем

$$\gamma_k = \begin{cases} \dfrac{4}{\sigma_1^2}, & \text{если } \sigma_1 = \sigma_2; \\ 2\max\left(\dfrac{4\sigma_k^2}{\sigma_j^4}, \mu\left(\dfrac{\sigma_k^2}{\sigma_j^2} - 1\right)\right) \Big/ \min\left(\dfrac{1}{\sigma_1^2} + \dfrac{1}{\sigma_2^2}, \dfrac{\sigma_1^2}{\sigma_2^2} + \dfrac{\sigma_2^2}{\sigma_1^2} - 2\right), & \text{если } \sigma_1 \neq \sigma_2. \end{cases}$$

В условиях 2 в качестве процессов $\widetilde{z}(t)$ и $\widetilde{z}_m(t)$ можно использовать процессы

$$\widetilde{z}(t+1) = \eta(t+1) + \frac{2\sigma_k}{\sigma_j^2}\xi(t+1)\sum_{s=0}^{\infty} \Delta^T \langle \Lambda_k^s \rangle_1 \sigma_k \xi(t-s);$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\widetilde{z}_m(t) = \eta(t) + \frac{2\sigma_\kappa}{\sigma_d^2}\,\xi(t)\sum_{s=0}^{m}\Delta^T < \Lambda_\kappa^s >_1 \sigma_\kappa\,\xi(t-s-1); \qquad (\text{П.I7})$$

где

$$\eta(t+1) = \frac{2\sigma_\kappa}{\sigma_d^2}\,\xi(t+1)\,g_\kappa + (-1)^\kappa\left(\frac{\sigma_\kappa^2}{\sigma_d^2} - 1\right)\left(\xi^2(t+1) - 1\right);$$

$$\Delta^T = (\Delta_1, \ldots, \Delta_p)^T,\ \Delta_i = \lambda_i^{(2)} - \lambda_i^{(1)},\ g_\kappa = m_2 - m_1 + m_\kappa\,\frac{\sum_{i=1}^{p}\Delta_i}{1 - \sum_{i=1}^{p}\lambda_i^{(\kappa)}};$$

$$\Lambda_\kappa = \begin{pmatrix} \lambda_1^{(\kappa)} & \cdots & \lambda_p^{(\kappa)} \\ E_{p-1} & & 0 \end{pmatrix};$$

$< A >_1$ обозначает первый столбец матрицы A. Процессы (П.I7) получены путем подстановки в (П.I5) аппроксимаций процесса x, представленного в векторной форме [I0]. Из (П.I7) следует, что

$$M_\kappa \widetilde{z}^2(t) = \frac{4\sigma_\kappa^2}{\sigma_d^4}\left(g_\kappa^2 + \Delta^T F_\kappa \Delta\right) + \mu\left(\frac{\sigma_\kappa^2}{\sigma_d^2} - 1\right)^2 = f_\kappa > 0. \qquad (\text{П.I8})$$

Проверим последнее условие леммы 2. Для этого используем асимптотическое свойство процесса авторегрессии. Представим процесс $\widetilde{x}(t)$ в векторной форме

$$\widetilde{x}(t+1) = m_\kappa' \cdot \mathbb{1} + \Lambda_\kappa \widetilde{x}(t) + \widetilde{\xi}_\kappa(t+1),$$

где

$$\widetilde{x}(t) = (x(t), \ldots, x(t+1-p))^T,\ m_\kappa' = \frac{m_\kappa}{1 - \sum_{i=1}^{p}\lambda_i^{(\kappa)}},\ \mathbb{1} = (1, \ldots, 1)^T,\ \widetilde{\xi}_\kappa(t) = \{\sigma_\kappa \xi(t+1), 0\}^T$$

Известно, что

$$\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\widetilde{x}(t)\widetilde{x}^T(t) = F_\kappa + (m_\kappa')^2\,\mathbb{1}\cdot\mathbb{1}^T \quad \text{п.н.}$$

где матрица F_κ определяется из уравнения

$$F_\kappa - \Lambda_\kappa F_\kappa \Lambda_\kappa^T = \Sigma_\kappa,\ \Sigma_\kappa = \begin{Vmatrix} \sigma_\kappa^2 & 0 & \cdots & 0 \\ 0 & \cdot & \cdots & 0 \end{Vmatrix}.$$

Заменяя теперь знаменатель отношения $\tau(H)/H$ на левую часть равенства (П.I6), получаем

$$\lim_{H\to\infty}\frac{\tau(H)}{H} = h_\kappa \quad \text{п.н.},$$

$$h_\kappa = 2\Big/\left[\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\left(g_\kappa^2 + \Delta^T F_\kappa \Delta\right) + \frac{\sigma_1^2}{\sigma_2^2} + \frac{\sigma_2^2}{\sigma_1^2} - 2\right]. \qquad (\text{П.I9})$$

Таким образом, установлено, что для процессов (I5) и (П.I5) выполнены условия леммы 2. Из леммы 2 и равенства

$$P_\kappa\left[\widetilde{\varepsilon}_{\kappa i}(H) \le x \,/\, \mathcal{F}_{\tau_{i-1}+1}\right] = \sum_{\ell\ge 0} P_\kappa\left[\varepsilon_\kappa(H,\ell) \le x \,/\, \mathcal{F}_\ell\right]\chi_{[\tau_{i-1}+1=\ell]},$$

где χ_A - индикатор события A, следует, что случайная величина $\widetilde{\varepsilon}_{\kappa i}(H_i)$ при условии σ-алгебры $\mathcal{F}_{\tau_{i-1}+1}$ также является

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

асимптотически гауссовской с теми же параметрами, что и величина $\varepsilon(H,\ell)$. Из (П.18) и (П.19) следует, что асимптотическая дисперсия D_K шума $\mathcal{E}_{K,i}(H)$ определяется формулой

$$D_K = \frac{2\left[4\frac{\sigma_K^2}{\sigma_j^4}\left(g_K^2+\Delta^T F_K \Delta\right)+\mu\left(\frac{\sigma_K^2}{\sigma_j^2}-1\right)^2\right]}{\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)\left(g_K^2+\Delta^T F_K \Delta\right)+\frac{\sigma_1^2}{\sigma_2^2}+\frac{\sigma_2^2}{\sigma_1^2}-2} \qquad (П.20)$$

Отсюда с учетом (П.5), полагая $H_i = H$, получаем соотношения (17). Формулы (18) нетрудно получить [3], используя (П.6) и учитывая, что

$$P_1\left[\varepsilon_{1,i}(H) > 1+\delta \,/\, \mathcal{F}_{\tau_{i-1}+1}\right] = \alpha;$$

$$P_2\left[\varepsilon_{2,i}(H) < \delta-1 \,/\, \mathcal{F}_{\tau_{i-1}+1}\right] = \beta.$$

Теорема 2 доказана.

Литература

1. Клигене Н., Телксниc Л. Методы обнаружения моментов изменения свойств случайных процессов. (Обзор). - АиТ, 1983, № 10, с. 5-56.

2. Статистические проблемы управления. Вильнюс, 1984, вып. 65. - 271 с.

3. Воробейчиков С.Э., Конев В.В. Последовательный метод обнаружения разладок случайных процессов рекуррентного типа. - АиТ, 1984, № 5, с. 27-38.

4. Ширяев А.Н. Статистический последовательный анализ. - М.: Наука, 1976. - 272 с.

5. Basseville M., Benveniste A. Sequential detection of abrupt changes in spectral characteristics of digital signal. - IEEE Trans. on Information Theory, 1983, v. IT-29, N.5, p. 709-724.

6. Никифоров И.В. Последовательное обнаружение изменения свойств временных рядов. - М.: Наука, 1983. - 199 с.

7. Кульбак С. Теория информации и статистика. - М.: Наука, 1967. - 408 с.

8. Конев В.В., Конева Е.С. О достаточных условиях сходимости последовательных оценок параметров нелинейных рекуррентных

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

процессов. - В кн.: Математическая статистика и её приложения, Томск: Изд-во Том .. ун-та, 1981, вып. 7, с. 82-92.

9. Конев В.В., Пергаменщиков С.М. Об оценивании числа наблюдений при последовательной идентификации параметров динамических систем. - АиТ, 1984, № 12, с. 56-63.

10. Андерсон Т. Статистический анализ временных рядов. - М.: Мир, 1976. - 760 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ $C(\alpha)$-КРИТЕРИЯ ПРИ НАРУШЕНИИ УСЛОВИЯ НЕЗАВИСИМОСТИ СЛУЧАЙНЫХ ВЕЛИЧИН

Л.С.Воробьев

I. Введение. Постановка задачи

Будем рассматривать последовательность регулярных статистических экспериментов

$$\mathcal{E}_n = \left\{ \mathcal{X}_n, \mathcal{A}_n, P_{n,\omega} \right\}, \quad n = 1,2,\dots,$$

порожденных выборкой наблюдений над стационарной в узком смысле последовательностью случайных величин $X_1, X_2, \dots$, где $\mathcal{A}_n$ - σ-алгебра, определенная на $\mathcal{X}_n$, а $P_{n,\omega}$, $\omega \in \Omega$ - семейство распределений на $\mathcal{A}_n$, относительно которого предполагается выполнение свойства контигуальности [I] распределений $\left\{ P_{n,\omega} \right\}$ и $\left\{ P_{n,\omega_n = \omega + \frac{\tau}{\sqrt{n}}} \right\}$, $n = 1,2,\dots$, $\tau = const$, $\omega_n \in \Omega$. Пусть существует (относительно меры Лебега) плотность $p(x, \xi, \theta)$ каждой из с.в. $X_1, X_2, \dots$ и совместные двумерные плотности $p_{ij}(x, y, \xi, \theta)$ пар случайных величин (X_i, X_j), зависящие от параметров $(\xi, \theta) = \omega$, $\xi \in \Xi \subset R^1$, $\theta = (\theta_1, \dots, \theta_s) \in \Theta \subset R^s$, $\Xi \times \Theta = \Omega \subset R^{s+1}$. По выборке наблюдений $X = (X_1, \dots, X_n)$ требуется проверить гипотезу $H_0 : \xi = 0$ против альтернатив $H_n : \xi = \xi_n = \frac{\tau}{\sqrt{n}}$, $\tau = const$, при мешающем параметре θ.

Наиболее исследованным критерием для проверки таких гипотез является критерий отношения правдоподобия. Но, как отмечается во многих работах, при построении статистик этого критерия в большинстве практических задач возникают трудности вычислительного характера. Для случая, когда $X_1, X_2, \dots$ - последователь-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ность независимых одинаково распределенных величин, Дж.Нейманом [2] предложен критерий, названный $C(\alpha)$, который асимптотически эквивалентен критерию отношений правдоподобия, но строится более простым способом. В силу простоты построения $C(\alpha)$-критерий получил в последнее время широкое применение для решения различных естественно-научных задач.

В настоящей статье исследуется устойчивость некоторых свойств классического $C(\alpha)$-критерия в случае, когда $X_1, X_2, \ldots$ не являются независимыми, но лишь слабо зависимы.

2. Классический $C(\alpha)$-критерий

Согласно работе [2] $C(\alpha)$-критерий для проверки гипотезы $H_0: \xi = \xi_0$ против сближающихся альтернатив $H_n: \xi_n = \xi_0 + \frac{\tau}{\sqrt{n}}$ в присутствии мешающего параметра θ основывается на статистике вида

$$Z_n(\hat{\theta}_n) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi(X_i, \hat{\theta}_n), \qquad (1)$$

где $X = (X_1, \ldots, X_n)$ - выборка наблюдений над последовательностью независимых одинаково распределенных с.в. с распределением $P_{\xi\theta}$ и плотностью $p(x, \xi, \theta)$; $\hat{\theta}_n$ - $\sqrt{n}$- состоятельная оценка [1] параметра θ ; $\psi(x, \theta)$ - крамеровская функция [1], имеющая при H_0 нулевое математическое ожидание и единичную дисперсию. Предельная мощность $C(\alpha)$-критерия вычисляется в результате двойного перехода к пределу при $n \to \infty$ и $\xi_n \to \xi_0$, связанных таким образом, что $(\xi_n - \xi_0)\sqrt{n} = \tau = const$, и равна

$$\beta(\tau, \theta) = \lim_{\substack{n\to\infty \\ \xi_n \to \xi_0}} P_{\xi\theta}\{Z_n(\hat{\theta}_n) > v_\alpha\} = 1 - \Phi(v_\alpha - \mu\tau), \qquad (2)$$

где $\Phi(\cdot)$ - стандартное нормальное распределение; v_α - критическое значение, определяемое из уравнения $1 - \Phi(v_\alpha) = \alpha$; параметр μ равен ковариации при H_0 между $\psi(X_1, \theta)$ и $\frac{\partial}{\partial \xi} \ln p(X_1, \xi, \theta)\big|_{\xi=\xi_0}$.

Асимптотически локально наиболее мощный (асимптотически оптимальный) критерий в классе $C(\alpha)$-критериев получается при специальном выборе функции $\psi(x, \theta)$ [1].

Отметим, что так как при нулевой гипотезе $H_0: \xi = \xi_0$

$$\lim_{n\to\infty} P_{\xi_0\theta}\{Z_n(\hat{\theta}_n) > v_\alpha\} = 1 - \Phi(v_\alpha) = \alpha,$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

т.е. уровень значимости достигается асимптотически при всех $\theta \in \Theta$, то $C(\alpha)$-критерии являются асимптотически подобными критериями уровня α.

Довольно полный обзор по асимптотически подобным критериям, включая класс $C(\alpha)$-критериев, сделан А.В. Бернштейном в [3].

3. Исследование устойчивости классического $C(\alpha)$-критерия

Пусть $X = (X_1, \dots, X_n)$ - выборка наблюдений над стационарной в узком смысле последовательностью слабо зависимых с.в., удовлетворяющих условию сильного перемешивания [4]. С помощью классического $C(\alpha)$-критерия со статистикой (I) проверяется гипотеза $H_0: F=0, \theta \in \Theta$ против $H_1: F>0, \theta \in \Theta$.

Найдем асимптотическое выражение для мощности $C(\alpha)$-критерия при условии сильного перемешивания с.в. $X_1, X_2, \dots$.

ТЕОРЕМА I. Пусть стационарная в узком смысле последовательность $X_1, X_2, \dots$ удовлетворяет условию сильного перемешивания с коэффициентом $\gamma(h)$ и $\gamma(h) < K h^{-(6+\varepsilon)}$, $\varepsilon > 0$, $K > 0$. Пусть $M|f(X_1, \theta)|^3 \le C < \infty$. Тогда для функции мощности $\beta_n(F, \theta)$ асимптотически оптимального $C(\alpha)$-критерия существует при $n \to \infty$, $F_n = \frac{\tau}{\sqrt{n}} \to 0$, $\tau = const$ и условии $\sqrt{n}$-состоятельности оценки мешающего параметра θ предел

$$\beta^*(\tau, \theta) = 1 - \Phi\left((\nu_\alpha - \tau\mu) \frac{\mu}{B_0} \right), \qquad (3)$$

где

$$\mu = G_{0\theta} f(X_1, \theta), \quad 0 < B_0^2 = M_{0\theta} f^2(X_1, \theta) + 2 \sum_{i=2}^{\infty} M_{0\theta}\left(f(X_1, \theta) f(X_i, \theta) \right) < \infty.$$

Доказательство этой теоремы приведено в работе [5].

СЛЕДСТВИЕ. В условиях теоремы I асимптотически достигаемый уровень значимости $C(\alpha)$-критерия равен

$$\alpha^* = 1 - \Phi\left(\frac{\nu_\alpha \mu}{B_0} \right). \qquad (4)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Аналогично можно получить выражения (3) и (4) при других условиях перемешивания с.в.

Из теоремы I и следствия следует, что при нарушении условия независимости с.в. $X_1, X_2, \ldots$ $C(\alpha)$-критерий "теряет" свойство асимптотической подобности, т.е. при $n \to \infty$ не достигается номинальный уровень значимости α.

4. Асимптотически подобный критерий для слабо зависимых случайных величин

Рассмотрим для проверки гипотезы H_0 критерий из семейства $C(\alpha)$-критериев, статистика которого имеет вид

$$Z_n(\hat{\theta}_n) = B_n^{-1}(\hat{\theta}_n) \sum_{i=1}^{n} f(X_i, \hat{\theta}_n), \qquad (5)$$

где

$$B_n^2(\theta) = \sigma_{0\theta}^2 \left(\sum_{i=1}^{n} f(X_i, \theta) \right).$$

Покажем, что этот критерий является асимптотически подобным уровня α. Будем предполагать, что при сильном перемешивании с.в. с коэффициентом $\gamma(h)$ выполняется условие $\gamma(h) < K h^{-(6+\varepsilon)}$, $\varepsilon > 0$ и $M_{\gamma\theta} |f(X_1, \theta)|^3 < \infty$, а при условии равномерно сильного перемешивания с коэффициентом $\psi(h)$ выполняется $\sum_{h=1}^{\infty} (\psi(h))^{1/2} < \infty$.

Тогда справедлива следующая

ТЕОРЕМА 2. Критерий $C(\alpha)$ со статистикой (5) и критической областью (v_α, ∞), $\Phi(v_\alpha) = 1 - \alpha$, является асимптотически подобным критерием уровня α для стационарной последовательности сильно перемешанных или равномерно сильно перемешанных с.в., и его предельная мощность при $n \to \infty$, $\tau_n = \frac{\tau}{\sqrt{n}} \to 0$ и $\sqrt{n}$-состоятельности оценки $\hat{\theta}_n$ равна

$$\beta(\tau, \theta) = 1 - \Phi(v_\alpha - \tau\lambda),$$

$$\lambda = B_{\theta}^{-1} \sigma_{0\theta}^2 f(X_1, \theta), \quad 0 < B_\theta^2 = M_{0\theta} f^2(X_1, \theta) + 2 \sum_{i=2}^{\infty} M_{0\theta}\left(f(X_1, \theta) f(X_i, \theta) \right) < \infty$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Доказательство теоремы 2 проводится по полной аналогии с доказательством теоремы I.

Теорема 2 дает критерий, который одинаково хорошо работает как для слабо зависимых величин, так и для независимых. В последнем случае построенный критерий просто совпадает с классическим $C(\alpha)$-критерием.

Литература

I. Русас Дж. Контигуальность вероятностных мер. -М.: Мир, I975,- с.256.
2. Neyman J. Optimal asymptotik tests of composite statistical hypotheses. The Harald Cramer volume.- New York: J.Wiley and Sons, I959, p.2I3-234.
3. Берштейн А.В. Асимптотически подобные критерии.- Итоги науки и техники. Теория вероятностей. Математическая статистика. Теоретическая кибернетика. -М.: ВИНИТИ, I979, т.I7, с.3-56.
4. Ибрагимов И.А., Линник Ю.В. Независимые и стационарно связанные величины.- М.: Наука, I965,- с.524.
5. Воробьев Л.С., Зуев Н.М. Исследование мощности критериев проверки гипотез о параметрах распределений стационарных случайных последовательностей. - Минск, I985, с.2I.- Рукопись деп. в ВИНИТИ.II.I2.85, № 85I8-В.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

О ПОСТРОЕНИИ ЭФФЕКТИВНЫХ РЕШАЮЩИХ ПРАВИЛ НА КОРОТКИХ ОБУЧАЮЩИХ ВЫБОРКАХ В ЗАДАЧАХ РАСПОЗНАВАНИЯ ОБРАЗОВ

А.Б.Глаз

Особенность задач распознавания образов заключается в построении решающих правил, позволяющих определять принадлежность объектов к тому или иному классу, на основе ограниченных обучающих выборок. Поскольку в практических задачах распознавания обучающие выборки обычно существенно ограничены, то возникает проблема построения таких решающих правил, которые позволяли бы минимизировать вероятность ошибочной классификации.

I. Постановка задачи

Рассмотрим множество $\mathcal{M}$ объектов X, описываемых n-мерными векторами признаков, которое может быть разбито на два непересекающихся подмножества (класса) $\mathcal{M}_1$ и $\mathcal{M}_2$ в соответствии с некоторой, априорно неизвестной, общностью для объектов одного класса признаков x_i, $i \in [1:n_1]$, $n_1 < n$. Примем, что из $\mathcal{M}_1$ и $\mathcal{M}_2$ делается случайная обучающая выборка объектов $X_h \in \mathcal{N}$ ограниченной длины L, $h \in [1:L]$, т.е. объекты обучающей выборки $\mathcal{N}$ появляются случайно и независимо в соответствии с некоторым распределением вероятностей $P(X)$. Для каждого объекта обучающей выборки априорно известна принадлежность к классу $\mathcal{M}_1$ или $\mathcal{M}_2$, распределение $P(X)$ неизвестно. Поскольку классы $\mathcal{M}_1$ и $\mathcal{M}_2$ не пересекаются, то на множестве $\mathcal{M}$ существует оптимальное решающее правило

$$R(X) = \begin{cases} 1, & \forall X \in \mathcal{M}_1; \\ 0, & \forall X \in \mathcal{M}_2. \end{cases} \qquad (1)$$

Так как для построения решающего правила (I) может быть использовано не все множество $\mathcal{M}$, а только подмножество $\mathcal{N}$, то неизвестная характеристическая функция (1) аппроксимируется функцией $R'(X)$, построенной на обучающей выборке $\mathcal{N}$.

Поскольку выборка $\mathcal{N}$ всегда ограничена, то найти такое приближение $R'(X)$, которое гарантировало бы любую точность аппроксимации, принципиально не представляется возможным. Поэтому вели-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

орно выберем класс аппроксимирующих функций, например, параметрический $R'(X,V)$, где V – неизвестный m –мерный вектор параметров, и введем меру уклонения $R'(X,V)$ от $R(X)$ – функционал среднего риска, определяющий вероятность ошибочной классификации:

$$I(V) = \int_{Y} [R(X) - R'(X,V)]^2 P(Y)\,dY, \qquad (2)$$

где Y – $(n+1)$ –мерный вектор, значения первых n координат которого совпадают со значениями координат вектора X, а значение $(n+1)$ -й координаты равно значению $R(X)$; $P(Y)$ – плотность распределения вероятностей; Y – $(n+1)$ –мерное пространство (множество всех векторов Y).

Так как плотность распределения $P(Y)$ априори неизвестна, найти по ограниченной выборке N оптимальное решающее правило $R'(X,V^*)$, доставляющее точный минимум функционалу (2), невозможно. Именно поэтому в качестве меры уклонения обычно выбирается другой функционал – функционал эмпирического риска, определяющий частоту ошибочной классификации на выборке

$$I_э(V) = \frac{1}{L} \sum_{k=1}^{L} [R(X_k) - R'(X_k,V)]^2, \qquad (3)$$

и в заданном классе определяется решающее правило, минимизирующее эмпирический риск (3). Но при этом ниоткуда не следует, что выбранное из условия минимума эмпирического риска решающее правило будет обязательно доставлять минимум функционалу (2). Так, В.Н.Вапником было показано [1], что для того, чтобы с ростом L $\min_{V \in E^m} I_э(V)$ стремился к $\min_{V \in E^m} I(V)$, необходимо выполнение условия равномерной сходимости частот к вероятностям, поэтому в случае ограниченной емкости классов решающих правил для конкретного правила, минимизирующего эмпирический риск, может быть определен доверительный интервал для вероятности ошибочной классификации. Из определения этого интервала следовало, что в случае малой длины L обучающей выборки выбор решающего правила только из условия минимума эмпирического риска является далеко не оптимальным.

2. Алгоритм построения эффективного решающего правила на короткой обучающей выборке

Для построения решающего правила предлагается следующая процедура.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Сначала в заданном классе решающих правил $R'(X,V) = sg(V,\Phi(X))$, где $\Phi(X)$ – некоторая система функций из априорно выбранного класса $\mathcal{L}_\Phi$, определяющая структуру решающего правила, строится множество $\mathcal{R}$ решающих правил, минимизирующих эмпирический риск, т.е. каждое правило определяется из условия

$$\sum_{X\in\mathcal{M}} [R(X) - R'(X,V)]^2 \longrightarrow \min_{\Phi(X)\in\mathcal{L}_\Phi} \min_{V\in E^m} . \quad (4)$$

При этом равновероятно может быть образована любая из возможных структур $\Phi(X)$, т.е. решающие правила, минимизирующие эмпирический риск, генерируются случайно в соответствии с равномерным законом распределения вероятностей.

Далее на множестве правил $\{R^i_3(X)\} \subset \mathcal{R}$, $i \in [1:T]$, безошибочно классифицирующих обучающую выборку, определяется окончательное решающее правило, классифицирующее все объекты $X \in \mathcal{M}$, не вошедшие в обучающую выборку, следующим образом:

$$R^*(X) = sg\left[\sum_{i=1}^{T} R^i_3(X) - \frac{T}{2}\right], \quad (5)$$

где T – используемое нечетное число таких правил.

В тех случаях, когда при генерации множества $\mathcal{R}$ не удается построить правила, безошибочно классифицирующие обучающую выборку, окончательное решающее правило определяется как

$$R^{**}(X) = sg\left[\sum_{i=1}^{T} \alpha_i R^i(X) - \frac{1}{2}\right], \quad (6)$$

где α_i – вес i-го решающего правила $R^i(X)$, величина которого зависит от надежности q_i классификации этим правилом обучающей выборки следующим образом:

$$\alpha_i = q_i \left(\sum_{i=1}^{T} q_i\right)^{-1}, \quad (7)$$

а решающие правила $\{R^i(X)\} \subset \mathcal{R}$, $i \in [1:T]$ отбираются как первые T правил из множества $\mathcal{R}$, в котором в этом случае все решающие правила упорядочены по величине эмпирического риска и перенумерованы так, что первому правилу соответствует наименьшее значение величины эмпирического риска.

Такой подход к построению решающего правила на коротких обучающих выборках дает возможность получить дополнительную информацию о выборке и существенно понизить чувствительность к "плохому расположению" объектов обучающей выборки, что позволяет минимизировать вероятность ошибочной классификации. Доказывается сходимость такой процедуры и показывается ее эффективность на примере решения различных практических задач распознавания.

Литература

1. Вапник В.Н. Восстановление зависимостей по эмпирическим данным. – М.: Наука, 1979. – 448с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

АЛГОРИТМ УСТРАНЕНИЯ АНОМАЛЬНЫХ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ

В.А.Гныря, А.А.Светлаков

Одной из важнейших задач, возникающих при обработке экспериментальных данных, является задача выявления аномальных результатов измерений. В литературе есть сведения о том, что экспериментальные данные могут содержать до 5 - I0 % аномальных результатов измерений [I, 2] , которые могут дать сильное смещение при оценке параметров, особенно дисперсии. Проблеме выявления грубо искаженных (аномальных) результатов измерений посвящен ряд работ [2, 3] , однако их результаты мало пригодны для автоматизированной обработки экспериментальных данных. Автоматизация обработки результатов измерений, в том числе с использованием ЭВМ, привела к необходимости формализации и алгоритмизации процедур исключения аномальных результатов измерений из поступающего на обработку потока измерений. Ниже приводится алгоритм устранения аномальных результатов измерений из последовательности экспериментальных данных, который может быть использован при автоматизированной обработке измерений.

Сущность предлагаемого алгоритма заключается в следующем. Пусть имеется выборка $X_1, X_2 \ldots, X_n$, элементы которой являются результатами измерений некоторой переменной величины X. Здесь n - некоторое конечное натуральное число, а все значения X являются случайными величинами, подчиняющимися одному и тому же нормальному распределению с параметрами m и σ^2, значения которых неизвестны. Пусть поступило еще одно измерение X_{n+1}, необходимо проверить, является ли это измерение аномальным. Если оно аномальное, его необходимо выкинуть из ряда приведенных измерений. В противном случае X_{n+1} сохраняется для дальнейшего использования. Критерий проверки основан на статистике t, определяемой соотношением

$$t = \frac{X_{n+1} - m_n}{S_\sigma} ,$$

где $S_\sigma = \sqrt{\frac{n+1}{n}}\, S_n$; $S_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - m_n)^2}$; $m_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. (I)

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Статистика t подчиняется распределению Стьюдента, поэтому вероятность того, что мы откидываем "неаномальное измерение имеет следующий вид:

$$P(|x_{n+1} - m_n| \geq \varepsilon) = 2\int_{\varepsilon/S_\sigma}^{\infty} S(t, n-1)\,dt,$$

где $S(t,n)$ - плотность t - распределения. Существуют таблицы значений величины t_β, таких что

$$2\int_{t_\beta}^{\infty} S(t, n-1) = \beta. \qquad (2)$$

Задав достаточно малую доверительную вероятность β, с помощью таблиц находят значение порога t_β, который используется в соотношении вида

$$\frac{|x_{n+1} - m_n|}{S_\sigma} \geq t_\beta. \qquad (3)$$

Если данное неравенство выполняется, то x_{n+1} можно считать аномальным измерением и отбросить его как непригодное для дальнейшего анализа.

Формулы (I), (2) позволяют вычислить значения m_n, S_n, t_β только при фиксированном значении n, между тем в реальных условиях часто приходится сталкиваться с ситуациями, когда у нас уже есть значения m_n, S_n, t_β, и мы получаем новые результаты измерений переменной X. В данном случае необходимо использовать рекуррентные алгоритмы вычисления m_n, S_n, t_β. Для вычисления m_n и S_n^2 синтеза подобных алгоритмов не составляет особого труда: для этого достаточно записать формулы (I) для случая, когда выборка x_i, $i=\overline{1,n}$ пополнилась еще одним $(n+1)$-м значением x_{n+1}, и в правых частях полученных при этом выражений выделить значения m_n и S_n^2. Действуя указанным образом, приходим соответственно к формулам вида

$$m_n = \frac{1}{n+1}(n \cdot m_n + x_{n+1}); \quad S_{n+1}^2 = \frac{n-1}{n} S_n^2 + \frac{1}{n+1}(x_{n+1} - m_n)^2. \qquad (4)$$

Как было уже сказано выше, значение t_β находится из таблиц значений функции $t_\beta = \tau(n, \beta)$, являющейся нижним пределом интеграла (2). Использование данных таблиц и метода интерполирования позволяет отыскать значения t_β при любом фиксированном значении $n > 1$ и общепринятых в математической статистике значениях β, равных 0,05; 0,01; 0,001. Несложно придать данной процедуре и рекуррентный характер, но при его реализации необходимо

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

хранить в памяти вычислительного устройства большие таблицы многоразрядных чисел. Кроме того, поиск решения при этом требует значительных затрат времени. Однако, если заметить, что функция $\tau(n,\beta)$ при всех значениях β является гладкой и асимптотически убывающей при $n \to \infty$, то ее можно аппроксимировать достаточно простой функцией. Приведенные исследования показали, что функция t_β наиболее успешно аппроксимируется дробно-рациональной функцией вида

$$\hat{\tau}(n,\beta) = \frac{a_1(\beta) + a_2(\beta)\cdot n}{1 + a_3(\beta)\cdot n + a_4(\beta)\cdot n^2}, \qquad (5)$$

где $a_i(\beta)$, $i = \overline{1,4}$, - параметры аппроксимирующей функции, определяемые из минимума среднеквадратической погрешности аппроксимации функции $\tau(n,\beta)$. На рис. I приведены графики функций $\tau(n,\beta)$ и $\hat{\tau}(n,\beta)$ при $\beta = 0{,}05$. Значения параметров $a_i(\beta)$, где $i = \overline{1,4}$, а также величины погрешностей $\Delta\tau(\beta)$, вычисляемых по формуле

$$\Delta\tau(\beta) = \frac{\sum_{i=4}^{100} (\tau(i,\beta) - \hat{\tau}(i,\beta))^2}{96},$$

представлены в таблице.

Применение аппроксимации (5) и формул (3)-(4) позволяют получить следующий алгоритм устранения аномальных результатов измерений:

$$\frac{|X_{n+1} - m_n|}{S_n} \geqslant \sqrt{\frac{n}{n-1}} \cdot \hat{\tau}(n,\beta) \qquad (6)$$

Приведем некоторые результаты исследований рассматриваемого алгоритма устранения аномальных результатов измерений. Для проверки работоспособности алгоритма (6) реализован численный эксперимент, в котором элементы выборки X_i получались с помощью стандартной подпрограммы GAUSS. Данная подпрограмма генерирует псевдослучайные числа, распределенные по нормальному закону с заданными значениями средней величины и среднеквадратического отклонения. Использовались псевдослучайные числа с нулевым средним и единичной дисперсией. Аномальные значения генерировались также с помощью подпрограммы GAUSS. Среднее значение также равнялось нулю, но среднеквадратическое отклонение соответствовало десяти, и эти значения составляли примерно 30% выборки X_i. На рис. 2 представлены результаты работы алгоритма (6), где

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

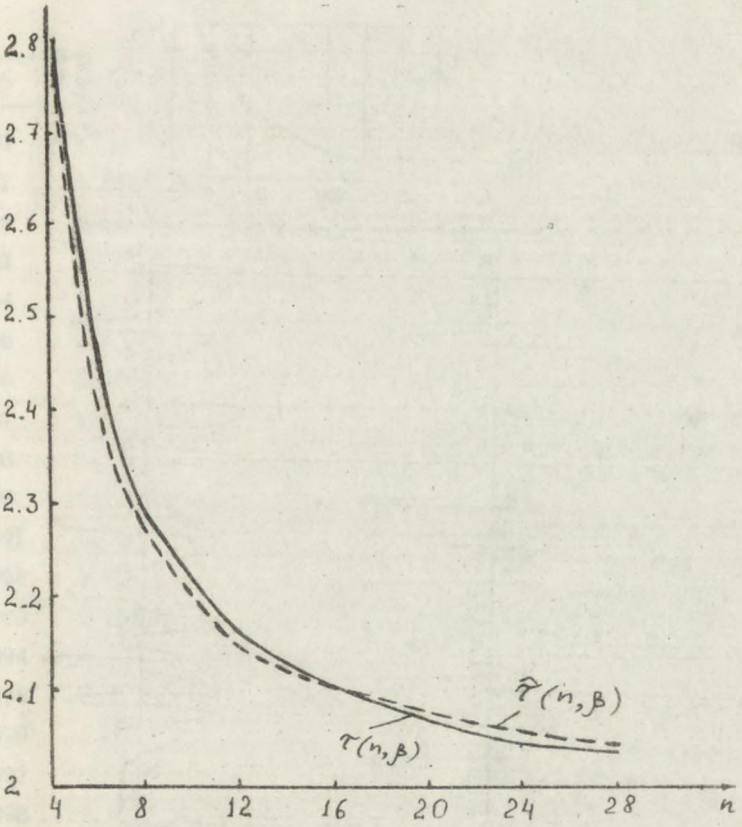

Рис. I

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Рис. 1

Рис. 2

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$1-\beta$	$\Delta\tau(\beta)$	$a_1(\beta)$	$a_2(\beta)$	$a_3(\beta)$	$a_4(\beta)$
0,90	0,0007I	-00,8I739	-2,484I5	-I,5I473	0,00005
0,9I	0,00086	-00,90395	-2,49223	-I,4750I	0,00006
0,92	0,00I04	-0I,008II	-2,50227	-I,43497	0,00007
0,93	0,00I29	-0I,I3629	-2,5I496	-I,39449	0,00008
0,94	0,00I64	-0I,29800	-2,53III	-I,35332	0,00009
0,95	0,002I5	-0I,50883	-2,55I87	-I,3I098	0,000II
0,96	0,00293	-0I,79695	-2,57934	-I,2669I	0,000I3
0,97	0,00427	-02,22I46	-2,6I73I	-I,220I8	0,000I7
0,98	0,00696	-02,93034	-2,67389	-I,I6895	0,00023
0,99	0,0I466	-04,49273	-2,772I	-I,I0878	0,00039
0,99I	0,00253	-I4,43692	-7,9378I	-3,06226	0,00026
0,992	0,00280	-I4,2I449	-7,60I82	-2,89055	0,00027
0,993	0,003I5	-I4,03299	-7,27I66	-2,72I57	0,00028
0,994	0,000358	-I3,88828	-6,94098	-2,55239	0,00029
0,995	0,004I6	-I3,82676	-6,6I5I9	-2,6I5I9	0,0003I
0,996	0,00497	-I3,85797	-6,28402	-2,2I363	0,00034
0,997	0,0062I	-I4,05588	-5,94374	-2,03634	0,00038
0,998	0,00833	-I4,56556	-5,57582	-I,84334	0,00044
0,999	0,0I335	-I5,96737	-5,I4360	-I,6I374	0,00057

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$\beta=0{,}01$, $n=\overline{1,100}$. На данном рисунке изображены значения всей выборки x_n , где аномальные результаты отмечены крестиками , а также значения выборки $\hat{x}_n$, в которую не входят значения x_n , отброшенные в результате работы алгоритма. На месте отброшенных значений у функции $\hat{x}_n$ нарисованы кружки.

Полученные результаты подтверждают работоспособность предложенного алгоритма и позволяют заключить, что он может быть применен на практике при решении задач устранения аномальных результатов измерений. При этом рекуррентность соотношений (4) и компактность формулы (5) помогут сократить трудоемкость решения задачи устранения аномальных измерений по сравнению с ее решением классическим методом. Данное обстоятельство делает реальной возможность применения предложенного алгоритма в различного рода системах автоматизации обработки измерительной информации, функционирующих в реальном масштабе времени.

ЛИТЕРАТУРА

1. Стогов Г.В., Макшанов А.В., Мусаев А.А. Устойчивые методы обработки результатов измерений (обзор). - Зарубежная радиоэлектроника, 1982, № 9, с. 28-41.

2. Микшенина Н.Г. Выявление и исключение аномальных значений (обзор). - Заводская лаборатория, 1966, № 3, с. 25-32.

3. Голубков А.С., Коротаев В.П. Отбраковка недостоверных результатов телеизмерений. - Метрология, 1973, № 2, с. 54-67.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ПОСТРОЕНИЕ ПРОЦЕДУР ПОИСКА РЕШЕНИЯ НА ОСНОВЕ АДАПТИВНЫХ СТОХАСТИЧЕСКИХ МОДЕЛЕЙ

С.Ю. Городецкий, Ю.И. Неймарк

В настоящей работе излагается единый, основанный на использовании непараметрических вероятностных моделей поведения функций подход к построению вычислительных процедур в задачах поиска решения следующего вида. В гиперкубе $\Pi \subset R^n$ задана вектор-функция $f = (f_0, f_1, \ldots, f_N)$, значениями которой являются векторы $z = (z_0, \ldots, z_N)$ из R^{N+1}. В R^{N+1} определено (может быть, неявно) множество E^*. Решением задачи являются точки $x^* \in \Pi$, для которых $f(x^*) \in E^*$. В символической форме задача приобретает вид

$$f(x) \to E^*, \quad x \in \Pi. \qquad (1)$$

Так представимы, например, задачи решения систем уравнений, систем неравенств, задачи поиска глобального минимума $f_0(x)$ при ограничениях-неравенствах $f_k(x) \leqslant 0$ $(k = \overline{1,N})$, $x \in \Pi$. В последнем случае $E^* = \{ z : z_0 \leqslant f_0^*, z_1 \leqslant 0, \ldots z_N \leqslant 0 \}$, где

$$f_0^* = \min f_0(x), \; x \in D_{f_1 \ldots f_N} = \{ x \in \Pi : f_1(x) \leqslant 0, \ldots, f_N(x) \leqslant 0 \}.$$

О функции f имеется некоторая априорная информация. Дополнительная информация извлекается в результате проведения испытаний функции f в точках $x_1, x_2, \ldots, x_m, \ldots$ области поиска Π. Результаты испытаний $I_1, I_2, \ldots, I_m, \ldots$ могут определять, например, значения и градиенты компонент функции f в точках испытаний или только значения. В настоящей работе

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

предполагается последнее, т.е. $I_i = (x_i, f_i)$, где $f_i = (f_0(x_i), \ldots, f_N(x_i))$, хотя все результаты легко обобщаются на случай других испытаний.

После проведения m испытаний имеется поисковая информация $\omega_m = \{I_1, \ldots, I_m\}$. Основной проблемой построения процедуры поиска решения (предполагается, что испытания дорогостоящи) является выбор решающего правила G_m размещения точки очередного испытания $x_{m+1} = G_m(\omega_m) \subset \Pi$. Излагаемый подход основан на двух основных принципах. Во-первых, на построении по результатам испытаний ω_m стохастических моделей поведения компонент вектор-функции f . Во-вторых, на задании локальных целей и их оптимальном удовлетворении (в рамках принятой модели функции) на данном шаге поиска. Этот подход является обобщением [1,2] . В несколько ином аспекте он изложен в [3] .

Рассмотрим конкретные способы задания стохастических моделей.

1. Пусть $h(x)$- скалярная нелипшицева функция и имеется множество результатов ее испытаний ω_m . Неопределенности значения $y = h(x)$ в произвольно взятой точке x припишем стохастический характер. Будем считать y реализацией случайной величины с функцией распределения

$$F^h_{x,m}(y) = F^h(y; x, \omega_m) = \varphi\left(\frac{y - M(x, \omega_m)}{d^0_h d_h \rho^{s_h}_m(x)}\right), \quad (2)$$

где $\varphi(u)$ - заданная дифференцируемая строго возрастающая функция распределения некоторой случайной величины со средним 0, дисперсией 1 и значениями в (0,1) при конечных u , $\rho_m(x) = \min\{\rho(x, x_i) : i = \overline{1,m}\}$, $\rho^2(x, x_i) = \sum_{j=1}^{n} (x^j - x^j_i)^2$; $M(x, \omega_m) = M^h(x, \omega_m)/d_h$;

$$M^h(x, \omega_m) = \sum_{\ell=1}^{\tau} h(x_{i_\ell}) \rho^{-2}(x, x_{i_\ell}) \Big/ \sum_{\ell=1}^{\tau} \rho^{-2}(x, x_{i_\ell}) . \quad (3)$$

Суммирование ведется по τ точкам, ближайшим к x среди точек $x_1, x_2, \ldots, x_m$; d^0_h, d_h - масштабирующие коэффициенты.

2. Если функция $h(x)$ имеет известную структуру

$$h(x) = H(q_1(x), \ldots, q_k(x))$$

(известна функция H), причем при испытаниях функции h известны результаты испытаний функций $q_1, \ldots, q_k$, то в (2) принимаем

$$M(x, \omega_m) = H(M^{q_1}(x, \omega_m), \ldots, M^{q_k}(x, \omega_m)), \quad (4)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

3. Если $h(x)$ - скалярная липшицева функция с константой L_h, тогда вместо (2) предполагаем равномерную распределенность значения $y=h(x)$ в интервале $[h_m^-(x), h_m^+(x)]$ возможных значений $h(x)$, порожденном свойством липшицивости, т.е.

$$F^h_{x,m}(y)=\max\{0;\min\{1;(y-h_m^-(x))/(h_m^+(x)-h_m^-(x))\}\}. \quad (5)$$

В стохастическую модель задачи (I) включим стохастические модели компонент f_i функции f (в виде совокупностей распределений $\{F^{f_i}_{x,m}(y): x\in\Pi\}$ типа (2), (3) или (2), (4), или (5)) и предположение о статистической независимости значений $f_i(x), f_j(x)$ при $i\neq j$.

Рассмотрим принципы выбора локальных целей поиска и построения решающего правила G_m. Вместо множества E^* в (I) введем явно задаваемые множества $E^*_{\omega_m}$ (примеры их выбора приводятся ниже) и определим промежуточные (локальные) цели как $f(x)\to E^*_{\omega_m}$. Введем функцию

$$g(x;\omega_m)=P_{\omega_m}(f(x)\in E^*_{\omega_m}), \quad (6)$$

определяющую в рамках принятой стохастической модели вероятность того, что для данного x $f(x)\in E^*_{\omega_m}$.

Будем считать, что первые m_0 точек испытаний размещены в Π равномерно. Для $m\geq m_0$ постулируем локально-оптимальное решающее правило G_m:

$$x_{m+1}=\arg\max\{g(x;\omega_m): x\in\Pi\}. \quad (7)$$

Заметим, что в случае, когда лебегова мера $\mu(E^*_{\omega_m})=0$, (6) следует заменить на

$$g(x,\omega_m)=\lim_{\varepsilon\to 0} P_{\omega_m}(f(x)\in E^*_{\omega_m,\varepsilon})/\mu(E^*_{\omega_m,\varepsilon}), \quad (8)$$

где использована специальная последовательность вложенных множеств $E^*_{\omega_m,\varepsilon_1}\supset E^*_{\omega_m,\varepsilon_2}\supset\dots$, что $\mu(E^*_{\omega_m,\varepsilon})\neq 0$, $\bigcap_\varepsilon E^*_{\omega_m,\varepsilon}=E^*_{\omega_m}$.

Рассмотрим построение конкретных процедур поиска решения.

I. Задача решения системы уравнений $f_i(x)=0$ $(i=\overline{0,N})$ с нелипшицевыми функциями. Здесь $E^*=\{0,0,\dots,0\}\subset R^{N+1}$. Полагаем $E^*_{\omega_m}=E^*$. Тогда $\mu(E^*_{\omega_m})=0$. Используя $E^*_{\omega_m,\varepsilon}=\{z: -\varepsilon\leq z_i\leq\varepsilon, (i=\overline{0,N})\}$, из (2), (8) находим

$$g(x,\omega_m)=\prod_{i=0}^{N}\frac{d}{dy}F^{f_i}_{x,m}(0) \quad (9)$$

В (9) используем (2) с $S_{f_0}=\dots=S_{f_N}=S$. Пусть Φ_K - класс не-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

прерывных в Π функций $h(x)$, для которых, во-первых, непусто множество глобальных минимумов функции $|h(x)|$ со значениями на нем функции, равными нулю, и, во-вторых, при стремлении X ко всякой точке X_h^* этого множества величина $|h(x)|$ мажорируется сверху бесконечно малой того же порядка, что $\rho^K(x, x_h^*)$.

УТВЕРЖДЕНИЕ. Для функций $f_i \in \Phi_K$ решающее правило (7), (9), (2), (3) при $S \le \min(2, K)$ порождает последовательность испытаний X_m, все предельные точки которой являются решениями системы уравнений $f_i(x) = 0 \quad (i = \overline{0, N})$.

2. Задача решения систем неравенств $a_i \le f_i(x) \le b_i \quad (i = \overline{0, N})$ $x \in \Pi$ для нелипшицевых функций; $E^* = \{z : a_i \le z_i \le b_i \quad (i = \overline{0, N})\}$. Определяем $E_{\omega_m}^* = \{z : a_i + \delta d_{s_i} \le z_i \le b_i - \delta d_{s_i}\}$ для достаточно малого $\delta > 0$. Согласно (2), (6) получаем

$$g(x, \omega_m) = \prod_{i=0}^{N} \left(F_{x,m}^{f_i}(b_i - \delta d_{s_i}) - F_{x,m}^{f_i}(a_i + \delta d_{s_i}) \right). \quad (10)$$

УТВЕРЖДЕНИЕ. Для непрерывных функций f_i, если множество решений системы неравенств имеет внутренние точки, то точка решения с помощью процедуры (7), (10) будет найдена за конечное число шагов.

3. Задача поиска глобального минимума $f_0(x)$ в Π при ограничениях-неравенствах $f_i(x) \le 0 \quad (i = \overline{1, N})$ для нелипшицевых функций предварительно сводится к задаче с одним ограничением $h(x) = \max\{f_1(x)/d_{s_1}, \ldots, f_N(x)/d_{s_N}\} \le 0$. Тогда $E^* = \{z : z_0 \le f_0^*;\ z_1 \le 0\} \in R^2$. Определим

$$f_{0m}^* = \begin{cases} \min\{f_0(x_k) : (k = \overline{1, m}),\ x_k \in D_h\}, & \text{если } \exists x_k \in D_h; \\ +\infty, & \text{если } x_k \notin D_h \ (k = \overline{1, m}) \end{cases}$$

и $E_{\omega_m}^* = \{z : z_0 \le f_{0m}^* - \eta_m d_{s_0};\ z_1 \le -\delta d_h\}$, где $\eta_m \ge \eta > 0$, $\delta > 0$. Согласно (2), (5)

$$g(x, \omega_m) = F_{x,m}^{f_0}(f_{0m}^* - \eta_m d_{s_0}) F_{x,m}^{h}(-\delta d_h), \quad (11)$$

причем в (11), в выражении (2) для $F_{x,m}^{h}$ используется представление вида (4) с

$$M(x, \omega_m) = \max\{M^{s_1}(x, \omega_m)/d_1, \ldots, M^{s_N}(x, \omega_m)/d_N\} \text{ и } S_h > S_{s_0}.$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Процедура поиска вида (II), (7) порождает последовательность точек испытаний, относительная концентрация которых в допустимой области D_h при $m \to \infty$ имеет более высокий порядок, чем вне этой области. Точная формулировка свойств метода приведена в [3].

4. Предыдущая задача для липшицевых с константами $L_{f_0}, \ldots, L_{f_N}$ функций может быть решена с помощью процедуры, основанной на стохастических моделях (5). При этом выбираем

$$E^*_{\omega_m} = \{z: z_0 \le f^*_{om} - \eta_m;\ z_1 \le 0, \ldots, z_N \le 0\}$$

и, используя (6), определим

$$g(x, \omega_m) = F^{f_0}_{x,m}(f^*_{om} - \eta_m) \prod_{i=1}^{N} F^{f_i}_{x,m}(0).$$

Закон изменения η_m примем в следующем виде. При заданном начальном η_{m_0}

$$\eta_{m+1} = \eta_m; \quad \eta_m := \begin{cases} \eta_m & \text{при } g(x_{m+1}, \omega_m) > 0; \\ \eta_m / 2 & \text{при } g(x_{m+1}, \omega_m) = 0. \end{cases}$$

Если одновременно выполнятся два условия $\eta_m < \varepsilon$ и $g(x_{m+1}, \omega_m) = 0$, то f^*_{om} определяет решение задачи с точностью $\varepsilon > 0$ по значению функции. Число точек испытаний, размещаемых построенной процедурой вне любой δ-окрестности множества решений, ограничено равномерно по ε.

Литература

1. Неймарк Ю.И. Динамические системы и управляемые процессы. - М.:Наука, 1978.-336с.
2. Городецкий С.Ю.,Неймарк Ю.И. О поисковых характеристиках алгоритма глобальной оптимизации с адаптивной стохастической моделью.- В кн.:Проблемы случайного поиска. Рига,1981, Вып.9,с.83-105.
3. Городецкий С.Ю. Учет ограничений в многоэкстремальной оптимизации с адаптивными стохастическими моделями.- В кн.:Труды IУ конференции молодых ученых НИИ ПМК и факультета ВМК ГГУ. Горький 1985. с. 72-90. Деп. в ВИНИТИ 10 июня 1986, № 4227-В86.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ БУТСТРЭП-МЕТОДА В ЗАДАЧАХ СТАТИСТИЧЕСКОГО АНАЛИЗА

В.В.Губарев, Е.Н.Наумова, А.А.Наумов

1. Постановка задачи

В последнее время в отечественной [1,2,3] и зарубежной [4-9] статистической литературе все чаще стали появляться работы, в которых используется новый, так называемый "бутстрэп"-метод с широким, как предполагает автор Б.Эфрон, спектром применения. И одновременно опубликован ряд статей, указывающих на серьезные проблемы применения этого метода в различных статистических задачах [3,6,9]. Поэтому цель настоящей работы - определить возможности применения бутстрэп-метода в задачах оценивания, в частности для определения точностных показателей разных статистических характеристик, оценить неблагоприятные моменты и последствия такого использования метода в задачах статистического анализа моделей регрессионного типа.

Для достижения поставленной цели решались следующие задачи:

- проверка возможности оценивания закона распределения с помощью бутстрэп-метода на "образцовых" выборках;
- исследование возможности применения бутстрэп-метода для определения точностных показателей оценок типовых характеристик - математического ожидания, дисперсии, коэффициентов асимметрии, эксцесса и корреляции, найденных по одной выборке;
- исследование возможности применения бутстрэп-метода в задачах регрессионного анализа.

2. Методика исследования

Бутстрэп-метод, предложенный Б.Эфроном в 1979 году [7,8], более детально рассмотрен в целом ряде работ [4,5,6,7,8,9].

Для исследования нами разработана обобщенная схема (рис.1), которая включает в себя три схемы изучения применимости бутстрэп-метода, для решения первой задачи предлагается первая схема исследования, для решения второй - вторая, рассчитанная на экспери-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

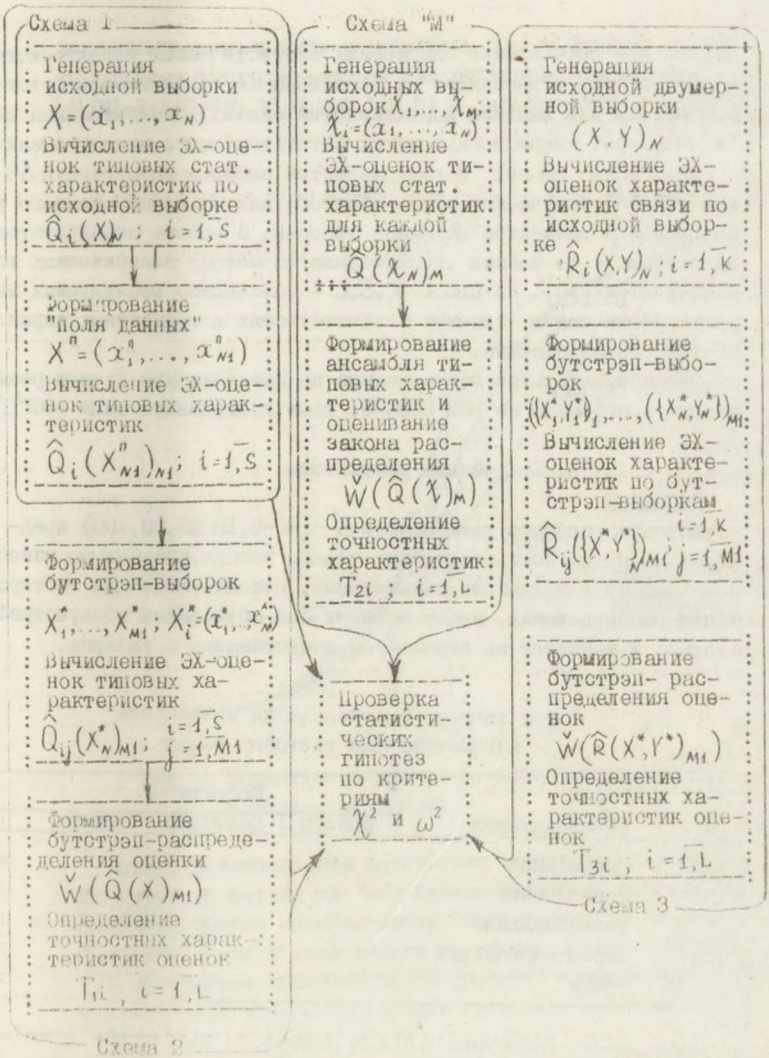

Рис.1.Обобщенная схема исследования применимости бутстрэп – метода

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ментальное определение одномерных статистических характеристик, и третья, ориентированная на исследование применимости бутстрэп-метода для анализа характеристик связи, дополнительная схема "М" отражает традиционный метод математического моделирования.

Под "полем данных" подразумевается выборка размерности $N1$, где $N1 \gg N$, полученная путем обычного или модифицированного тиражирования элементов исходной выборки. В работе используются ЭХ-оценки, т.е. оценки, полученные по методу эмпирических характеристик [10]. На рис.1 K, S, L - количество исследуемых характеристик связи, типовых статистических и точностных характеристик соответственно.

Экспериментальное исследование по предложенным схемам реализовано программно на ЭВМ БЭСМ-6, язык программирования - ФОРТРАН.

3. Полученные результаты

Задача 1

Исходная выборка размерности N (N=5,10,25,50,100) представляла набор псевдослучайных чисел, распределенных по известному закону с заданными параметрами. Используемые для тестирования распределения, выбор которых был обусловлен общеупотребимостью, и значения их параметров представлены в таблице.

Набор законов распределения и значения параметров для тестовых задач

№ п/п	Закон распределения	Параметры: сдвига	масштаба	формы
1	Нормальный	0	1	-
2	Равномерный	0	1	-
3	Логнормальный	0	1	1
4	Экспоненциальный	0	1	-
5	Релея	0	1	-

Поле данных размерности $N1$ для заданного количества бутстрэп-выборок - $M1$ ($M1$=100,250,500,1000) формировалось многократным тиражированием исходной выборки методом Монте-Карло с помощью псевдослучайного датчика псевдослучайных чисел, равномерно распределенных на участке $[0,1]$. Авторами также рассмотрены различные модификации поля данных в части его формирования. Результа-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ты тестирования в виде усредненных оценок функций распределения, полученных бутстрэп-методом для различных соотношений N/MI , где $(N/MI=2;1;0,5;0,25;0,2;0,1;0,01)$, представлены на рис.2,3,4,5,6.

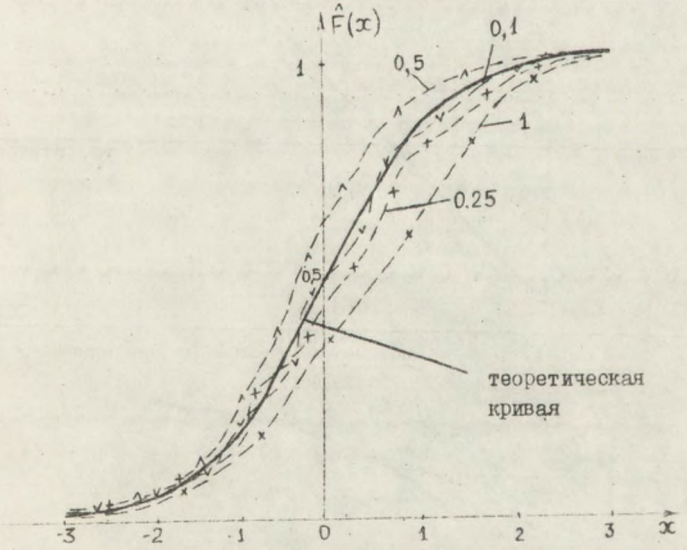

Рис.2. Усредненные оценки функции распределения нормального закона

Проверка возможности оценивания закона распределения с помощью бутстрэп-метода на "образцовых выборках" проводилась путем сравнения закона распределения "поля данных" с теоретическим по критериям χ^2 и ω^2 в зависимости от объема выборочных данных.

В дальнейшем предполагается провести аналогичные исследования с помощью непараметрических критериев согласия.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

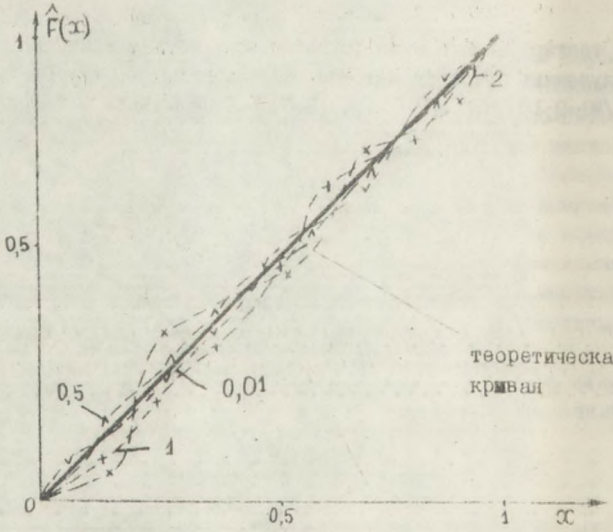

Рис.3. Усредненные оценки функции распределения равномерного закона

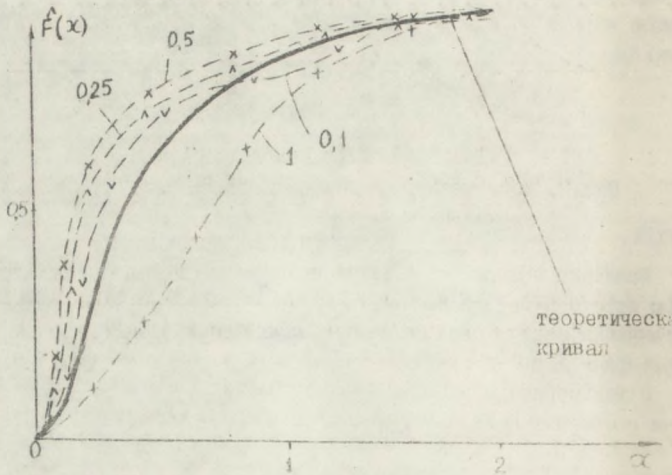

Рис.4. Усредненные оценки функции распределения логнормального закона

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

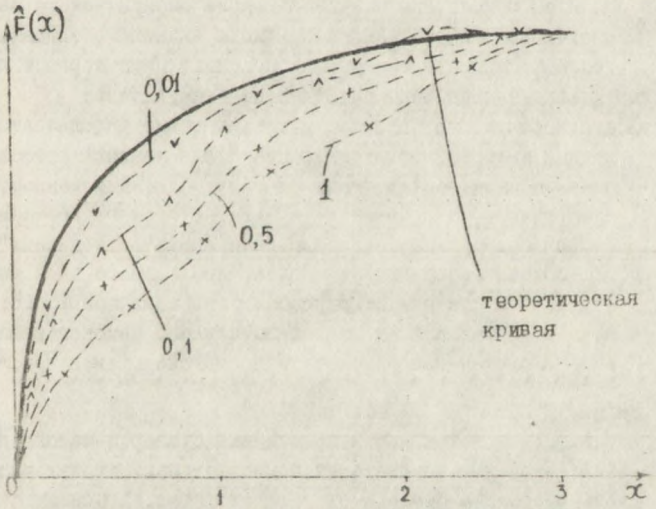

Рис.5. Усредненные оценки функции распределения экспоненциального закона

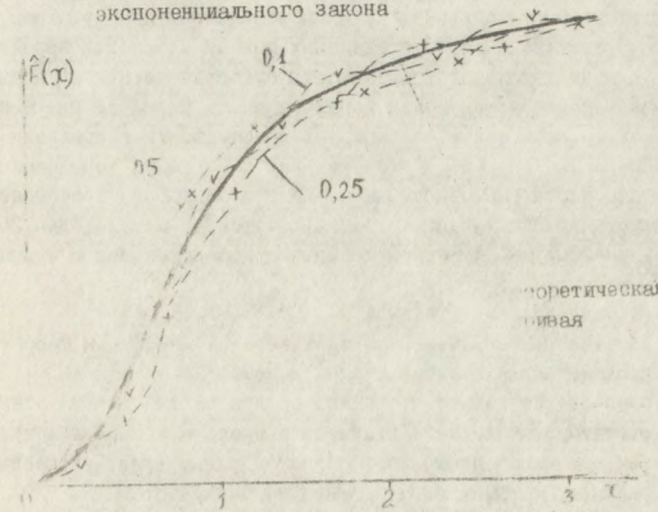

Рис.6. Усредненные оценки функции распределения закона Релея

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Основные выводы по решению первой задачи:

- в целом по группе гипотеза о согласии эмпирического закона распределения, построенного по "полю данных" с заданным теоретическим распределением, как правило, не отвергается;
- для конкретных законов гипотеза не отвергалась:
  а) для нормального закона при объеме свыше 25 элементов в исходной выборке, если генерация "поля данных" проводилась обычным способом, и свыше 5 – при модифицированном;
  б) для равномерного закона – всегда не отвергалась при $N \geqslant 10$ с модифицированным способом генерации "поля данных";
  в) для закона распределения Релея, проявившего свойства устойчивости для разных способов генерации при $N \geqslant 15$;
  г) для логнормального и экспоненциального законов распределения положительные результаты достигались лишь при $N \geqslant 50$.

Задача 2.

Исследование возможности применения бутстрэп-метода для определения точностных показателей оценок типовых характеристик проводилось с помощью Схемы 2 (см. рис.1). Основные результаты получены для нормального и логнормального законов распределения; многочисленные таблицы обобщены и представлены на рис.7,8,9,10 в виде усредненных оценок плотности вероятности.

Для исследования применимости бутстрэп-метода в задачах анализа характеристик связи использовалась Схема 3, где в качестве характеристик связи рассмотрены коэффициенты корреляции Пирсона и Спирмена. На рис.11 схематически изображены усредненные оценки плотности распределения парного и рангового коэффициентов корреляции, полученных с помощью бутстрэп-метода, для 25 пар, извлеченных из двумерной нормальной совокупности с параметрами $N(0;0;1;1;0,5)$.

Выводы по второй задаче:

- полученные результаты не противоречат известным теоретическим исследованиям и литературным данным;
- в работе показано, что закон распределения для ЭХ-оценок математического ожидания сходится в пределе к нормальному, дисперсия – к Хи-квадрат, коэффициентов асимметрии, эксцесса и корреляции – нормальному, но крайне медленно;
- скорость сходимости нелинейно зависит от соотношения $N/M1$.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

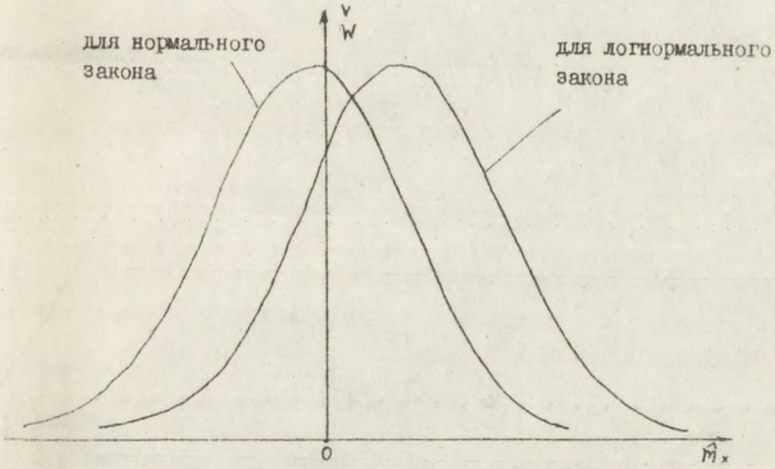

Рис.7 Усредненная оценка плотности распределения для оценок математического ожидания

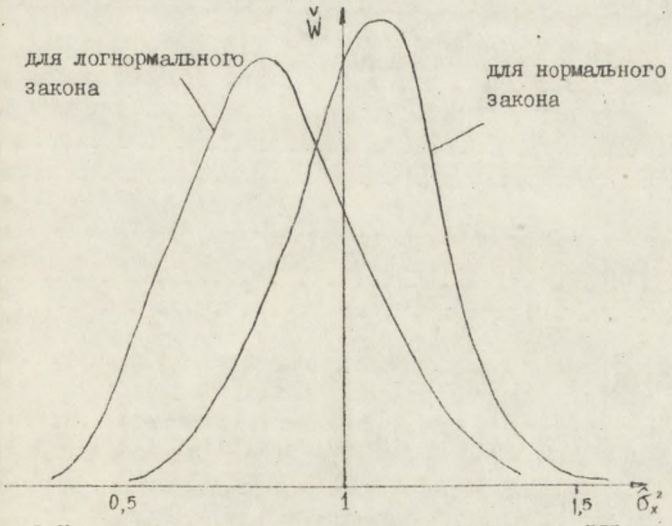

Рис.8 Усредненная оценка плотности распределения для оценок дисперсии

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

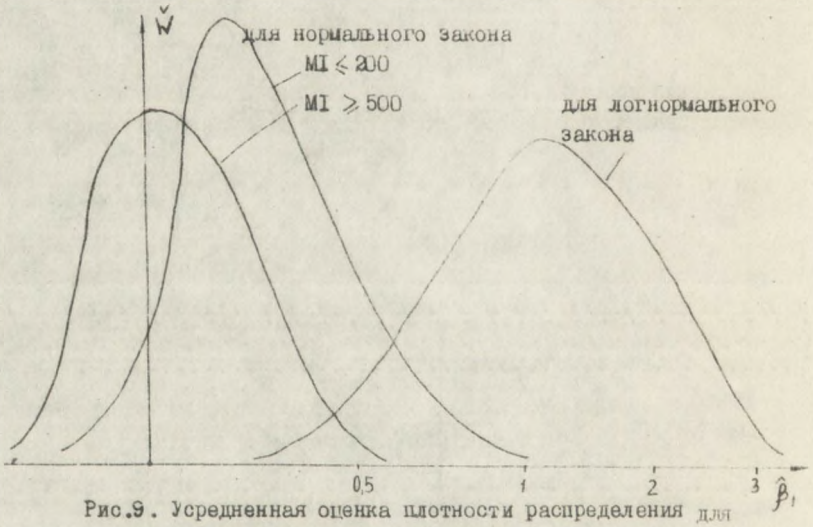

Рис.9. Усредненная оценка плотности распределения для оценок коэффициентов асимметрии

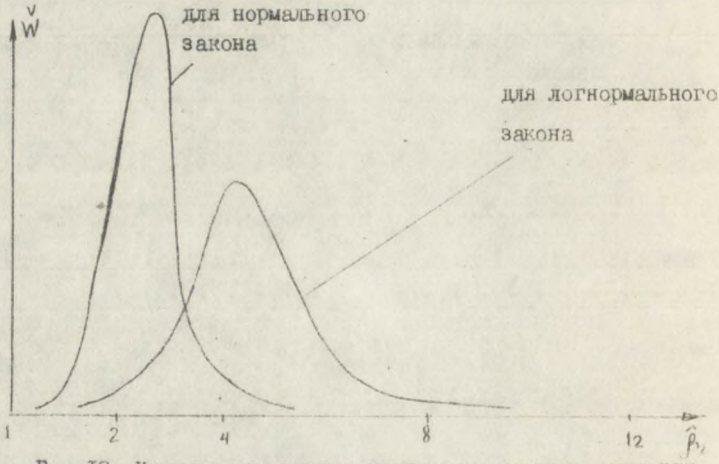

Рис.10. Усредненная оценка плотности распределения для оценок коэффициентов эксцесса

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

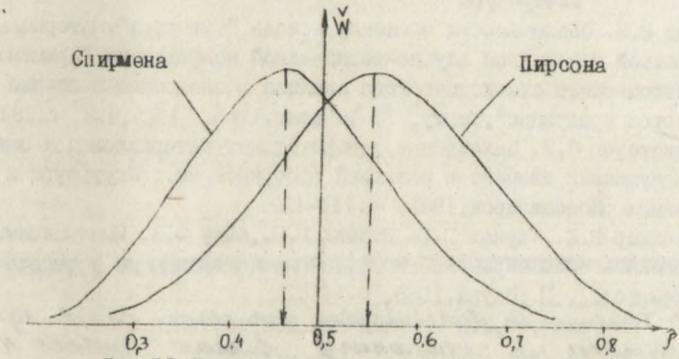

Рис.II Усредненная оценка плотности распределения для оценок коэффициентов корреляции Пирсона и Спирмена

Задача 3

В работе исследуется проблема применения бутстрэп-метода в задачах регрессионного типа как для параметрического так и для непараметрического случая. В результате теоретического исследования получены следующие выводы:

- бутстрэп-процедура в моделях регрессионного типа может ухудшать модель объекта в смысле введенного в работе критерия соответствия бутстрэп-оценки реальной оценке характеристики модели;
- даже в случае, наиболее благоприятном для бутстрэп-метода, дисперсия отклика вдвое больше исходной;
- полученные результаты очевидным образом обобщаются на модели систем одновременных уравнений и на операторы оценивания типа 2-МНК, 3-МНК и др.[11,12] ;
- упомянутые выше модели систем одновременных уравнений в плане их использования в задачах прогнозирования в части построения модели прогноза - один из вариантов использования бутстрэп-метода.

4. Заключение

Результаты работы убедительно показали, что для решения поставленных задач можно использовать бутстреп-метод, целесообразно проведение более глубокого и детального исследования возможностей его применения в других задачах статистического анализа.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Литература

I. Унт М.В. Овозможности обощения метода "бутстрэп". Материалы третьей Всесоюзной научно-технической конференции "Применение многомерного статистического анализа в экономике и оценке качества продукции", Тарту, 17-18 сент. 1985, 1985, ч.2, с.184-186.

2. Жанатаров С.У. Нахождение доверительных интервалов для оценок пропущенных элементов реальной выборки. В кн.: Структуры и анализ данных. Новосибирск, 1985, с.111-122.

3. Пинскер И.Ш., Трунов В.Г., Волчик В.М., Айду Э.А. Имитационные оценки качества решения. В кн.: Поиск зависимости и оценка погрешности. М.: Наука, 1985, с.14-32.

4. D. Freedman. On bootstrapping two-stage least-squares estimates in stationary linear models // Ann. Statist., 1984, Vol. 12, N3, 827-842

5. D. Freedman. Bootstrapping regression models // Ann. Statist., 1981, Vol. 9, N6, 1218-1228

6. S. Nathaniel. Qulms about bootstrap confidence intervals // J. Amer. Stat. Assoc., 1985, Vol. 80, N390, 360-361

7. B. Efron. Bootstrap - another look at Jackknife // Ann. Statist., 1979, Vol. 7, N1, 1-26

8. B. Efron. Bootstrap confidence intervals for a class of parametric problems // Biometrika, 1985, Vol. 72, N1, 45-58

9. D. Boos, J. Monahan. Bootstrap methods using prior information // Biometrika, 1986, Vol. 73, N1, 77-83

10. Губарев В.В. Алгоритмы статистических измерений. - М.: Энергоатомиздат, 1985. - 272 с.

11. Федоровский Ю.П., Пирогов Г.Г. Проблемы структурного оценивания в эконометрии. - М.: Статистика, 1979. - 327 с.

12. Дубровский С.А. Прикладной многомерный статистический анализ. - М.: Финансы и статистика, 1982. - 216 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

РОБАСТНОЕ ОЦЕНИВАНИЕ В МНОГОМЕРНОМ СТАТИСТИЧЕСКОМ АНАЛИЗЕ

А.М.Дубров, Н.Д.Вилегжанин, В.Б.Королев

Робастное оценивание в математической статистике появилось в связи с тем, что реальные статистические данные, как правило, не подчиняются табличным законам распределения. Особенно сильно сказывается отказ от робастного оценивания в многомерном статистическом анализе совокупности экономических показателей при изучении их на ограниченном объеме предприятий. Ограничение объема не позволяет удалять из рассмотрения предприятие, у которого значение хотя бы одного из показателей оказалось существенно отличным от остальных ("грубой ошибкой").

Исследования показали, что при обработке информации без учета робастности и с учетом получаются различные результаты, влекущие к различной экономической (содержательной) интерпретации. При этом выявлено, что при робастном оценивании исходных признаков модель оказывается более адекватной реальному процессу.

При факторном анализе, например без учета робастности, существенно меняется содержательная интерпретация полученных результатов. Это легче проследить по следующей таблице:

Признаки	Нагрузки на общие факторы							
	a_1	a_1^P	a_2	a_2^P	a_3	a_3^P	a_4	a_4^P
y_1	-0,31	0,10	0,87	0,38	0,43	0,72	0,20	0,18
y_2	0,46	0,41	0,04	-0,12	0,26	0,32	0,46	0,66
y_3	0,76	0,62	0,41	0,59	0,33	0,21	0,06	0,24

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

В таблице $\bar{a}$ – вектор, весовых коэффициентов, полученный без робастного оценивания, $\bar{a}^{p}$ – вектор, полученный при робастной оценке исходных данных. В данной задаче $a_{кр1} = 0,20; a_{кр2} = 0,50$. Так, $a_{12} = 0,87$, значит, признак y_1 должен участвовать в формировании названия второго общего фактора, $a_{12}^{p} = 0,38$, следовательно, учет робастности показывает, что предыдущее решение ошибочно, так как $0,38 < a_{кр2}$. Такое же положение наблюдается при $a_{13} = 0,43$ и $a_{13}^{p} = 0,72$. В первом случае y_1 не должен участвовать в формировании названия третьего общего фактора, а робастное оценивание показывает, что этот показатель дает большой вклад в дисперсию этого фактора, равный $0,72^2 = 0,5184$. Характерным является также то, что $a_{34} = 0,06 < a_{кр1}$, зато $a_{34}^{p} = 0,24 > a_{кр1}$ и относится к значимым весовым коэффициентам. Если же факторный или компонентный анализ используется для классификации многомерных наблюдений, более близкие результаты получаются при проверке на известных полигонных классах, статистические характеристики исходных данных были получены робастными методами.

При решении экономических задач на основе канонических корреляций более близкие результаты к известным связям получаются при робастной оценке. Кластерный анализ по исходным статистическим показателям позволяет при робастном оценивании получать более близкое к реальному разбиение на классы. Данные исследования послужили основой для создания статистического пакета прикладных программ, в котором робастное оценивание является процедурой, которая может быть использована при разных методах многомерного статистического анализа. Его особенностью является наличие программы, учитывающей наличие засорения при асимметричных распределениях. Это программа джекнайф – оценки.

Для оперативного проведения многомерного статистического анализа используется автоматизированное рабочее место (АРМ) статистика-экономиста на базе персональной ЭВМ I840. Математическое обеспечение, разработанное для АРМ позволяет проводить корреляционный, регрессионный, факторный, компонентный, дискриминантный, кластерный анализ. В пакете АРМа процедура винзоризации позволяет проводить робастное оценивание исходных признаков.

Таким образом, опыт проведения многомерного статистического анализа в экономических исследованиях показал, что пренебреже-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ние робастностью может привести к выработке рекомендаций, не адекватных требованиям реального процесса.

Литература

1. Дубров А.М. Обработка статистических данных методом главных компонент. М.Статистика, 1978. - 203с.
2. Хьюбер П. Робастность в статистике. М.Мир, 1984. - 303 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

АЛГОРИТМЫ УПРАВЛЕНИЯ ДИНАМИЧЕСКИМИ ОБЪЕКТАМИ С ЗАПАЗДЫВАНИЕМ В УСЛОВИЯХ НЕПАРАМЕТРИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ

А.А.Иванилов

В работе [1] развиваются общие принципы непараметрического подхода к решению задач управления различными, в том числе и динамическими, объектами. Принципиальное отличие данного подхода – отсутствие этапа параметризации математической модели объекта, структуры регулятора. Априорная информация для построения непараметрических систем представляет собой сведения качественного характера, что позволяет ориентироваться на учет реальных сведений о свойствах исследуемого процесса. В работе [2] приводится описание непараметрического регулятора для управления линейным динамическим объектом. В качестве модели объекта используется численная оценка его весовой функции. О помехах, действующих в каналах измерения и на сам объект предполагается, что их математические ожидания равны нулю, а дисперсии ограничены. В данной работе алгоритм управления [2] обобщается на случай, когда в линейном динамическом объекте содержится звено с запаздыванием. Величина запаздывания предполагается известной.

Постановка задачи управления

Рассмотрим линейный динамический объект, описываемый дискретной весовой функцией $h(s)$, удовлетворяющей условию физической реализуемости и содержащей звено с запаздыванием. Дискретная величина запаздывания τ считается известной. Действующие на сам объект и на устройство измерения сигналов входа-выхода помехи имеют неизвестные плотности распределения. Помехи имеют нуле-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

вые математические ожидания и ограниченную дисперсию. Начальные условия принимаются нулевые.

Пусть имеется модель объекта, представленная оценкой весовой функции $H_N(s)$, на конечном интервале дискретного времени $s=\overline{1,m}$, зависящей от параметра N. Следуя [I], оценивание весовой функции может производиться по реакции объекта на стандартные входные сигналы в экспериментальном режиме. В этом случае N - количество точек наблюдения реакции объекта на стандартный входной сигнал. При $N \to \infty$ оценка весовой функции сходится в среднеквадратическом к истинной весовой функции объекта.

Задача управления состоит в том, чтобы найти такое управляющее воздействие x_i, $i=\overline{1,m}$, реакция объекта на которое минимизировала бы среднеквадратический критерий $\mathcal{J}=M\{(y_\xi(s)-y^*(s))^2\}$, где $y^*(s)$ - желаемый выход объекта. Общая структура системы управления представлена на рис. I.

Введем случайную величину - погрешность оценивания весовой функции объекта: $\eta_s = H_N(s)-h(s)$. Тогда, используя в качестве связи входа-выхода объекта дискретное уравнение свертки, с учетом известной величины запаздывания τ, оптимальное, в смысле принятого критерия, управляющее воздействие записывается в виде

$$x^\circ_{s-\tau} = \frac{1}{H_N(1)-\eta_1}\left[H_N(1)\,x^M_{s-\tau} + \sum_{i=1}^{s-\tau-1}\eta_{s-\tau+i+1}\,x^\circ_i\right], \qquad (I)$$

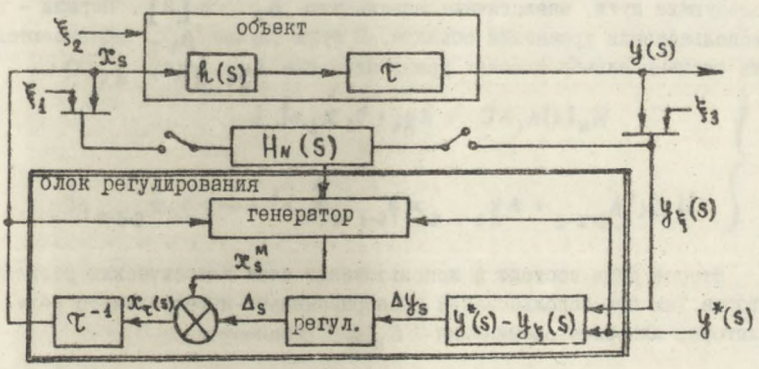

Рис. I

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

где $$x^{M}_{s-\tau} = \frac{1}{H_N(1)}\left[\frac{y^{*}(s)}{\Delta t} + \sum_{i=1}^{s-\tau-1} H_N(s-\tau-i+1)\,x^{\circ}_{i}\right]. \qquad (2)$$

Из (I),(2) следует, что для получения на s -м шаге выходной величины, минимизирующей принятый критерий, необходимо подать соответствующее управляющее воздействие на τ шагов раньше.

Выражение для оптимального $x^{\circ}_{s-\tau}$ содержит неизвестные величины $h_1,\dots,h_{s-\tau}$ и поэтому непосредственно не вычисляется. Таким образом, необходимо, чтобы алгоритм для решения поставленной задачи управления оценивал погрешности восстановления весовой функции на основе наблюдений результатов $s-1$ предыдущих шагов управления.

Алгоритм управления

Оптимальное управляющее воздействие состоит из суммы двух слагаемых. Первое зависит только от оценки весовой функции объекта и может быть вычислено по формуле (2) на основе имеющейся информации. Второе слагаемое обусловлено погрешностью восстановления весовой функции. В соответствии со структурой (I) предложен алгоритм управления, оценивающий оптимальное $x^{\circ}_{s-\tau}$:

$$\hat{x}_{s-\tau} = x^{M}_{s-\tau} + \Delta_{s-\tau}, \qquad (3)$$

где $x^{M}_{s-\tau}$ дается формулой (2), а $\Delta_{s-\tau}$ – корректирующая добавка. Для оценивания корректирующей добавки в докладе обсуждаются следующие пути, аналогичные оцениванию Δ_s в [2]. Первый – в использовании уравнений объекта. В этом случае $\Delta_{s-\tau}$ оцениваются из системы алгебраических уравнений, где $\Delta y_s = y^{*}(s) - y_{\xi}(s)$:

$$\begin{cases} H_N(1)\Delta_1\Delta t + \Delta y_\tau = h_1\hat{x}_1\Delta t; \\ \quad\cdots\cdots\cdots\cdots\cdots\cdots \\ H_N(1)\Delta_{s-\tau-1} + \Delta y_{s-1-\tau} = h_{s-1-\tau}\hat{x}_1\Delta t + \dots + h_1\hat{x}_{s-1-\tau}\Delta t. \end{cases}$$

Второй путь состоит в использовании схем классических регуляторов. Так при использовании пропорционально-интегрального регулятора, алгоритм оценивания $\Delta_{s-\tau}$ запишется

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\Delta_{s-\tau} = \Delta_{s-\tau-1} + K_1 \Delta y_{s-\tau-1} + K_2 \sum_{i=1}^{s-\tau-1} \Delta y_i .$$

Структурная схема действия алгоритма управления представлена на рис. I. Генератор осуществляет построение основного управляющего воздействия

$$x_s^M = \frac{1}{H_N(1)} \left[\frac{y^*(s)}{\Delta t} - \sum_{i=1}^{s-1} H_N(s-i+1) \hat{x}_i \right].$$

Регулятор по наблюдениям Δy_i строит оценку корректирующей добавки без учета запаздывания. Суммируя полученные составляющие, формируем сигнал $x_\tau(s)$. Упредитель "τ-1" сдвигает сигнал $x_\tau(s)$ на τ шагов назад по времени. То есть, чтобы "получить" на s-м шаге выходной сигнал $y^*(s)$, необходимо на шаге $s-\tau$ подать на вход управление $x_{s-\tau} = x_\tau(s)$.

Литература

I. Медведев А.В. Непараметрические системы адаптации.- Новосибирск: Наука, 1983.- 174 с.
2. Иванилов А.А. Управление линейными динамическими системами в условиях непараметрической неопределенности.- В кн.: Применение ЭВМ в задачах управления. Красноярск, 1985, с. 36-50.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

СИНТЕЗ РОБАСТНОЙ ПРОЦЕДУРЫ РАСПОЗНАВАНИЯ И ОПРЕДЕЛЕНИЯ ОРИЕНТАЦИИ ОБЪЕКТОВ ПО ЗАШУМЛЕННЫМ ДВУМЕРНЫМ ИЗОБРАЖЕНИЯМ В СИСТЕМЕ ТЕХНИЧЕСКОГО ЗРЕНИЯ РОБОТА

В.А.Ильин, Н.А.Харламова

Задачи распознавания и определения положения и ориентации объектов по их проекциям возникают в системах технического зрения адаптивных промышленных роботов. Подобные задачи решаются с применением самых разных подходов, в том числе и на основе применения признаков, инвариантных сдвигу и развороту. Отличительной особенностью большинства известных методов распознавания и определения ориентации плоских объектов является то, что задачи распознавания и определения ориентации решаются раздельно. Для распознавания используются одни признаки, как правило , инвариантные развороту и сдвигу, для определения ориентации - другие (см., например, /I,2/). Подобное состояние проблемы представлено в обзорах /3,5/, где отдельно освещаются методы распознавания и алгоритмы определения ориентации.

В статье предлагается подход, который дает возможность одновременно решать задачи распознавания и определения разворота плоских объектов, основанный на использовании функций общего вида, инвариантных вращению и сдвигу. Значения функций подсчитываются по координатам точек контура объекта. В качестве характеристики положения используется центр тяжести. Суть подхода состоит в том, что наряду с центром тяжести вводится еще одна характерная точка объекта, также вычисляемая по координатам точек контура. Эта точка выбирается лежащей на окружности, которая имеет центр, совпадающий с центром тяжести объекта. Характерная точка не обязательно лежит на самом объекте, как правило, она находится вне его и даже вне изображения. Важно то, что она жестко связана с объектом. Вводятся инвариантные признаки, вычисляемые по координатам точек контура и координатам характер-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ной точки. Алгоритм распознавания и определения ориентации сводится к поиску на окружности характерной точки по значениям инвариантных признаков. Найденное положение точки определяет ориентацию объекта, а значения инвариантных признаков, вычисленные в ходе поиска, определяют тип объекта. Конкретные значения признаков, соответствующие объектам, определяются на стадии обучения до начала распознавания. Количество признаков зависит от количества объектов и их сходства.

Пусть дано двумерное дискретное изображение размером $N \times N$ На изображении имеется область, соответствующая конфигурации некоторого объекта. Предполагается, что контур объекта выделен. Контур представляет собой набор точек $x_1, x_2, \dots, x_K$ с координатами (x_i^1, x_i^2), $i = 1, \dots, K$, задаваемыми в некоторой системе координат, неподвижно связанной с экраном видеосенсорной системы. Не обязательно, чтобы объект обладал связным контуром. Пусть $u = (u^1, u^2)$ некоторая точка жестко связанная с объектом, подсчитываемая по точкам контура и не меняющая свое положение относительно объекта при любых сдвигах и разворотах. Если через $u_0 = (u_0^1, u_0^2)$ обозначить координаты центра тяжести объекта, то ориентация объекта однозначно определится ориентацией вектора $u - u_0$. Задача состоит в том, чтобы найти общий подход к формированию точек u на основе последовательности точек контура.

Пусть u -точка, жестко связанная с объектом, имеющим контур $x_1, \dots, x_K$. Для определения точки u предлагается использовать скалярные функции вида

$$y = F(x_1, \dots, x_K, u), \qquad (1)$$

которые инвариантны сдвигу, вращению и перестановкам последовательности $x_1, \dots, x_K$.

Свойство инвариантности сдвигу

$$F(x_1, \dots, x_K, u) = F(x_1 + \delta, \dots, x_K + \delta, u + \delta), \qquad (2)$$

где $\delta = (\delta^1, \delta^2)$ - вектор сдвига объекта на изображении.

Свойство инвариантности вращению

$$F(x_1, \dots, x_K, u) = F(B(\varphi)x_1, \dots, B(\varphi)x_K, B(\varphi)u), \qquad (3)$$

где $B(\varphi)$ - матрица, задающая вращение объекта на угол φ в плоскости изображения; $B(\varphi)x$ - произведение матрицы на вектор.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$B(\varphi)=\begin{Vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{Vmatrix}$$

Свойство инвариантности перестановкам

$$F(x_1,\dots,x_K,u)=F(x_{i_1},\dots,x_{i_K},u), \tag{4}$$

где $(i_1,\dots,i_K)$ – произвольная перестановка чисел $1,2,\dots,K$.

Свойство (4) введено в силу следующих соображений. Из-за произвольного расположения объекта в поле зрения выделение контура может начаться с любой точки. Поэтому расположение точек в последовательности $x_1,\dots,x_K$ в общем случае может быть произвольным. Конечный же результат не должен зависеть от этого расположения, что и выражается свойством (4).

Из свойств (2), (3) при малых δ и φ элементарным образом следуют дифференциальные уравнения в частных производных первого порядка, во множестве решений которых содержатся функции вида

$$y=\Phi(\rho^2(x_1,u),\dots,\rho^2(x_K,u)), \tag{5}$$

где $\Phi(\eta_1,\dots,\eta_K)$ – произвольная дифференцируемая функция K скелярных переменных; $\rho^2(x,u)=(x^1-u^1)^2+(x^2-u^2)^2$. Так как (5) зависит от расстояний, не меняющихся при сдвигах и вращениях, то очевидно, что (5) удовлетворяет свойствам (2),(3). Среди функций вида (5) необходимо выделить функции, удовлетворяющие (4). В дальнейшем будет использоваться подкласс функций, удовлетворяющих (4), вида

$$\Phi=\sum_{i=1}^{K} f(\rho^2(x_i,u)), \tag{6}$$

где $f(\eta)$ – произвольная дифференцируемая функция одного переменного.

Два признака Φ_1 и Φ_2 вида (5) назовем ортогональными на параллелепипеде P вида $\{0\le\eta_i\le r,\ i=1,\dots,K\}$, если выполняется соотношение

$$\int_P \Phi_1(\eta_1,\dots,\eta_K)\cdot\Phi_2(\eta_1,\dots,\eta_K)\,d\eta_1\dots d\eta_K=0$$

Функции Φ_1 и Φ_2 имеющие вид

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\Phi_1(h_1,\dots,h_k)=\sum_{i=1}^{k} f_1(h_i),$$

$$\Phi_2(h_1,\dots,h_k)=\sum_{i=1}^{k} f_2(h_i),$$

ортогональны на P, если функции f_1 и f_2 ортогональны на промежутке $[0,r]$.

Точка u^* выбирается лежащей на окружности W с центром в u_0 и радиусом r, одинаковым для всех обьектов, подлежащих распознаванию.

Точку u^* для каждого из обьектов предлагается находить с использованием признаков общего вида (6). Признаки вида (6) представляют собой подкласс признаков общего вида, инвариантных вращению и сдвигу.

Признак $F(u,A)$ вида (6) назовём робастным для обьекта A, если существует такая точка $w^* \in W$, называемая точкой робастности признака $F(u,A)$, что для заданных $\beta>0$ и $\varepsilon_0>0$ величина ε, подсчитываемая по формуле

$$\varepsilon=\min_{w\in\{w: w\in W,\ \|w^*-w\|=\beta\}} |F(w^*,A)-F(w,A)| \qquad (7)$$

удовлетворяет неравенству $\varepsilon>2\varepsilon_0$ и множество

$$W_\varepsilon=\{w: w\in W,\ |F(w^*,A)-F(w,A)|\leqslant\varepsilon\}$$

содержится во множестве

$$W_*=\{w: w\in W,\ \|w^*-w\|\leqslant\beta\}.$$

Величина β характеризует требуемую точность определения разворота обьекта, ε_0 характеризует тот суммарный уровень искажений, возникающих вследствие шумов фотоприёмника и влияния дискретности, при наличии которого использование признака $F(u,A)$ обеспечивает точность β. Если уровень искажений ε_0, измеряемый величиной возможного предела изменения значений признака $F(u,A)$, удовлетворяет неравенству $2\varepsilon_0>\varepsilon$, где ε подсчитывается по (7), то, следовательно, данный признак $F(u,A)$ не обеспечивает заданную точность определения

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

разворота.

В качестве характерных точек u^* для объектов предлагается использовать точка робастности. Если система технического зрения работает с объектами A_j, $j = 1, \dots, M$, то необходимо выбирать признаки, робастные одновременно для всех объектов A_j, $j = 1, \dots, M$.

Доказывается, что, если для некоторого признака $F(u, A)$ вида (6) точки робастности существуют, то они обязательно являются точками глобального экстремума (минимума или максимума) функции $F(u, A)$, $u \in W$, либо находятся в локальной окрестности точек глобального экстремума.

Признаки $F_1(u, A)$ и $F_2(u, A)$ ортогональны на промежутке $[a, b]$, если функции $f_1(\rho)$ и $f_2(\rho)$ определяющие признаки согласно (6), являются ортогональными на $[a, b]$.

Путём экспериментов на ЭВМ установлено, что при приемлемых значениях ε_0 для широкого класса совокупностей объектов $A_j, (j = 1, \dots, M)$ точка u_j^*, выбираемая для каждого $A_j, j = 1, \dots, M$ из условия

$$F(u_j^*) = \max_{u \in W} F(u, A_j) \qquad (8)$$

является робастной, если функция $f(\rho)$, определяющая признак согласно (6), принадлежит системе полиномов, ортогональных на промежутке $[a, b]$:

$$a = \min_{x_{i,j}} \min_{u \in W} \| x_{i,j} - u \|^2,$$

$$b = \max_{x_{i,j}} \max_{u \in W} \| x_{i,j} - u \|^2,$$

где $x_{i,j}$, $i = 1, 2, \dots, K_j$ – точки контура объекта A_j, $j = \overline{1, M}$.

В частности, вышеприведённое утверждение справедливо для полиномов Лежандра и Чебышева.

До начала работы системы технического зрения с объектами совокупности A_j, $j = 1, \dots, M$ производится формирование эталонной информации об объектах этой совокупности. Для этого

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

вводится в рассмотрение множество ортогональных признаков $F_s(u,A)$, $s=1,\dots,Q$ вида (6). Достаточно, чтобы только один из них, например, $F_1(u,A)$ был робастным для всех объектов совокупности. Для объектов A_j, $j=1,\dots,M$ подсчитываются признаки $F_s(u_j^*, A_j)$, $s=1,\dots,Q$, где u_j^* выбирается согласно (2). В результате влияния дискретности изображения и шумов фотоприёмника значения признаков $F_s(u,A)$ для каждого A изменяются в некоторых пределах в зависимости от угла разворота объекта A. На стадии формирования эталонной информации находятся множества

$$T_j = \left\{ \Delta_{s,j}^{(1)} \leqslant F_s(u_j^*, A_j) \leqslant \Delta_{s,j}^{(2)},\ \begin{matrix} s=1,\dots,Q_1; \\ j=1,\dots,M. \end{matrix} \right\} \qquad (9)$$

Процедура формирования T_j при заданном ε_0 легко реализуется на ЭВМ. Количество Q_1 используемых признаков определяется требованием попарной непересекаемости множеств T_j, $j=1,\dots M$. Для каждого объекта A_j, $j=1,\dots,M$ в некотором эталонном положении запоминается угол Ψ_j ориентации вектора $u_j^* - u_{0,j}$ где $u_{0,j}$ - центр тяжести контура объекта A_j.

Некоторая особенность возникает при формировании эталонной информации для объектов, имеющих центр или оси симметрии. Для подобных объектов значения любого признака $F(u,A)$, $u \in W$ разбиваются на последовательность одинаковых секторов, в каждом из которых с точностью до порядка повторяется весь набор значений признака. Поэтому для симметричных фигур при формировании эталона достаточно просматривать значения признаков только в одном каком-либо секторе и к эталонной информации необходимо добавлять угол между векторами, определяющими положение двух устойчивых точек в любых двух соседних секторах.

Процесс распознавания и определения ориентации объектов производится следующим образом. По выделенному контуру $x_1,\dots,x_K$ наблюдаемого объекта A находится центря тяжести u_0 и точка $u^* \in W$ доставляющая максимум признаку $F_1(u,A)$. Подсчитываются значения признаков $F_s(u^*,A)$, $s=1,\dots,Q$ и согласно (9) определяется тип обьекта. По углу Ψ эталонного разворота и углу ориентации вектора $u^* - u_0$ определяется ориентация объекта.

При работе с реальными изображениями значения признаков

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

изменяются в некоторых небольших пределах в результате влияния самых разнообразных помех при формировании изображений. Как уже было сказано, количество Q используемых признаков в (4) определяется требованием попарной непересекаемости множеств T_j, $j=1,\dots,M$. Найдём аналитические выражения для величин $\Delta^{(1)}_{s,j}$ и $\Delta^{(2)}_{s,j}$, $s=1,\dots,Q$, $j=1,\dots,M$, а затем поставим и решим задачу выбора минимального числа признаков $F_s(u,A)$, $s=1,\dots,Q$, обеспечивающих полное распознавание объектов A_j, $j=1,\dots,M$ при условии, что о помехах, вызывающих искажение контура нам известно только то, что возможные колебания координат точек контура находятся в заданных, известных пределах.

Пусть задан некоторый признак

$$F_s(u,A)=\frac{1}{K}\sum_{i=1}^{K} f(\rho^2(u,x_i)).$$

Пусть задано некоторое положение объекта A, определяемое последовательностью точек контура $x_1,\dots,x_K$, и точка u^*, являющаяся точкой робастности признака $F_s(u,A)$. Величину $\Delta F_s(u^*,A)$ определим в виде

$$\Delta F_s(u^*,A)=\frac{1}{K}\sum_{i=1}^{K}\left\{\left|\frac{\partial f}{\partial x_i^1}(\rho^2(u^*,x_i))\right|\cdot|\Delta x_i^1|+\left|\frac{\partial f}{\partial x_i^2}(\rho^2(u^*,x_i))\right|\cdot|\Delta x_i^2|\right\}. \qquad (10)$$

Величина $\Delta F_s(u,A)$ является оценкой сверху предела колебания признака $F_s(u,A)$ в результате действия шумовых искажений.

Пусть задана совокупность объектов A_j, $j=1,\dots,M$. Ищется признак $F_1(u,A)$ разделяющий максимальное число объектов из заданного класса. Для этого признак $F_1(u,A)$ представляется в виде

$$F_1(u,A)=\sum_{r=1}^{m} a_r \Phi_r(u,A), \qquad (11)$$

где $\Phi_r(u,A)$ - признак вида (6) с функцией $f(\gamma)$, взятой из некоторой системы ортогональных полиномов, коэффициенты a_r, $r=1,\dots,m$ - неизвестны.

Коэффициенты a_r предлагается находить из условия максимума

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

следующей функции

$$\sum_{\substack{j,s=1 \\ j\neq s}}^{M} \left\{ \left[F_1(u_j^*, A_j) + \Delta F_1(u_j^*, A_j) - F_1(u_s^*, A_s) + \Delta F_1(u_s^*, A_s) \right]^2 + \right. \tag{I2}$$

$$\left. + \left[F_1(u_j^*, A_j) - \Delta F_1(u_j^*, A_j) - F_1(u_s^*, A_s) - \Delta F_1(u_s^*, A_s) \right]^2 \right\},$$

где u_j^*, $j=1,\dots,M$ – точка робастности признака вида (II) для обьекта A_j, величины $\Delta F_1(u_j^*, A_j)$, $j=1,\dots,M$ подсчитываются по (I0). Максимум (I2) ищется при ограничении, являющимся условием существования точки робастности для части обьектов из совокупности A_j, $j=1,\dots,M$.

В результате решения задачи оптимизации находится признак вида (II), разделяющий в общем случае часть обьектов из заданной серии. Далее рассматриваются обьекты, не разделимые признаком $F_1(u,A)$. Вводится признак $F_2(u,A)$ вида (II) и снова решается задача оптимизации по поиску коэффициентов в признаке $F_2(u,A)$, теперь уже только для обьектов, не разделимых признаком $F_1(u,A)$ и т.д.

Вышеизложенный алгоритм распознавания и определения разворота моделировался на ЭВМ для серии обьектов. Ниже приведены результаты моделирования для трёх обьектов(рис1). Для однозначного распознавания и определения ориентации использовались признаки вида

$$\Phi_1 = \frac{1}{K} \sum_{i=1}^{K} P_4(\rho^2(x_i, u)); \tag{I3}$$

$$\Phi_2 = \frac{1}{K} \sum_{i=1}^{K} P_2(\rho^2(x_i, u)), \tag{I4}$$

где $P_4(\xi)$ и $P_2(\xi)$ полиномы Лежандра четвертой и второй степени /4/, ортогональные на промежутке $[0, r^2]$ (r-радиус окружности W):

$$P_1(\xi) = \frac{2}{r^2}\xi - 1,$$

$$P_2(\xi) = \frac{1}{2}\left[3\left(\frac{2\xi}{r^2} - 1\right)^2 - 1\right],$$

$$P_4(\xi) = \frac{1}{8}\left(35\left[P_1(\xi)\right]^4 - 30\left[P_1(\xi)\right]^2 + 3\right).$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Множитель $1/\kappa$ введен с целью уменьшения влияния дискретности изображения.

Алгоритм работал с бинарными изображениями размером 50x50. Моделирование проводилось на машинах серии ЕС. Изображение объекта в поле 50x50 имитировалось программно. Поворот объектов осуществлялся также программно. После чего работал алгоритм распознавания и определения ориентации. Ошибка определения разворота определялась дискретностью расположения точек u на окружности W. При моделировании эта дискретность бралась равной одному дуговому градусу. Радиус r брался равным 64. Значение радиуса никак не связано с размером изображения, так как поиск точки u^* не связан с обработкой изображения. Точка u^* высчитывается по точкам контура.

В таблице приведены диапазоны разброса значений признаков для трех обьектов, изображенных на рисунке. Как видно из таблицы, уже для двух признаков (I3), (I4) множества T_1, T_2, T_3 не пересекаются.

№ обьекта	Диапазон изменения признака		Диапазон изменения признака	
	Δ_1^1	Δ_1^2	Δ_2^1	Δ_2^2
I	87,92	92,54	3,98	4,I0
2	72,70	75,60	3,44	4,54
3	87,I9	92,85	3,66	3,8I

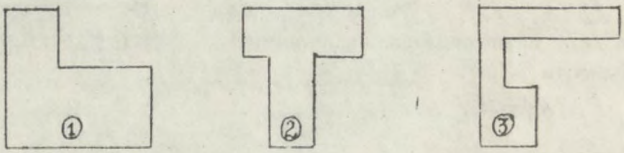

Рис. 1

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ЛИТЕРАТУРА

1. Колискор А.Ш., Лашко Е.Б., Модель Б.И., Саламандра Б.Л. Об одном регулярном методе получения признаков для идентификации плоских объектов. - Изв. АН СССР. Техническая кибернетика, 1982, № 4, с.132-135.
2. Cipovic H., Mulivojevic D., Kajmakovic Z. Planar object recognition by the computer vision method.- „Proceedings of the 3-rd International Conference on Robot Vision and Sensory Controls, Cambridge, Mass, 6-10 Nov., 1983. "Bedford, Amsterdam, 1983.
3. Петров А.А. Алгоритмическое обеспечение информационно-управляющих систем адаптивных роботов (алгоритмы технического зрения роботов). - В кн.: Итоги науки и техники. Техническая кибернетика. -М.: ВИНИТИ, 1984, т.17, с.251-291.
4. Суетин П.К. Классические ортогональные многочлены. -М.: Наука, 1976. - 327 с.
5. Жаботинский Ю.Д., Сердцев А.А. Системы технического зрения для промышленных роботов. - Зарубежная радиоэлектроника, 1985, № 12, с.23-33.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

НЕКОТОРЫЕ МОДИФИКАЦИИ НЕПАРАМЕТРИЧЕСКИХ АЛГОРИТМОВ ОПТИМИЗАЦИИ

Е.В.Ильин

Постановка задачи

Неизбежное наличие случайных помех при измерениях на реальных объектах обусловливает необходимость учета этого фактора при математической постановке задач, возникающих на этих объектах. Это касается и задачи оптимизации. Во многих случаях ее можно сформулировать в следующем виде.

Необходимо минимизировать функцию $y(x)$ типа среднего риска

$$y(x) = M_\omega \{Q(x,\omega)\} = \int_\Omega Q(x,\omega)\, dP(\omega), \qquad (1)$$

где M – символ математического ожидания, функция $Q(x,\omega)$ известна, однако распределение $P(\omega)$ не задано. Вместо этого имеется выборка статистически независимых наблюдений ω_i, $i=\overline{1,n}$ из этого распределения [1].

В этих условиях задача оптимизации формулируется в следующем виде. Пусть имеется выборка статистически независимых наблюдений $\{x_i, y_i\}$, $i=\overline{1,N}$, $x \in R^n$. Предположим, что значения $y(x)$ измерены неточно, $y_i = y(x_i) + \mu_i$, где μ_i – случайная помеха, обладающая ограниченной дисперсией и нулевым средним.

Требуется найти

$$\min_x y(x) \qquad (2)$$

при ограничениях

$$M_\omega \{\varphi_i(x,\omega)\} = 0, \quad i = \overline{1,\ell},$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$M_\omega\{\Psi_j(x,\omega)\} \le 0\,, \qquad j=\overline{1,m}\,, \tag{3}$$

где $x \in R^n$, функция $y(x)$ выпукла, а функции $\Phi(x,\omega)$, $\Psi(x,\omega)$ – неизвестны, однако имеются их значения в точках выборки $\{x_i, \Phi_i, \Psi_i\}$, $i=\overline{1,N}$, измеренные с помехой ω.

Для решения задачи (2),(3) воспользуемся методом штрафных функций. При этом будем учитывать все ограничения с помощью внешних квадратичных штрафов. Иначе говоря, решаем последовательность вспомогательных задач безусловной минимизации

$$\min_{x\in R^n} f_k(x)\,; \tag{4}$$

$$f_k(x) = y(x) + \sum_{i=1}^{\ell} K_i \Phi_i^2(x) + \sum_{i=1}^{m} K_{i+\ell}[\Psi_i^+(x)]^2\,.$$

Известно [2], что при условии непрерывной дифференцируемости функций Φ_i, Ψ_i для достаточно больших K найдется точка x_k локального минимума $f_k(x)$ в окрестности x^* локально единственного решения задачи (2), (3), и $x_k \to x^*$.

В формуле (4) знак плюс при $\Psi_i^+(x)$ означает, что суммируются только положительные значения функций Ψ_i в данной точке.

Непосредственное использование (4) для решения задачи оптимизации затруднительно, поскольку неизвестны функции $y(x)$, $\Phi_i(x)$, $\Psi_i(x)$. Вместо них применим непараметрические оценки функции регрессии

$$\hat{y}_N(x) = \sum_{i=1}^{N} y_i \prod_{j=1}^{n} \Phi\left(\frac{x-x_i^j}{c}\right) \Big/ \sum_{i=1}^{N} \prod_{j=1}^{n} \Phi\left(\frac{x-x_i^j}{c}\right)\,, \tag{5}$$

где $\Phi(\cdot)$ – финитная функция, c – коэффициент "размытости", удовлетворяющие условиям сходимости [2].

Таким образом, вместо (4) будем использовать

$$\hat{f}_k(x) = \hat{y}_N(x) + \sum_{i=1}^{\ell} K_i \hat{\Phi}_i^2(x) + \sum_{i=1}^{m} K_{i+\ell}[\hat{\Psi}_i^+(x)]^2\,, \tag{6}$$

где $\hat{\Phi}_i$ и $\hat{\Psi}_i$ – непараметрические оценки типа (5). В случае, когда функции Φ_i и Ψ_i известны, используем их в (6) в явном виде.

Можно модифицировать выражение (6). Введем переменную z

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$z_i^K = y_i + \sum_{j=1}^{k_1} K_j \varphi_j^2(x_i) + \sum_{j=1}^{m} K_{j+s} [\psi_j^+(x_i)]^2 \qquad (7)$$

и непараметрическую оценку функции регрессии

$$\hat{f}_{KN}(x) = \sum_{i=1}^{N} z_i \prod_{j=1}^{n} \Phi\left(\frac{x - x_i^j}{c}\right) \Big/ \sum_{i=1}^{N} \prod_{j=1}^{n} \Phi\left(\frac{x - x_i^j}{c}\right) . \qquad (8)$$

Таким образом, вместо (4) решаем задачу безусловной минимизации

$$\min_{x \in R^n} \hat{f}_{KN}(x) .$$

Описание непараметрического алгоритма оптимизации

Большинство итерационных методов решения задачи поиска экстремума формально описываются выражением

$$x_{s+1} = x_s + \tau(x_s) \quad , \qquad (9)$$

где x_s – решение, полученное на предыдущем итерационном шаге, $\tau(x_s)$ – некоторая функция, определяющая данный метод и его свойства. Так, например, для градиентного алгоритма $\tau(x_s) = -\gamma_s \nabla f(x_s)$. Рассмотрим непараметрический алгоритм решения задачи безусловной минимизации

$$\min_{x \in R^n} \hat{f}(x) \quad . \qquad (10)$$

Здесь $\hat{f}(x)$ – непараметрическая оценка функции регрессии типа (5), построенная на основе имеющейся статистически независимой выборки $\{x_i, f_i\}, i = \overline{1,N}$.

В качестве x_s в (9) берем

$$x_{s+1} = \sum_{i=1}^{N} x_i \varkappa_i \Big/ \sum_{i=1}^{N} \varkappa_i \quad , \qquad (11)$$

где x_i – точки исходной выборки, $\varkappa_i = \Phi\left(\frac{f_i - f_s^*}{d}\right)$, $\Phi(\cdot)$ – финитная функция, f_s^* – значение функции $\hat{f}(x_s)$, полученное на предыдущем шаге, d – коэффициент "размытости". Выражение (11) не обеспечивает получения на каждой итерации решения, доставляющего меньшее значение минимизируемой функции $f(x)$. Поэтому используем поисковый шаг $\tau(x_s)$ – второй член в (9).

Приведем для примера некоторые виды алгоритмов.

Многошаговый

$$x_{s+1} = x_s + \sum_{i=1}^{m} d_i (x_{s-i} - x_{s-i-1}) \quad ,$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

где m – число, учитывающее предысторию поиска на m шагов, α_i – нормирующие множители.

Стохастического градиента

$$x_{s+1} = x_s - \gamma_s \nabla \hat{f}(x_s) ,$$

здесь $\nabla \hat{f}(x_s)$ – стохастический градиент.

Случайный поиск

$$x_{s+1} = x_s + \Delta x_s , .$$

Δx_s – выбираем методом наилучшей пробы.

Выражение (II) определяет основную идею алгоритма оптимизации непараметрического типа.

Спишем его более подробно, применительно к задаче (4) сразу для двух модификаций, соответствующих выражениям (6) и (8). Т.е. будем использовать для оценки $f_K(x)$ в (4) выражения (6) для первой модификации и (8) – для второй.

Шаг 0. Полагаем $s = 0$ и задаемся начальными значениями коэффициентов штрафа K_i, $i = \overline{1, m+l}$.

Шаг I. $s = s + 1$.

Шаг 2. Берем произвольную начальную точку x_0^* и вычисляем по формуле (5) значение $y_0^* = \hat{y}_N(x_0^*)$, либо используя (8), $f_0^* = \hat{f}_{KN}(x_0^*)$. Для первого случая применяя (6), вычисляем $\hat{f}_0^* = \hat{f}_K(x_0^*)$.

Шаг 3. Находим оценку $\hat{x}_s$ по формуле

$$\hat{x}_s = \sum_{i=1}^{N} x_i \mathcal{Z}_i \Big/ \sum_{i=1}^{N} \mathcal{Z}_i \quad , \text{ где}$$

$$\mathcal{Z}_i = \Phi\left(\frac{y_s^* - y_i}{\alpha_s}\right) \prod_{j=1}^{m} \operatorname{sgn}\left\{\psi_j(x_i) - \varepsilon_{js}\right\} \prod_{j=1}^{l} \Phi\left(\frac{\varphi_j(x_i)}{\beta_{js}}\right)$$

для первой модификации и $\mathcal{Z}_i = \Phi\left(\frac{f_s^* - z_i}{\alpha_s}\right)$ для второй.

Здесь функция

$$\Phi(z) = \begin{cases} 0, \text{ если } |z| \geqslant 1 ; \\ 1 - z^2 , \text{ если } |z| < 1, \end{cases}$$

а ε_{js}, α_s, β_{js} – некоторые положительные, убывающие с ростом s, параметры метода.

Шаг 4. В полученной точке $\hat{x}_s$ вычисляем $\hat{f}_K(\hat{x}_s)$ по формуле (6) и $\hat{f}_{KN}(\hat{x}_s)$ по формуле (8), соответственно, для обеих модификаций. Если полученные оценки удовлетворяют условиям

$$\hat{f}_K(\hat{x}_s) < \hat{f}_K(x_{s-1}^*) , \quad \hat{f}_{KN}(\hat{x}_s) < \hat{f}_{KN}(x_{s-1}^*) , \qquad (I2)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

то идем на шаг 5. В противном случае в действие вступает второй член правой части выражения (9). Одним из способов, например, методом наилучшей пробы, в окрестности точки x^*_{s-1} выбираем такое приращение Δx_s, чтобы точка $\hat{x}_s = x^*_{s-1} + \Delta x_s$ удовлетворяла соотношению (I2).

Шаг 5. Присваиваем

$$x^*_s = \hat{x}_s, \quad y^*_s = \hat{y}_N(x_s), \quad f^*_{kN} = \hat{f}_{kN}(x_s).$$

Шаг 6. Увеличиваем коэффициенты штрафа K_i, $i = \overline{1,l+m}$, уменьшаем значения параметров $\varepsilon_{js}, \alpha_s, \beta_{js}$ и идем на Шаг 3. Поясним влияние параметров $K_i, \varepsilon_{js}, \alpha_s, \beta_{js}$ на поведение алгоритмов.

Выбор изначально достаточно больших K_i приводит к тому, что мы получим на первом шаге точку, лежащую близко к ограничениям. Это потребует частого применения методе наилучшей пробы, например, на шаге 4, поскольку при больших K_i, критерий носит овражный характер вдоль линий ограничений равенств, и задача становится плохо обусловленной.

Параметр ε_{js} характеризует возможность нарушения ограничений неравенств в ходе поиска решения, а параметры α_s и β_{js} определяют количество точек из исходной выборки, участвующих в вычислении $\hat{x}_s$.

Численное моделирование

В качестве модельной задачи условной оптимизации была рассмотрена двумерная функция Розенброка.

$$y(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2; \qquad (I3)$$

$$g(x) \equiv x_1^2 + x_2^2 - 0{,}9 = 0; \qquad (I4)$$

$$\psi(x) \equiv x_1 - 0{,}5 \le 0. \qquad (I5)$$

Точное решение этой задачи находится в точке $x = (0{,}5; 0{,}806)$.

Выборка $\{x_i\}$, $i = \overline{1,100}$ случайным образом генерировалась в области Q, удовлетворяющей условию

$$Q = \{x \mid 0 < x < 1, \ |g(x)| < 0{,}2, \ \psi(x) < 0{,}05\}.$$

Значения y_i, $i = \overline{1,100}$ вычислялись по формуле $y_i = y(x_i) + \mu_i$, где μ_i – случайная, равномерно распределенная помеха с нулевым средним. Начальные значения коэффициентов штрафа вычислялись по формулам:

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

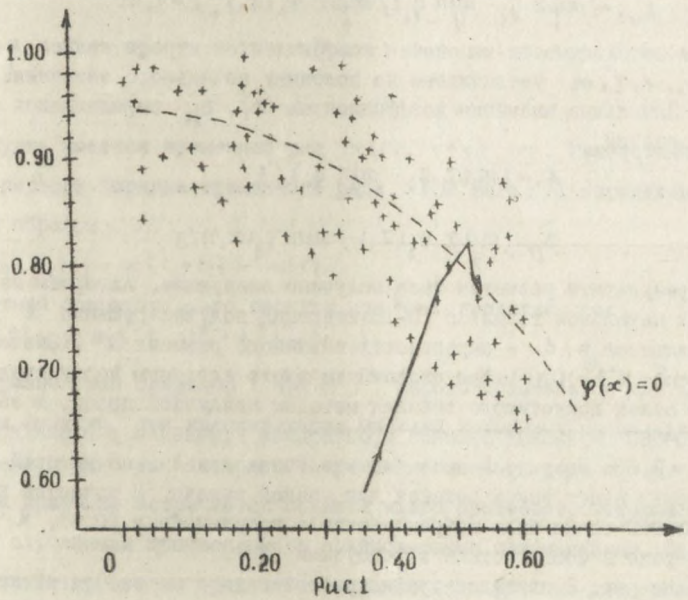

Рис.1

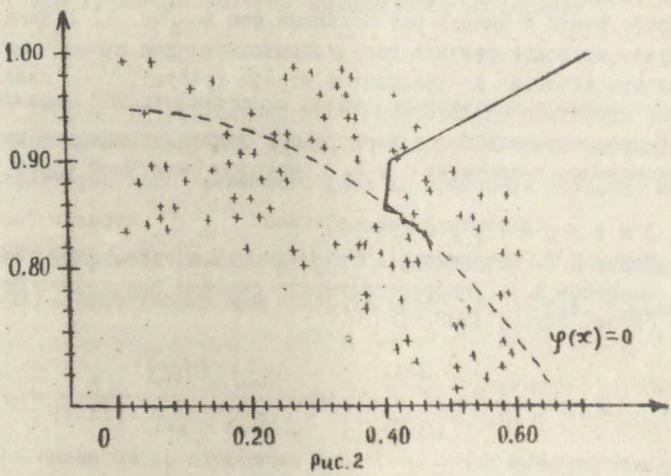

Рис.2

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$K_i=(\max_j y_j - \min_j y_j) / \max_i |\varphi_i(x_i)| \ , \ i=1,l$$
$$K_{i+l}=(\max_j y_j - \min_j y_j) / \max_t \psi_i(x_t) \ , \ i=1,m . \qquad (16)$$

При необходимости изменения коэффициентов штрафа каждое K_i, $i=1,l+m$ увеличиваем на половину начального значения.

Начальные значения коэффициентов α_s, β_{js} определялись из выражений

$$\alpha_0=(\max_i y_i - \min_i y_i)/3$$

$$\beta_{j0}=(\max_i \varphi_j(x_i) - \min_i \varphi_j(x_i))/3 .$$

В результате расчетов было получено следующее. Алгоритм из любой начальной точки за 10-15 итераций получал решение $\tilde{x}$, находящееся в ε - окрестности истинного решения x^*. Значение $\varepsilon=\|\tilde{x}-x^*\|^{1/2}=0.05$. При дальнейшем счете алгоритм переключался на поиск допустимого решения методом наилучшей пробы, и за 300-500 случайных бросков происходило уточнение решения до $\varepsilon \approx 0.03$. На рис. 1 приведена типичная кривая, по которой алгоритм ведет поиск решения для данной задачи. В качестве начальной точки была выбрана точка с координатами (0,0), K_i - выбраны в соответствии с формулами (16).

На рис. 2 приведен график работы алгоритма из другой начальной точки, для случая, когда начальные значения коэффициентов штрафа взяты в десять раз большими, чем на Рис. 1. В данном случае на поиск решения ушло большее машинное время, т.к. большие значения коэффициентов штрафа требуют, чтобы линия поиска проходила вдоль ограничения-равенства, что влечет за собой необходимость более частого обращения к процедуре выбора приемлемого приращения Δx_s методом наилучшей пробы.

Литература

1. Поляк Б.Т. Введение в оптимизацию.- М.: Наука, 1983.- 384 с.
2. Медведев А.В. Непараметрические системы адаптации.- Новосибирск, Наука, 1983.- 174 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ПРИМЕНЕНИЕ РАЗНОСТНОГО ОПЕРАТОРА ДЛЯ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ

В.А. Йоала

Пусть имеется временной ряд $x(t)$, $t = 1, \dots, M$. Разностный оператор первого порядка временного ряда $x(t)$ с шагом N определяют следующим образом:

$$\Delta(t, N) = x(t+N) - x(t).$$

Разностный оператор p-го порядка при $N=1$ определяют как

$$\Delta^p(t, 1) = \sum_{j=0}^{p} (-1)^{p-j} \binom{p}{j} x(t+j).$$

Такой разностный оператор (при $N=1$) широко применяется для интерполяции функций, при идентификации моделей временных рядов, для определения гладкости (степени) тренда временного ряда и т.д.

На практике встречается большой класс процессов, которые называются случайными процессами со стационарными приращениями. Для описания таких процессов предложено использовать функцию дисперсии приращений (ФДП), оценка которой определяется по формуле

$$D(N) = \frac{1}{2(M-N)} \sum_{t=1}^{M-N} \left[\Delta(t, N) \right]^2 , \qquad N = 1, 2, \dots .$$

С помощью ФДП относительно просто описывается кроме интенсивности временного ряда также его тренд, по этому ФДП является такой же важной характеристикой временного ряда во временной области, как в частотной области спектральная плотность.

Пусть измеряется сигнал $y(t) = z(t) + x(t)$, где $z(t)$ - полезный сигнал, а $x(t)$ - собственный шум измерительного устройства. Используя функцию

$$V(L) = \frac{1}{[M/L]-1} \sum_{i=0}^{[M/L]-2} \left[\sum_{t=Li+L+1}^{Li+2L} x(t) - \sum_{t=Li+1}^{Li+L} x(t) \right]^2 , \qquad L = 1, 2, \dots ,$$

где $[M/L]$ - целая часть отношения M/L (L - шаг интегрирования), можно определить оптимальное время интегрирования сигнала $y(t)$ с точки зрения наименьшего влияния собственного шума $x(t)$ интегрирующего

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

измерительного устройства.

Пусть $x(t)=a_0+a_1t+u_t$, $t=1,\dots,M$, где все u_t некоррелированы имеют нулевые средние и конечные дисперсии. Для оценивания коэффициентов a_0 и a_1 обычно применяют метод наименьших квадратов. Вычисления оценок $\hat{a}_0$ и $\hat{a}_1$ при этом требуют M умножений и $2M$ сложений. При использовании разностного оператора вычисления оценок $\hat{a}_0$ и $\hat{a}_1$ требуют только M сложений. При существенном уменьшении времен вычислений оценки

$$\hat{a}_1=\frac{4}{M^2}\sum_{t=1}^{M/2}\left[x(t+\frac{M}{2})-x(t)\right]=\frac{4}{M^2}\left[S_2-S_1\right];$$

$$\hat{a}_0=\frac{1}{M}\left[S_1+S_2-\frac{M(M+1)}{2}\hat{a}_1\right],$$

где $S_1=\sum_{t=1}^{M/2}x(t)$ и $S_2=\sum_{t=M/2+1}^{M}x(t)$

являются несмещенными и состоятельными оценками коэффициентов a_1 и a_0 и имеют при $M\to\infty$, соответственно, только в 4/3 и 5/4 раза большую дисперсию, чем оценки метода наименьших квадратов.

Литература

1. Андерсон Т. Статистический анализ временных рядов.- М.: Мир, 1976.- 755 с.
2. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. Вып.I.- М.: Мир, 1974.- 406 с.
3. Йоала В.А., Космачев В.М. Исследование низкочастотного шума электронных устройств.- В кн.: Первый Всемирный конгресс общества математической статистики и теории вероятностей им. Бернулли. Тезисы. Том I. М.: Наука, 1986, с.242.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

РЕГРЕССИЯ КВАНТИЛЕЙ КАК НЕПАРАМЕТРИЧЕСКАЯ ПРОЦЕДУРА ПРИКЛАДНОЙ СТАТИСТИКИ

М.П.Каминский

Одним из наиболее часто встречающихся в приложениях нарушений предпосылок нормального регрессионного анализа является несоблюдение нормальности распределения отклика. Вместе с тем во многих приложениях регрессия среднего без какой-либо потери содержательной ценности решаемой задачи может быть заменена регрессией медианы, а во многих исследованиях (надежность, биометрика и т.п.) регрессия квантилей вообще отвечает содержательной постановке решаемых задач. Далее мы увидим, что в условиях довольно мягких непараметрических ограничений на функцию распределения (ф.р.) отклика ряд известных оценок квантилей обладает свойством асимптотической нормальности, использование которой позволяет проводить нормальный регрессионный анализ для медианы или других квантилей. Заметим, что асимптотические свойства часто используются в прикладной статистике, например, широко используется асимптотическая нормальность оценок максимального правдоподобия, хотя практически проверить справедливость условий, при которых асимптотические свойства существуют, часто не удается. Напротив, рассматриваемые далее оценки квантилей обладают асимптотической нормальностью, гарантируемой в достаточно широких классах ф.р.

Наиболее известной оценкой квантили является следующая. Пусть $F(y)$ - непрерывная ф.р. случайной величины Y. $Y_1, Y_2, \ldots, Y_n$ - выборка объема n из $F(y)$. Оценкой квантили уровня

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

p Y_p является выборочная квантиль $\hat{Y}_p$:

$$\hat{Y}_p = \begin{cases} Y_{(\mu+1)}, \text{ где } \mu = [np], \text{ если } np \text{ не есть целое;} \\ \text{любому значению } Y \in [Y_{(np)}, Y_{(np+1)}], \text{ если } np \text{ - целое,} \end{cases} \quad (I)$$

где $Y_{(i)}$ - i - я порядковая статистика данной выборки. Если [I] $F(y)$ - непрерывная вместе с первыми двумя производными функция, то $\hat{Y}_p$ имеет асимптотически нормальное распределение со средним равным Y_p и средним квадратическим отклонением

$$\sigma = \frac{1}{f(Y_p)} \sqrt{\frac{p(1-p)}{n}}, \quad (2)$$

где $f(\cdot)$ - плотность распределения Y.

Рассмотрим наиболее типичный случай, когда регрессия представляет собой преобразование сдвига. Запишем модель для квантили Y_p в виде

$$\hat{Y}_p = \eta(A, X) + \varepsilon_0, \quad (3)$$

где A - параметры модели; X - неслучайные регрессоры; $\varepsilon_0 \sim N(0, \sigma^2)$. Если в качестве $\hat{Y}_p$ используется оценка (I), которая для всех точек плана строится по выборкам одинакового объема n, то дисперсии погрешности ε одинаковы во всех точках плана, так как $f(Y_p)$ не изменяется при преобразовании сдвига, и построение оценок параметров модели (3) не требует введения весов. Если же квантили в точках плана эксперимента X_i оцениваются по выборкам разного объема n_i, выбор весов w_i очевиден: $w_i \sim n_i$.

Кроме оценки (I) в настоящее время известны другие непараметрические оценки квантилей в классе непрерывных ф.р., обладающие свойством асимптотической нормальности и представляющие собой линейные комбинации порядковых статистик - это оценка Кая и Лахенбруха [2] и оценка Харрелла и Дэвиса [3].

Пусть из выборки $Y_1, Y_2, \ldots, Y_n$ извлекается без возвращения случайная подвыборка объема $k < n$, которая упорядочивается в вариационный ряд:

$$Y_{1:k;n}, Y_{2:k;n}, \ldots, Y_{k:k;n}. \quad (4)$$

Оценкой квантили уровня p, построенной по подвыборке (4), будет $Y_{[(k+1)p]:k;n}$. Оценка Кал и Лахенбруха (К-Л-оценка) представляет собой усредненную по всем $\binom{n}{k}$ подвыборкам (4) оценку

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$K_{[(k+1)p]} = \sum_{j=r}^{r+n-k} \left[\binom{j-1}{r-1} \binom{n-j}{k-r} \binom{n}{k}^{-1} \right] Y_{j:n}, \qquad (5)$$

где $Y_{j:n}$ – j-я порядковая статистика исходной выборки объема n, $r = [(k+1)p]$, при этом выбор k никак не оговаривается.

Оценка Харрелла и Дэвиса (Х–Д–оценка) строится на основе следующих соображений. Поскольку математическое ожидание порядковой статистики с номером $[(n+1)p]$ – $E(Y_{[(n+1)p]})$ сходится к $Y_p = F^{-1}(p)$, то в качестве оценки Y_p предлагается использовать некоторую оценку $E(Y_{[(n+1)p]})$ независимо от того, целое $(n+1)p$ или нет. Так как

$$E(Y_{(k)}) = \frac{1}{B(k, n-k+1)} \int_{-\infty}^{\infty} Y F(Y)^{k-1} [1-F(Y)]^{n-k} dF(Y),$$

где $B(\ell, m)$ – бета–функция, оценку $E(Y_{[(n+1)p]})$ предлагается строить путем замены $F(Y)$ на соответствующую выборочную ф.р. $F_n(Y)$. В результате Х–Д–оценка квантили уровня p имеет вид

$$Q_p = \frac{1}{B\{(n+1)p, (n+1)(1-p)\}} \int_0^1 F_n(Y) Y^{(n+1)p-1} (1-Y)^{(n+1)(1-p)-1} dy, \qquad (6.1)$$

или в терминах неполной бета–функции

$$Q_p = \sum_{i=1}^{n} W_{n,i} Y_{(i)}, \qquad (6.2)$$

где $W_{n,i} = I_{i/n}\{p(n+1), (1-p)(n+1)\} - I_{\frac{i-1}{n}}\{p(n+1), (1-p)(n+1)\}$, $I_x(a,b)$ – неполная бета–функция.

Использование непараметрических оценок (1), (5), (6) в задачах нормальной регрессии квантилей в первую очередь ставит вопрос о скорости сходимости этих оценок к асимптотическому нормальному распределению. В работе [3] приводятся качественные результаты исследований сходимости оценок (6) методом Монте–Карло с использованием статистики Колмогорова–Смирнова для равномерного и нормального распределений Y. Авторы указанной работы заключают, что нормальная аппроксимация распределений выборочных квантилей приемлема для выборок объемов порядка $n = 20$ при $p = 0,5$ и $n = 30 + 50$ для $p = 0,95$. Высказывается также предположение, что для асимметричных распределений, таких как экспоненциальное, для соблюдения нормальности может понадобиться $n = 80 + 100$ для квантилей уровня $p = 0,9$ и выше. Проведенные нами аналогичные исследования сходимости распределения оценки (1) для экспоненциального и вейбулловского распределений при $p = 0,1; 0,2; 0,5; 0,9$ и $n = 19 + 600$ в целом под-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

тверждают описанную выше картину - сходимость распределения квантилей на хвостах оцениваемых распределений хуже, чем в районе медианы. При этом для приложений можно рекомендовать следующую оценку: если сходимость распределения выборочной квантили на левом хвосте является приемлемой при p_1 и n_1, то она оказывается настолько же приемлемой при таких p_2 и n_2, что $p_1 n_1 = p_2 n_2$ при $p_{1,2} < 0,5$. Для правого хвоста распределения $p_{1,2} > 0,5$ аналогичная оценка дает $(1-p_1)n_1 = (1-p_2)n_2$. Сходимость К-Л-оценки (5) к нормальному распределению не исследовалась, однако в работе [3] высказывается предположение, что сходимость оценок квантилей хвостов распределений может потребовать больших объемов выборки. В этой же работе методом Монте-Карло исследовалась эффективность оценки (6) относительно (I) с помощью обобщенного ламбда-распределения [4]. В качестве оценки эффективности использовалось среднее квадратическое отклонение оценки квантили относительно истинного значения. Исследовались распределения с легкими, средними и очень тяжелыми (типа t-распределения с 5 степенями свободы) хвостами, а также подобные нормальному, Коши, экспоненциальному распределени . На основе результатов моделирования авторы работы [3] делают вывод о более высокой эффективности оценок (6) во всех случаях, за исключением распределений подобных Коши, хотя нам представляется, что на основе тех же данных для $p = 0,05$; 0,I и $n \leq 30$ можно сделать и противоположный вывод. Аналогичное сравнение эффективности оценок (I) и (5), выполненное в работе [2], показало более высокую эффективность оценки (5), за исключением распределений с тяжелыми хвостами - экспоненциального, двойного экспоненциального, Коши и случаев оценивания квантилей хвостов распределений.

Таким образом, выбор оценок в задачах регрессии квантилей как преобразования сдвига должен основываться на предварительном исследовании распределения. Вместе с тем возможны и общие рекомендации, например, регрессия медианы требует меньших объемов выборки, чем регрессия квантилей хвостов распределения. Кроме того, наличие цензурирования типа I или II однозначно определяет выбор оценки (I).

Рассмотрим далее случай, когда регрессия представляет собой линейное преобразование масштаба. Такого рода регрессии

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

рассматриваются как модели долговечности технических и биологических объектов в теории надежности [5] и биометрике [6].

Будем исходить из оценки квантили (I), поскольку в данных приложениях, как правило, имеют дело с цензурированными справа данными. Используя асимптотическую нормальность оценки (I), запишем

$$\hat{Y}_p \sim N\left(\zeta(A,X), \frac{b^2}{f^2(\zeta(A,X))}\right),$$

где $b = \sqrt{\frac{p(1-p)}{n}}$, $\zeta(A,X) = Y_p$, откуда

$$\hat{Y}_p = \zeta(A,X)\,\varepsilon, \qquad (7)$$

где $\varepsilon \sim N\left(1, \frac{b^2}{\zeta^2(A,X)\, f^2(\zeta(A,X))}\right)$.

Попытаемся привести модель (7) к нормальной регрессии, представляющей преобразование сдвига. Выделим далее такой достаточно часто рассматриваемый класс функций $\zeta(A,X)$, которые логарифмическим преобразованием приводятся к линейному по параметрам виду, т.е. $\ln(\zeta(A,X)) = XA$. Параметрические аналоги задачи построения таких регрессионных моделей рассмотрены в [7] для вейбулловского распределения, в [8] - для гамма-распределения, непараметрическим аналогом могут служить модели Кокса [6].

Представим теперь модель (7) в виде

$$\ln(\hat{Y}_p) = XA + \ln\varepsilon, \quad \varepsilon \sim N\left(1, \frac{b^2}{Y_p^2 f^2(Y_p)}\right), \qquad (8)$$

причем, поскольку мы рассматриваем линейное преобразование масштаба, $Y_p^2 f^2(Y_p) = const$. Погрешность ε представим как

$$\varepsilon = \varepsilon' + 1, \quad \varepsilon' \sim N\left(0, \frac{b^2}{Y_p^2 f^2(Y_p)}\right). \qquad (9)$$

Сделаем еще одно приемлемое для рассматриваемых приложений допущение - будем считать Y положительно определенной случайной величиной с неубывающей на интервале $[0, Y_p]$ плотностью. Нетрудно убедиться, что последнее допущение отвечает условию старения ф.р. [9] на этом интервале. При сделанных допущениях для малых $p < 0.5$ справедливо неравенство

$$f(Y_p)\,Y_p \geq p. \qquad (10)$$

Используя (10), заменим погрешность ε' в (9) погрешностью ε'' с большей дисперсией

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\varepsilon'' \sim N\left(0, \frac{b^2}{p^2} \equiv \frac{1-p}{pn}\right).$$

Потребовав теперь выполнения условия $\varepsilon'' << 1$, эквивалентного

$$\left(\frac{1-p}{pn}\right)^{1/2} << 1, \qquad (11)$$

получаем вместо модели (7) нормальную регрессию с постоянной дисперсией погрешности для $\ln(\hat{Y}_p)$

$$\ln(\hat{Y}_p) = XA + \varepsilon''. \qquad (12)$$

Таким образом получено условие (II), позволяющее при достаточно общих непараметрических ограничениях на ф.р. привести задачу построения регрессии линейного преобразования масштаба к нормальной регрессионной схеме для преобразования одвига. Данное условие соответствует упомянутой выше лучшей сходимости выборочных квантилей к асимптотическому распределению при уровнях квантилей близких к медиане. В заключении заметим еще раз, что условие (II) относится только к левому хвосту распределения.

ЛИТЕРАТУРА

1. Крамер Г. Математические методы статистики.-М.: Мир, 1975.-648с.
2. Kaigh W.D., Lachenbruch. A generalized quantile estimator.-Commun. Statist. Theory and methods, 1982, vol.11, N 19, p.p. 2217-2238.
3. Harrell F.E., Davis C.E. A new distribution-free quantile estimator.- Biometrika, 1982, vol.69, p.p. 635-640.
4. Ramberg J. Setal. A probability distribution and its uses in fitting data.- Technometrics, 1974, vol.21, p.p.201-214.
5. Mann N.R., Schafer R.E., Singpurwalla N.D. Methods for statistical analysis of reliability and life data.- N-Y, Wiley, 1974, 560p.
6. Cox D.R. Regression models and life tables.- Journ. Roy. Stat. Soc., 1972, B34, p.p. 187-220.
7. McCool J.I. Confidence limits for Weibull regression with censored data.- IEEE Transaction of reliability, 1980, vol.R-29, N 2, p.p. 145-150.
8. Jones R.A.etal. Tolerance bounds for log gamma regression models. - Technometrics, 1985, vol.27, N 2, p.p.109-118.
9. Барлоу Р., Прошан Ф. Статистическая теория надежности и испытания на безотказность/ Пер. с анг.-М.: Наука, 1984.-327с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ОБРАБОТКА ИЗОБРАЖЕНИЙ НА ОСНОВЕ МОДЕЛИ МНОЖЕСТВЕННОЙ РЕГРЕССИИ

С.А.Кислицын, М.П.Ангелов, В.М.Мохов

Среди специальных методов статистической обработки изображений: сжатия и аппроксимации, фильтрации и восстановления, наиболее употребительными на практике являются методы, основанные на непараметрическом подходе [I-5]. В данной работе исследуются статистические свойства оценок метода сжатия данных с фрагментов изображений, опубликованного в [5]. Метод основан на приближении функций двух независимых переменных, заданных таблично, ортогональными многочленами Чебышёва одного дискретного переменного в разделимом базисе. Метод выгодно отличается в вычислительном отношении от ряда известных процедур сжатия или восстановления функций многих переменных, представленных таблицами данных, прежде всего разделимым базисом [2-4], рекомендованным ещё в [I], и рекуррентным характером оценок, не требующим пересчёта данных в случае неудовлетворительной точности приближения [5]. Со статистической точки зрения, обработка изображения, заключающаяся в его сжатии, представляет собой процесс построения модели множественной регрессии двух переменных. Исследование такой модели, помимо обоснования метода сжатия изображений, представляет самостоятельный интерес.

I. Постановка задачи

Пусть на $S \subset R^2$ задана функция f, описываемая уравнением

$$f(x,y) = h(x,y) + \xi(x,y), \qquad (I)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$h(x,y)$ - из класса обобщённых многочленов, $f(x,y)$ известна своими значениями на множестве точек (x_i, y_j) из S $(i \in \bar{I},\ j \in \bar{J})$, где $\bar{I} = \{0,1,\dots,I-1\}$, $\bar{J} = \{0,1,\dots,J-1\}$, $f(x,y) \in \mathcal{L}_2(\Omega, \mathcal{M}, P_f)$ - вещественные случайные величины, зависящие от параметра $(x,y) \in S$, $(\Omega, \mathcal{M}, P_f)$ - выборочное вероятностное пространство [6], связанное с фрагментом изображения, представленного функцией f. Обозначим через $\mathcal{L}_2' = \mathcal{L}_2(f_{ij} : i \in \bar{I},\ j \in \bar{J})$ гильбертово пространство случайных величин, порождённое сеткой $(x_i, y_j) \in S$, а через $\mathcal{M}' = \sigma\{f_{ij} : i \in \bar{I}, j \in \bar{J}\} \subset \mathcal{M}$ - соответствующую σ-алгебру.

Сжатие данных, связанное с наилучшим приближением h^* функции $f \in \mathcal{L}_2$ элементами пространства $\mathcal{L}_2' \subseteq \mathcal{L}_2$, осуществим путём нахождения обобщённого многочлена h, составленного из линейно независимых и ортогональных функций $\{\vartheta\}$ так, чтобы в $\mathcal{L}_2'$ величина

$$\|f - h\|^2 = M \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} [f(x_i, y_j) - h(x_i, y_j)]^2 \qquad (2)$$

достигала минимума. Здесь M - символ математического ожидания по мере P_f,

$$h(x,y) = \sum_{k=0}^{K-1} \sum_{\ell=0}^{L-1} \theta_{k\ell}\, \vartheta_{k\ell}(x,y) \quad (K\ell \le IJ). \qquad (3)$$

Задача состоит в том, чтобы помимо вычисления коэффициентов $\theta_{k\ell}$ исследовать свойства полученных оценок.

2. Вид оценки множественной регрессии

Для построения базиса $\{\vartheta\}$ оценки достаточно зафиксировать известную линейно независимую систему ортогональных функций одной переменной и использовать разделимость функций [1,2,5].

ЛЕММА 1. Пусть $\{\varphi_k\}$, $\{\psi_\ell\}$ - полные линейно независимые системы ортогональных функций на X и Y $(X, Y \subset R^2)$. Тогда

$$\vartheta_{k\ell}(x,y) = \varphi_k(x)\psi_\ell(y) \qquad (4)$$

образует полную, ортогональную в

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

метрике $\mathcal{L}_2'$, систему линейно независимых векторов на $S \subseteq X \times Y \subseteq R^2$.

Доказательство. Пусть для любых двух функций f и g, заданных на R^2, известных своими значениями в конечном числе точек, $(f|g)$ - естественное скалярное произведение. Тогда на $S \subset R^2$ с учетом (4)

$$(v_{ke}|v_{mu})_S = \sum_{i,j} v_{ke}(x_i,y_j)\, v_{mu}(x_i,y_j) =$$

$$\sum_{i,j} \varphi_k(x_i)\psi_e(y_j)\varphi_m(x_i)\psi_u(y_j) = (\varphi_k|\varphi_m)_X (\psi_e|\psi_u)_Y .$$

Отсюда следует ортогональность базиса (4). В силу полноты и линейной независимости систем $\{\varphi_k\}$ и $\{\psi_e\}$ $(k,e=0,1,\dots)$ система $\{v_{ke}\}$ будет также линейно независимой полной системой [7]. Лемма I доказана.

В теории приближений решена задача выбора оптимального базиса для конкретной таблицы в виде обобщённых многочленов, в качестве которых могут выступать многочлены Чебышёва дискретного переменного $T_{k,L}(x)$ $(k \le L, x \in R)$ [7], где L - количество измерений, k - степень многочлена. Положим [5]:

$$\begin{cases} \varphi_k(x) = T_{k,I}(x); & (k \le I-1,\ x \in X); \\ \psi_e(y) = T_{e,J}(y); & (e \le J-1,\ y \in Y). \end{cases} \qquad (5)$$

ТЕОРЕМА I. Пусть базис $\{v_{ke}(x,y)\}$ $((x,y) \in S)$ определён соотношениями (4), (5). Наилучшее в метрике (2) с.к. - приближение (3) функции f, заданной таблицей f_{ij} $(i \in \bar{I},\ j \in \bar{J})$, существует, единственно $(\mathrm{mod}\, P_f)$ и есть произведение многочленов, степеней в точности равных $K \le I-1$, $L \le J-1$. При этом $(k \in \bar{K}, e \in \bar{L})$

$$\theta^*_{ke} = M(\theta_{ke}|M') = \alpha_k^{-1}\beta_e^{-1} \sum_{i=0}^{I-1}\sum_{j=0}^{J-1} f_{ij}\varphi_{ki}\psi_{ej}, \qquad (6)$$

где $\varphi_{ki} = T_{k,I}(x_i)$, $\psi_{ej} = T_{e,J}(y_j)$, $\alpha_k = \sum_{i=0}^{I-1} \varphi_{ki}^2$, $\beta_e = \sum_{j=0}^{J-1} \psi_{ej}^2$.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Доказательство. Подставляя (3) в (2), дифференцируя (2) по θ_{mn} $(m \in \bar{K}, n \in \bar{L})$ и приравнивая результат к нулю, находим

$$\sum_{i,j} f_{ij} \varphi_{mi} \psi_{nj} - \sum_{k,e} \theta_{ke} \sum_{i,j} \varphi_{ki} \psi_{ej} \varphi_{mi} \psi_{nj} = 0, \qquad (7)$$

где обозначено

$$f_{ij} = f(x_i, y_j) \qquad (i \in \bar{I}, j \in \bar{J});$$

$$\varphi_{mi} = \varphi_m(x_i) \qquad (m \in \bar{K} = \{0, 1, \dots, K-1\});$$

$$\psi_{nj} = \psi_n(y_j) \qquad (n \in \bar{L} = \{0, 1, \dots, L-1\})$$

для всяких i, j таких, что $(x_i, y_j) \in S$.

Учитывая формулу (5), ортогональность базиса $\{\theta\}$ (лемма I), измеримость f_{ij} для всех $i \in \bar{I}$, $j \in \bar{J}$ относительно σ-алгебры $\mathcal{M}'$, порождённой экспериментом, из уравнения (7) сразу получаем формулу (6). Доказательство теоремы I закончено.

3. Свойства оценок множественной регрессии

Исследуем свойства статистики (6). Предположение А: пусть $\forall i, j$ $(i \in \bar{I}, j \in \bar{J})$, $M\xi_{ij} = 0$, $M\xi_{ij}\xi_{pq} = \sigma^2 \delta_{ip} \delta_{jq}$, $M\xi_{ij}^2 = \sigma^2$, где δ_{ij} - символ Кронекера. Рассмотрим модель изображения (I), (3).

ТЕОРЕМА 2. Пусть выполнены условия теоремы I и предположечие А. Тогда

I) оценка θ^*_{ke} является несмещённой

$$M\theta^*_{ke} = \theta_{ke} \quad (k \in \bar{K}, l \in \bar{L});$$

2) состоятельной (в среднеквадратичном смысле)

$$M(\theta^*_{mn} - \theta_{mn})(\theta^*_{pq} - \theta_{pq}) = \sigma^2 \alpha_m^{-1} \beta_n^{-1} \delta_{mp} \delta_{nq}, \qquad (8)$$

где при $I, J \to \infty$ по определению $\alpha_m, \beta_n \to \infty$;

3) несмещённой оценкой для σ^2 будет величина

$$\sigma^2_* = (IJ - KL)^{-1} \sum_{i=0}^{\bar{I}-1} \sum_{j=0}^{\bar{J}-1} \zeta_{ij}^2, \qquad (9)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

где $\zeta_{ij} = f_{ij} - f^*_{ij}$,

$$f^*_{ij} = M(f_{ij} | \mathcal{M}') = \sum_{k=0}^{K-1} \sum_{\ell=0}^{L-1} \theta^*_{k\ell} \varphi_{ki} \psi_{\ell j} \; ; \qquad (10)$$

4) остаточная дисперсия есть

$$s_0^2 = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} f_{ij}^2 - \sum_{k=0}^{K-1} \sum_{\ell=0}^{L-1} (\theta^*_{k\ell})^2 \lambda_k \beta_\ell \; ; \qquad (11)$$

5) дисперсия коэффициентов θ^*_{mn} $(m \in \bar{K}, n \in \bar{L})$ является оценкой несмещённой и равна

$$\sigma_\theta^2(m,n) = M\{(\theta^*_{mn} - \theta_{mn})^2 | \mathcal{M}'\} = (IJ - KL)^{-1} \lambda_m^{-1} \beta_n^{-1} s_0^2 . \qquad (12)$$

Доказательство. Убедимся в несмещённости оценки θ^*. Для этого запишем оценку (6) через (I) для произвольно фиксированных $k \in \bar{K}$, $\ell \in \bar{L}$:

$$\lambda_k \beta_\ell \theta^*_{k\ell} = \sum_{i,j} f_{ij} \varphi_{ki} \psi_{\ell j} = \sum_{i,j} \sum_{m,n} \theta_{mn} \varphi_{mi} \psi_{nj} \varphi_{ki} \psi_{\ell j} +$$

$$+ \sum_{i,j} \varphi_{ki} \psi_{\ell j} \xi_{ij} = \lambda_k \beta_\ell \theta_{k\ell} + \sum_{i,j} \xi_{ij} \varphi_{ki} \psi_{\ell j} . \qquad (13)$$

Отсюда $M\theta^*_{k\ell} = \theta_{k\ell}$.

Докажем второе утверждение. С учётом доказанного пункта и (13)

$$M(\theta^*_{mn} - M\theta^*_{mn})(\theta^*_{pq} - M\theta^*_{pq}) = M(\theta^*_{mn} - \theta_{mn})(\theta^*_{pq} - \theta_{pq}) =$$

$$= M \sum_{i,j} \sum_{r,s} \xi_{ij} \xi_{rs} \varphi_{mi} \psi_{nj} \varphi_{pr} \psi_{qs} \lambda_m^{-1} \lambda_p^{-1} \beta_n^{-1} \beta_q^{-1} =$$

$$= \sigma^2 \lambda_m^{-1} \beta_n^{-1} \delta_{mp} \delta_{nq} .$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Если $I \to \infty$ либо $J \to \infty$, то по построению базиса (4), (5)

$$\alpha_m = \sum_{i=0}^{I-1} \varphi_{mi}^2 = \frac{(m+1+I)^{[m+1]}}{(2m+1)\, I^{[m]}} \to \infty ,$$

либо

$$\beta_n = \sum_{j=0}^{J-1} \gamma_{nj}^2 = \frac{(n+1+J)^{[n+1]}}{(2n+1)\, J^{[n]}} \to \infty ,$$

где $x^{[n]} = x(x-1)\dots(x-n+1)$ - факториальный многочлен порядка n [7]. Поэтому для всяких $m \in \overline{K}$, $n \in \overline{L}$

$$M(\theta^*_{mn} - \theta_{mn})^2 \to 0 \quad \text{при } I \to \infty \text{, либо } J \to \infty .$$

Оценим дисперсию $\sigma^2 = M\xi^2_{ij}$. Для этого рассмотрим величину

$$\zeta_{ij} = f_{ij} - h^*_{ij} \qquad (i \in \overline{I},\ j \in \overline{J}),$$

где h^*_{ij} определено соотношением (10). Тогда, с учётом (1)

$$\zeta_{ij} = \sum_{k=0}^{K-1} \sum_{e=0}^{L-1} (\theta_{ke} - \theta^*_{ke}) \varphi_{ki} \gamma_{ej} + \xi_{ij} .$$

Очевидно, $M\zeta_{ij} = 0$. Подсчитаем величину $\sum \zeta^2_{ij}$. Имеем

$$\zeta^2_{ij} = \Big[\sum_{k,e} (\theta_{ke} - \theta^*_{ke}) \varphi_{ki} \gamma_{ej} + \xi_{ij}\Big]\Big[\sum_{p,q} (\theta_{pq} - \theta^*_{pq}) \varphi_{pi} \gamma_{qj} + \xi_{ij}\Big] =$$

$$= \sum_{k,e} \sum_{p,q} (\theta_{ke} - \theta^*_{ke})(\theta_{pq} - \theta^*_{pq}) \varphi_{ki} \varphi_{pi} \gamma_{ej} \gamma_{qj} +$$

$$+ 2 \sum_{k,e} (\theta_{ke} - \theta^*_{ke}) \varphi_{ki} \gamma_{ej} \xi_{ij} + \xi^2_{ij} .$$

В силу (6), (13) и ортогональности базиса

$$\sum_{i,j} \zeta^2_{ij} = \sum_{k,e} \sum_{p,q} (\theta_{ke} - \theta^*_{ke})(\theta_{pq} - \theta^*_{pq}) \sum_m \varphi_{km} \varphi_{pm} \sum_n \gamma_{en} \gamma_{qn} +$$

$$+ 2 \sum_{k,e} (\theta_{ke} - \theta^*_{ke}) \sum_{m,n} \xi_{mn} \varphi_{km} \gamma_{en} + \sum_{m,n} \xi^2_{mn} =$$

$$= \sum_{k,e} (\theta_{ke} - \theta^*_{ke})^2 \alpha_k \beta_e + 2 \sum_{k,e} (\theta_{ke} - \theta^*_{ke})(\theta^*_{ke} - \theta_{ke}) \alpha_k \beta_e +$$

$$+ \sum_{m,n} \xi^2_{mn} = \sum_{i,j} \xi^2_{ij} - \sum_{k,e} (\theta_{ke} - \theta^*_{ke})^2 \alpha_k \beta_e .$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Следовательно, с учётом доказанного соотношения (8)

$$M\sum_{i=0}^{I-1}\sum_{j=0}^{J-1}\zeta_{ij}^2 = \sum_{i=0}^{I-1}\sum_{j=0}^{J-1} M\{\xi_{ij}^2 - \sum_{k=0}^{K-1}\sum_{e=0}^{L-1} M(\Theta_{ke}-\Theta_{ke}^*)^2 \lambda_k \beta_e =$$

$$= \sigma^2 IJ - \sigma^2 KL = \sigma^2(IJ-KL).$$

Отсюда вытекает, что несмещённая оценка для σ^2 может быть подсчитана по формуле (9). Действительно,

$$M\sigma_*^2 = (IJ-KL)^{-1}\sum_{i,j}\zeta_{ij}^2 = (IJ-KL)^{-1}(IJ-KL)\sigma^2 = \sigma^2.$$

Далее, с другой стороны, из соотношений (10) можем записать

$$\zeta_{ij}^2 = (f_{ij}-f_{ij}^*)^2 = [f_{ij} - \sum_{k,e}\Theta_{ke}^*\varphi_{ki}\psi_{ej}][f_{ij} - \sum_{p,q}\Theta_{pq}^*\varphi_{pi}\psi_{qj}] =$$

$$= f_{ij}^2 - 2\sum_{k,e}\Theta_{ke}^*\varphi_{ki}\psi_{ej}f_{ij} + \sum_{k,e}\sum_{p,q}\Theta_{ke}^*\Theta_{pq}^*\varphi_{ki}\psi_{ej}\varphi_{pi}\psi_{qj}.$$

Откуда, поскольку f_{ij} и Θ_{ke}^* - M'- измеримые случайные величины при каждом i, j, k, e, получаем формулу (11), которая в пределе является аналогом формулы Парсеваля. Именно

$$\ell_0^2 = \sum_{i=0}^{I-1}\sum_{j=0}^{J-1}\zeta_{ij}^2 = \sum_{i=0}^{I-1}\sum_{j=0}^{J-1} f_{ij}^2 - \ell\sum_{k=0}^{K-1}\sum_{e=0}^{L-1}(\Theta_{ke}^*)^2\lambda_k\beta_e +$$

$$+ \sum_{k=0}^{K-1}\sum_{e=0}^{L-1}\sum_{p=0}^{K-1}\sum_{q=0}^{L-1}\Theta_{ke}^*\Theta_{pq}^*\lambda_k\beta_e\delta_{kp}\delta_{eq} = \sum_{i=0}^{I-1}\sum_{j=0}^{J-1} f_{ij}^2 - \sum_{k=0}^{K-1}\sum_{e=0}^{L-1}(\Theta_{ke}^*)^2\lambda_k\beta_e.$$

Докажем последнее утверждение теоремы. Используем соотношения (8), (9), (11). Можем записать для любых $m \in \overline{K}$, $n \in \overline{I}$

$$\sigma_\Theta^2(m,n) = M\{(\Theta_{mn}^* - \Theta_{mn})^2 | M'\} =$$

$$= \lambda_m^{-1}\beta_n^{-1}\sigma_*^2 = (IJ-KL)^{-1}\lambda_m^{-1}\beta_n^{-1}\ell_0^2.$$

Несмещенность условной дисперсии σ_Θ^2 следует из несмещенности оценки σ_*^2. На самом деле

$$M\sigma_\Theta^2(m,n) = \lambda_m^{-1}\beta_n^{-1}M\sigma_*^2 = \lambda_m^{-1}\beta_n^{-1}\sigma^2,$$

что совпадает с (8). Теорема 2 доказана.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ЗАМЕЧАНИЕ I. Из полученных результатов интерес представляет формула (9), вид которой близок к классической несмещённой оценке дисперсии восстановления функции по типу одномерной полиномиальной регрессии. В этом смысле формулу (9) следует рассматривать как обобщение [6].

ЗАМЕЧАНИЕ 2. Наибольший практический интерес при обработке изображений путём их сжатия на основе модели множественной регрессии представляет формула (I2), из которой следует, что монотонно $\sigma^2_{\Theta}(m,n) \downarrow 0$ при $m \uparrow I-1$, либо $n \uparrow I-1$. Этот эффект обусловлен типом оценки и выбором базиса.

ЛИТЕРАТУРА

I. Ярославский Л.П. Введение в обработку изображений. -М.: Сов. радио, I979. - 3I2 с.
2. Джайн А.К. Сжатие видеоданных: Обзор. - ТИИЭР, I98I, т. 69, №3, с.7I-II7.
3. Галковский Т., Рутковский Л. Непараметрическое восстановление функций многих переменных с приложением к идентификации систем. - ТИИЭР, I985, т.73, №6, с.262-263.
4. Пытьев Ю.П. Задачи реставрации изображений. -ДАН СССР, I979, т.245, №I, с.42-46.
5. Кислицын С.А., Ангелов М.П., Мохов В.М. Алгоритм сжатия картографической информации. -В кн.: Методы и микроэлектронные средства цифрового преобразования и обработки сигналов, т.2. Рига, I986, с. 6I5-6I6.
6. Ибрагимов И.А., Хасьминский Р.З. Асимптотическая теория оценивания. - М.: Наука, I979. - 528 с.
7. Березин И.С., Жидков Н.П. Методы вычислений, т.I. - М.: Наука, I966. - 632 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ОЦЕНКИ МЕДИАННОГО ТИПА ДЛЯ КОЭФФИЦИЕНТОВ ЛИНЕЙНОЙ РЕГРЕССИОННОЙ МОДЕЛИ

А.В. Китаева

Рассмотрим модель линейной (относительно оцениваемых параметров) регрессии

$$y_i = \sum_{k=1}^{S} \theta_k x_{ik} + n_i \;, \quad i = \overline{1,N} \;, \qquad \text{где } X = (x_{ik},\ i = \overline{1,N},\ k = \overline{1,S})$$

– заданная матрица, $\theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_S \end{pmatrix}$ – вектор неизвестных параметров, n_i – независимые одинаково распределенные случайные величины с симметричной дифференцируемой плотностью распределения p. Будем предполагать, что ранг матрицы X равен S, и ее элементы ограничены при $N \to \infty$.

I. Будем искать оценки параметров θ_k – $\hat{\theta}_k$ из следующей системы уравнений:

$$\xi_m^I = \frac{1}{N} \sum_{i=1}^{N} x_{ik}\, sign\Big(y_i - \sum_{k=1}^{S} \hat{\theta}_k^I x_{ik}\Big) = 0 \;, \quad m = \overline{1,S}. \qquad (I)$$

Рассмотрим асимптотические характеристики получающихся оценок.

ТЕОРЕМА I. Если $p(0) \neq 0$, то оценки $\hat{\theta}_k^I$ $k = \overline{1,S}$ являются сильно состоятельными.

ДОКАЗАТЕЛЬСТВО. I) При $N \to \infty$ $\xi_m^I \xrightarrow{п.н.} M\xi_m^I$, $m = \overline{1,S}$, что непосредственно следует из усиленного закона больших чисел, т.к.

$$\mathcal{D}\big[sign(y_i - \sum_{k=1}^{S} \theta_k^I x_{ik})\big] < \infty.$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

2) Система уравнений

$$M\xi_m^I = 0, \quad m = \overline{1,s}, \qquad (2)$$

имеет единственное решение, если $p(0) \neq 0$. Действительно,

$$M\xi_m^I = \frac{1}{N}\sum_{i=1}^{N} x_{im} F\Big(\sum_{k=1}^{s} \Delta\theta_k^I x_{ik}\Big), \quad \text{где } \Delta\theta_k^I = \theta_k - \widehat{\theta}_k^I,$$

$F(x) = 2\int_0^x p(t)\,dt$. Умножим m-ое уравнение системы (2) на $\Delta\theta_m^I$ и, суммируя по m от I до s, получим

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{m=1}^{s} \Delta\theta_m^I x_{im} F\Big(\sum_{k=1}^{s}\Delta\theta_k^I x_{ik}\Big) = 0.$$

Отсюда, так как $F(x)$ является нечетной функцией и обращается в 0 лишь при $x = 0$, следует, что $\sum_{k=1}^{s}\Delta\theta_k^I x_{ik} = 0 \ \forall i$; т.к. $\operatorname{rang} X = s$, получаем, что $\Delta\theta_k^I = 0 \ \forall k$.

Из I) и 2) следует $\widehat{\theta}_k^I \xrightarrow[N\to\infty]{\text{п.н.}} \theta_k \ \forall k$.

ТЕОРЕМА 2. Асимптотическое распределение величин $\sqrt{N}\Delta\theta_k^I$, $k = \overline{1,s}$ является нормальным.

ДОКАЗАТЕЛЬСТВО. Представим систему (I) в виде

$$M\xi_m^I + \big(\xi_m^I - M\xi_m^I\big) = 0, \ m = \overline{1,s}$$

и разложим $M\xi_m^I$ в ряд по $\Delta\theta = \begin{pmatrix}\Delta\theta_1\\ \vdots \\ \Delta\theta_s\end{pmatrix}$ в т. $\Delta\theta = 0$.

$$M\xi_m^I = M\xi_m^I\big|_{\Delta\theta=0} + F'(0)\sum_{\tau=1}^{s}\frac{1}{N}\sum_{i=1}^{N} x_{im}x_{i\tau}\Delta\theta_\tau + \frac{1}{2}F''\big|_{\Delta\theta=\Delta\tilde{\theta}};$$

$$\sum_{\tau,e=1}^{s}\frac{1}{N}\sum_{i=1}^{N} x_{im}x_{i\tau}x_{ie}\Delta\theta_\tau\Delta\theta_e, \quad \Delta\tilde{\theta} = \begin{pmatrix}\Delta\tilde{\theta}_1\\ \vdots\\ \Delta\tilde{\theta}_s\end{pmatrix}, \ \Delta\tilde{\theta}_k \in [0, \Delta\theta_k];$$

$$M\xi_m^I\big|_{\Delta\theta=0} = 0 \ \forall m, \quad F'(x)\big|_{x=0} = 2p(0), \quad F''(x) = 2p'(x).$$

По теореме о сходимости [I] с учетом результата теоремы I получаем, что распределение $\Delta\theta_k$, $k = \overline{1,s}$ совпадает с распределением $-\lim_{N\to\infty} N(X^TX)_k^{-1}\Delta\xi^I/2p(0)$, где $\Delta\xi^I = \begin{pmatrix}\Delta\xi_1^I\\ \vdots\\ \Delta\xi_s^I\end{pmatrix}$.

$\Delta\xi_i^I = \xi_i^I\big|_{\theta^I=\theta} - M\xi_i^I\big|_{\theta^I=\theta}$, $(X^TX)_k^{-1}$ – k - ая строка матрицы $(X^TX)^{-1}$, откуда по центральной предельной теореме и следует асимптотическая нормальность $\sqrt{N}\theta_k$.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Найдем асимптотическую ковариационную матрицу вектора $\Delta\theta^I$:

$$M\{\Delta\theta^I\Delta\theta^{IT}\}=\lim_{N\to\infty}N(X^TX)^{-1}M\{\Delta\xi^I\Delta\xi^{IT}\}N(X^TX)^{-1}/4p^2(0).$$

Учитывая, что $M\{\Delta\xi^I\Delta\xi^{IT}\}\sim\frac{\lim N^{-1}X^TX}{N}$, получим $M\{\Delta\theta^I\Delta\theta^{IT}\}\sim\frac{(X^TX)^{-1}}{4p^2(0)}$.

Покажем, что наши оценки имеют минимальную возможную дисперсию для оценок, получаемых из системы вида

$$\xi_m=\frac{1}{N}\sum_{i=1}^{N}a_{im}\,sign\Big(y_i-\sum_{k=1}^{S}\hat\theta_k x_{ik}\Big)=0,\quad m=\overline{1,S}.$$

Для оценок $\hat\theta_k$ при выполнении соответствующих условий на матрицу $A=\{a_{im},\ i=\overline{1,N},\ m=\overline{1,S}\}$ ковариационная матрица будет иметь вид $\lim_{N\to\infty}N(A^TX)^{-1}M\{\Delta\xi\Delta\xi^T\}N(X^TA)^{-1}$, где $M\{\Delta\xi\Delta\xi^T\}\sim\frac{\lim N^{-1}A^TA}{N}$. Покажем, что матрица $C=(A^TX)^{-1}A^TA\cdot(X^TA)^{-1}-(X^TX)^{-1}$ положительно полуопределена. Для этого воспользуемся неравенством Коши-Буняковского:

$(a^Tb)^2\le(a^Ta)(b^Tb)$ $\forall$ векторов a и b. Если в качестве a возьмем $A(X^TA)^{-1}z$, в качестве b – $X(X^TX)^{-1}z$, то получим $z^TCz\ge0$.

П. В качестве оценок $\hat\theta_k^{II}$, $k=\overline{1,S}$ возьмем решение системы уравнений

$$\xi_m^{II}=\frac{1}{N^2}\sum_{i,j=1}^{N}[x_{im}-x_{jm}]\,sign\Big[y_i-y_j-\sum_{k=1}^{S}\hat\theta_k^{II}(x_{ik}-x_{jk})\Big]=0\quad m=\overline{1,S}. \tag{3}$$

Рассмотрим асимптотические характеристики $\hat\theta_k^{II}$, $k=\overline{1,S}$.

Пусть выполняется следующее предположение: ранг матрицы XR равен $S+1$, где $XR=\{xr_{ik},\ i=\overline{1,N},\ k=\overline{1,S+1}:\ xr_{ik}=x_{ik},\ i=\overline{1,N},\ k=\overline{1,S};\ xr_{iS+1}=1\}$. Тогда имеет место

ТЕОРЕМА 3. Оценки $\hat\theta_k^{II}$, $k=\overline{1,S}$ сильно состоятельны.

Доказательство по структуре аналогично доказательству теоремы I, а доказательство пунктов I) и 2) по существу аналогично доказательству соответственно лемм I и 2 из [2]. Вышеприведенное условие обеспечивает в пункте 2 единственность тривиального решения для системы $\sum_{k=1}^{S}(x_{ik}-x_{jk})\Delta\hat\theta_k^{II}=0$, $i,j=\overline{1,N}$, что в [2] достигалось засчет ортогональности матрицы XR. Обозначим $B=\{b_{ij},\ i,j=\overline{1,S}\}=\lim_{N\to\infty}\frac{1}{N^2}(NX^TX-X^TEX)^{-1}$, $E=\{e_{ij},\ i,j=\overline{1,N}:\ e_{ij}=1\}$.

ТЕОРЕМА 4. В асимптотике при $N\to\infty$ ве-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

личины $\sqrt{N}\Delta\theta_k^{II}$, $k=\overline{1,s}$ подчиняются распределению Гаусса с параметрами $(0, (2\sqrt{6}\int p^2(x)dx/\sqrt{B_{kk}^{-1}})^{-2}.2)$

Для доказательства сначала нужно показать аналогично как и в теореме 2, что асимптотическое распределение величин $\Delta\theta_k^{II}$ совпадает с распределением величин $\frac{B_k \Delta\xi^I}{-4\int p^2(x)dx}$.

Все обозначения аналогичны разделу I. Далее теорема доказывается также как и теорема 2 из [2].

Учитывая, что $M[\Delta\xi^I \Delta\xi^{IT}] \sim 4B^{-1}/3N$, асимптотическая ковариационная матрица оценок θ_k^I будет иметь вид

$$B/12\left[\int p^2(x)dx\right]^2 N.$$

В заключение отметим устойчивость предлагаемых оценок $\hat{\theta}_k^I$, θ_k^I, $k=\overline{1,s}$ к аномальным ошибкам наблюдений в сравнении с МНК - оценками. Оценки $\hat{\theta}_k^{II}$ можно считать предпочтительнее оценок $\hat{\theta}_k^I$ в смысле ограниченности сверху сомножителя $\left[\int p^2(x)dx\right]^2/12$, стоящего перед дисперсией оценок. Заметим, что оценки первого $(\hat{\theta}_k^I)$ и второго $(\hat{\theta}_k^{II})$ типов можно применять совместно для различных k, исходя из особенностей матрицы X, что в частности и сделано в [2].

Литература

1. Г. Крамер. Математические методы статистики.- М.: Мир, 1975.- 648с.
2. Китаева А.В. Оценки медианного типа для коэффициентов тренда временного ряда.- В кн.: Поиск сигнала в многоканальной системе. Томск: Изд-во ТГУ, 1987, вып.2, с. 15.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ОБ ОСОБЕННОСТЯХ СХОДИМОСТИ РЕКУРРЕНТНЫХ ПРОЦЕДУР ТИПА МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ

М.М.Коган, Ю.И.Неймарк

Изучается характер сходимости рекуррентных процедур оценивания в задаче линейной регрессии при вырожденных и близких к вырожденным входах. Установлено, что в первом случае, отвечающем неидентифицируемости системы, имеет место сходимость оценок к некоторому многообразию, которому принадлежит неизвестный вектор, а во втором случае происходит разделение движений на быстрые, соответствующие приходу оценок к многообразию, и медленные, соответствующие движению вдоль многообразия к оцениваемому вектору.

Рекуррентная процедура некоторого общего вида для идентификации неизвестной матрицы Ω_* порядка $m \times K$ по измерениям векторов $X_t \in R^K$ и $\Phi_t \in R^m$, связанных зависимостью

$$X_{t+1} = \Omega_*^T \cdot \Phi_t + \xi_{t+1}, \quad t \geq 0, \qquad (1)$$

где $\xi_t^T = (\xi_t^{(1)}, \xi_t^{(2)}, \ldots, \xi_t^{(K)})$ образуют последовательность независимых случайных векторов, $E\xi_t = 0$, $E\xi_t\xi_t^T = K_\xi$, определяется уравнением [1]:

$$\Omega_{t+1} = \Omega_t + \varkappa \cdot \Gamma_{t+1}^{-1} \Phi_t \cdot \left(X_{t+1} - \Omega_t^T \cdot \Phi_t\right)^T, \qquad (2)$$

в котором Ω_0 произвольная матрица $m \times K$. Матрицы Γ_t порядка $m \times m$ и величина $\varkappa$ задаются одним из соотношений

$$\Gamma_{t+1} = G_{t+1}, \quad \varkappa = 1,$$

$$G_{t+1} = \begin{cases} G_t + \Phi_t \Phi_t^T, & 0 \leq t < \tau - 1 \\ G_0 + \sum_{s=0}^{\tau-1} \Phi_{t-s} \cdot \Phi_{t-s}^T, & t \geq \tau - 1, \end{cases} \qquad (3)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

в котором $G_0 > 0$, $\tau \geq 1$ - целое число, характеризующее "глубину памяти" алгоритма, или

$$\Gamma_{t+1} = \gamma_{t+1} I_m, \quad \gamma_{t+1} = Spur\, G_{t+1}, \quad 0 < \varkappa < 2, \qquad (4)$$

где I_m -единичная матрица $m \times m$. Алгоритм матричной идентификации (2) распадается на k алгоритмов, идентифицирующих столбцы матрицы Ω_* , и , когда Ω_* вектор, содержит в себе алгоритм Качмажа (в этом случае Γ_t и $\varkappa$ определяются в (4), $\tau=1$) и рекуррентную процедуру метода наименьших квадратов (тогда Γ_t и $\varkappa$ задаются в (3), $\tau = \infty$).

Ковариационная матрица вектора $\Omega_t^{(i)}$ -оценки i -го столбца матрицы Ω_* , полученной в соответствии с алгоритмом (2),(3), $\tau = \infty$ вычисляется по формуле

$$K_{\Omega_t^{(i)}} = \sigma_i^2 \cdot \left(\Gamma_t^{-1} - \Gamma_t^{-1} \Gamma_0 \Gamma_t^{-1} \right), \quad \sigma_i^2 = E\, \xi_t^{(i)2},$$

в которой $\Gamma_t^{-1} = \left(\Gamma_0 + \sum_{s=0}^{t} \Phi_s \Phi_s^T \right)^{-1}$, и для состоятельности этой оценки в средне квадратичном необходимо и достаточно выполнения условия [2]

$$\lim_{t \to \infty} \min_{a} \left[a^T \left(\sum_{s=0}^{t} \Phi_s \Phi_s^T \right) a / \|a\|^2 \right] = \infty. \qquad (5)$$

Допустим, что это условие не выполнено, но существует ортогональное преобразование P и момент времени $T > 0$, что для всех $t > T$

$$\sum_{s=0}^{T} \Phi_s \Phi_s^T = P \begin{pmatrix} R_{11}(t) & R_{12}(t) \\ R_{12}^T(t) & R_{22}(t) \end{pmatrix} P^T, \qquad (6)$$

где матрицы $R_{11}(t), R_{12}(t), R_{22}(t)$ имеют порядки $\ell \times \ell, \ell \times (m-\ell), (m-\ell) \times (m-\ell)$ соответственно и удовлетворяют условиям

$$\lim_{t \to \infty} \lambda_{min}(R_{11}(t)) = \infty \; ; \; \lim_{t \to \infty} R_{11}^{-1}(t) \cdot R_{12}(t) = 0 ; \qquad (7)$$

$$R_{22}(t) > 0, \; t > T \; ; \; \lim_{t \to \infty} R_{22}(t) = R_{22}(\infty) < \infty$$

Введём новые переменные $\tilde{a} = P^T a$, представим вектор $\tilde{a}^T = (\tilde{a}^T(1), \tilde{a}^T(2))$, где $\tilde{a}(1) \in R^{\ell}$, $\tilde{a}(2) \in R^{m-\ell}$ и определим многообразие

$$G = \{ \tilde{a} \in R^m, \; \tilde{a}(1) = 0 \}.$$

УТВЕРЖДЕНИЕ I. При выполнении условий (6),(7) оценки $\Omega_t^{(i)}$ алгоритма (2),(3), $\tau = \infty$ при $t \to \infty$ стремятся к многообразию G в том смысле, что

$$\lim_{t \to \infty} E \left(\tilde{\Omega}_t^{(i)}(1) - \tilde{\Omega}_*^{(i)}(1) \right) \left(\tilde{\Omega}_t^{(i)}(1) - \tilde{\Omega}_*^{(i)}(1) \right)^T = 0 \qquad (8)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Доказательство. Отметим, что вид алгоритма (2),(3) не изменится при ортогональном преобразовании координат и замене $\tilde{\Gamma}_{t+1}^{-1} = P^T \Gamma_{t+1}^{-1} P$. В новом базисе ковариационная матрица оценки $\tilde{\Sigma}_t^{(i)}$ характеризуется матрицей $\sum_{s=0}^{t} \tilde{\Phi}_s \tilde{\Phi}_s^T$. Применяя известную формулу обращения блочной матрицы, получим

$$\left(\sum_{s=0}^{t} \Phi_s \Phi_s^T\right)^{-1} = \begin{pmatrix} R_{11}^{-1}(t) + R_{11}^{-1}(t) R_{12}(t) H^{-1} R_{12}^T(t) R_{11}^{-1}(t) & -R_{11}^{-1}(t) R_{12}(t) H^{-1} \\ -H^{-1} R_{12}^T(t) R_{11}^{-1}(t) & H^{-1} \end{pmatrix}, \quad (9)$$

где матрица $H = R_{22}(t) - R_{12}^T(t) R_{11}^{-1}(t) R_{12}(t)$. Из (7) следует

$$\lim_{t \to \infty} \left(\sum_{s=0}^{t} \tilde{\Phi}_s \tilde{\Phi}_s^T\right)^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & R_{22}^{-1}(\infty) \end{pmatrix},$$

что и требовалось доказать.

Далее, пусть для входа Φ_t выполнено условие (5) и имеет место представление (6), в котором

$$R_{11}(t) = R_{11} \cdot t, \; R_{12}(t) = R_{12} \cdot t \cdot \varepsilon^p, \; R_{12}(t) = R_{22} \cdot t \cdot \varepsilon, \quad (10)$$

где $p \geq 1$, $\varepsilon > 0$ – малый параметр. Тогда в системе (2),(3), $\tau = \infty$ происходит разделение движений на быстрые $\tilde{\Sigma}_t^{(1)}$ и медленные $\tilde{\Sigma}_t^{(2)}$.

УТВЕРЖДЕНИЕ 2. При выполнении условий (5), (6), (10) для оценок алгоритма (2),(3), $\tau = \infty$ имеет место

$$E\left[\left(\tilde{\Sigma}_t^{(i)} - \tilde{\Sigma}_*^{(i)}\right)^T \tilde{a}\right]^2 = \begin{cases} c_1 \cdot t^{-1}, & \tilde{a} \in G^{\perp} \\ (c_2 + c_3 \cdot \varepsilon^{-1}) \cdot t^{-1}, & \tilde{a} \in G, \end{cases} \quad (11)$$

где $c_1, c_2, c_3 > 0$, $G^{\perp}$ – ортогональное дополнение к G.

Доказательство. Учитывая (10), из (9) получим

$$\left(\sum_{s=0}^{t} \tilde{\Phi}_s \tilde{\Phi}_s^T\right)^{-1} = t^{-1} \begin{pmatrix} R_{11} + \varepsilon^{2p} R_{11}^{-1} R_{12} H^{-1} R_{12}^T R_{11}^{-1} & -\varepsilon^p R_{11}^{-1} R_{12} H^{-1} \\ -\varepsilon^p H^{-1} R_{12}^T R_{11}^{-1} & H^{-1} \end{pmatrix},$$

где $H = \varepsilon\left(R_{22} - \varepsilon^{2p-1} R_{12}^T R_{11}^{-1} R_{12}\right)$, откуда следует (11).

В качестве иллюстрации рассмотрим два примера, представляющих самостоятельный интерес.

А. Пусть Φ_t, $t > 0$ есть реализация стационарной последовательности независимых случайных векторов с нулевым математическим ожиданием и ковариационной матрицей $K_\Phi = E\Phi_t \Phi_t^T$. Если K_Φ вырожденная, то имеет место представление (6), в котором

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$R_{11}(t) = R_{11} \cdot t > 0$, матрицы $R_{12}(t), R_{22}(t)$-нулевые и $G = Ker\, K_\Phi$. Произведя в (2),(3) замены переменных $\tilde{\Phi}_t = P^T \Phi_t$, $\tilde{\Gamma}_t = P^T \Gamma_t P$, получим уравнения

$$\begin{cases} \tilde{\Sigma}^{(i)}_{t+1}(1) = \tilde{\Sigma}^{(i)}_t(1) + \tilde{\Gamma}^{-1}_{t+1}(1)\tilde{\Phi}_t(1)\left(\chi^{(i)}_{t+1} - \tilde{\Sigma}^{(i)T}_t(1)\tilde{\Phi}_t(1)\right) \\ \tilde{\Sigma}^{(i)}_{t+1}(2) = \tilde{\Sigma}^{(i)}_t(2) \\ \tilde{\Gamma}_{t+1}(1) = \tilde{\Gamma}_t(1) + \tilde{\Phi}_t(1)\Phi^T_t(1), \end{cases}$$

которые показывают, что оценки $\Sigma^{(i)}_t$ сходятся к многообразию G и предельная точка зависит от начальных условий. Если ковариационная матрица K_Φ близка к вырожденной, т.е. имеет вид

$$K_\Phi = P^T \begin{pmatrix} R_{11} & R_{12}\cdot\varepsilon^p \\ R^T_{12}\cdot\varepsilon^p & R_{22}\cdot\varepsilon \end{pmatrix} P,$$

и $\varepsilon > 0$ - малый параметр, то имеет место быстрая сходимость $\tilde{\Sigma}^{(i)}_t(1)$ к $\tilde{\Sigma}^{(i)}_*(1)$ и относительно медленный дрейф $\tilde{\Sigma}^{(i)}_t(2)$ к $\tilde{\Sigma}^{(i)}_*(2)$.

Б. Пусть Φ_t есть состояние линейной динамической системы

$$\Phi_{t+1} = A\Phi_t + \eta_{t+1}, \quad t > 0. \qquad (12)$$

Если $\eta_t \equiv 0$, то в динамической системе (2),(3), $\tau = \infty$ можно выделить три разнотемповых движения: а) уменьшение дисперсии проекции оценки $\Sigma^{(i)}_t$ на подпространство неустойчивых состояний системы (12) по показательному закону; б) линейное уменьшение дисперсии проекции $\Sigma^{(i)}_t$ на подпространство собственных векторов матрицы A, которым отвечает собственное число $|\lambda| = 1$; в) дисперсия проекции $\Sigma^{(i)}_t$ на подпространство устойчивых состояний системы (12) стремится к $Const > 0$. Помимо этого, если вектор Φ_0 не общего вида, то для $\Sigma^{(i)}_t$ имеется некоторое предельное многообразие G.

УТВЕРЖДЕНИЕ 3. Для оценок $\Sigma^{(i)}_t$ алгоритма (2),(3), $\tau = \infty$, в котором Φ_t определяется уравнением (12), $\eta_t \equiv 0$ и $\Phi^T_0 f_\kappa \neq 0$, выполнено

$$E\left[\left(\Sigma^{(i)}_t - \Sigma^{(i)}_*\right)^T e_\kappa\right]^2 = \begin{cases} d_1 |\lambda_\kappa|^{-2t}, & |\lambda_\kappa| > 1 \\ d_2\, t^{-1}, & |\lambda_\kappa| = 1 \\ d_3, & |\lambda_\kappa| < 1. \end{cases} \qquad (13)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

где $d_1, d_2, d_3 > 0$, а собственные векторы матрицы A $\{e_1, e_2, \dots, e_m\}$ и векторы $\{f_1, f_2, \dots f_m\}$ образуют квазибиортогональные множества, $Ae_k = \lambda_k e_k$.

Доказательство. Представим матрицу A в виде

$$A = \sum_{i=1}^{m} \lambda_i e_i f_i^T, \quad e_i^T f_j = \delta_{ij}, \quad \sum_{i=1}^{m} e_i f_i^T = I_m. \qquad (14)$$

Так как Φ_t удовлетворяет (12), $\eta_t \equiv 0$ и $A^s = \sum_{i=1}^{m} \lambda_i^s e_i f_i^T$, то

$$\sum_{s=0}^{t} \Phi_s \Phi_s^T = \sum_{i,j=1}^{m} \alpha_{ij}(t) e_i e_j^T, \qquad (15)$$

где $\alpha_{ij}(t) = (f_i^T \Phi_0)(f_j^T \Phi_0) \sum_{s=0}^{t} \lambda_i^s \lambda_j^s$. Если для всех $i = \overline{1,m}$ $f_i^T \Phi_0 \neq 0$, то из (14), (15) следует

$$\Big(\sum_{s=0}^{t} \Phi_s \Phi_s^T\Big)^{-1} = \sum_{i,j=1}^{m} \beta_{ij}(t) f_i f_j^T,$$

где $\beta_{ij}(t)$ элементы матрицы, обратной к матрице с элементами $\alpha_{ij}(t)$, $i,j = \overline{1,m}$. Отсюда, учитывая, что

$$E\big[(\Omega_t^{(i)} - \Omega_*^{(i)})^T e_k\big]^2 = \sigma_i^2 e_k^T \Big(\sum_{s=0}^{t} \Phi_s \Phi_s^T\Big)^{-1} e_k$$

и используя (14), придём к (13). Если для $1 \le j_1, j_2, \dots j_\tau \le m$ $f_{j_s}^T \Phi_0 = 0$, то матрица $\sum_{s=0}^{t} \Phi_s \Phi_s^T$ имеет вид (6), в котором $R_{12}(t) \equiv 0, R_{22}(t) \equiv 0$, а G есть линейное подпространство, натянутое на векторы $\{f_{j_1}, f_{j_2}, \dots, f_{j_\tau}\}$. В этом случае оценки $\Omega_t^{(i)}$ сходятся к этому многообразию, причём для $k \neq j_s$, $s = \overline{1,\tau}$ имеет место (13).

Если в (12) $\eta_t \not\equiv 0$ – последовательность независимых случайных векторов, $E\eta_t = 0$, $E\eta_t \eta_t^T = \varepsilon \cdot K_\eta > 0$ и $\varepsilon > 0$ – малый параметр, то в системе (2),(3), $\tau = \infty$ имеет место разделение движений на быстрые – к многообразию G и медленные – вдоль него к $\Omega_*^{(i)}$.

Литература

I. Коган М.М., Неймарк Ю.И. Адаптивное управление по скорости убывания функции Ляпунова.- В кн.: Тезисы докладов IX Всесоюзного совещания по проблемам управления, Ереван. - М., ИПУ, 1983, с. I6I-I62.

2. Линник Ю.В. Метод наименьших квадратов и основы математико-статистической обработки наблюдений.- М.: Физматгиз, 1958.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ИССЛЕДОВАНИЕ АСИМПТОТИЧЕСКОГО ПОВЕДЕНИЯ ФУНКЦИЙ ОТ МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ И КОВАРИАЦИОННОЙ МАТРИЦЫ

Т.Х. - А.Колло

Одной из широко применяемых возможностей для увеличения робастности статистических оценок является использование асимптотических результатов. Хотя асимптотические оценки сопровождаются ошибками, зависящими от объема выборки n, во многих случаях эти приближённые результаты вполне пригодны для применения в практических целях. Они имеют более простой вид, чем точные распределения и формулы, и являются более стабильными относительно изменений в распределении исходной выборки. Часто асимптотическое распределение является единственным источником информации для построения доверительных интервалов, так как точное распределение статистики неизвестно даже в случае нормально распределенной генеральной совокупности (например, собственные значения и векторы выборочной корреляционной матрицы).

Известно, что для каждой достаточно гладкой функции выборочной средней или выборочной ковариационной матрицы можно установить сходимость по распределению к нормальному закону при достаточно широких предположениях. В данной статье нас интересует асимптотическое поведение статистики $g(\bar{X},S)$, зависящей от выборочной средней $\bar{X}$ и ковариационной матрицы S. Использованная при этом матричная аппаратура не будет представлена, так как в современной литературе (как в монографиях, так и в статьях) эти вопросы в достаточной мере изложены. При необходимости можно ссылаться на статьи [3, 5, 7]. Чтобы облегчить чтение приведем объяснения только двух обозначений.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

vec-оператор для $p \times q$-матрицы M определяется равенством

$$\operatorname{vec} M = (m_{11}, \dots, m_{p1}, m_{12}, \dots, m_{p2}, \dots, m_{1q}, \dots m_{pq})'.$$

Переставленной единичной матрицей $I_{p,q}$ называется $pq \times pq$-блок-матрица, составленная из $p \times q$-блоков так, что в ij - м блоке ji - й элемент равен единице, а остальные элементы нулевые. Используем обозначения

$$Z = \begin{pmatrix} \bar{X} \\ \hline \operatorname{vec} S \end{pmatrix} \quad ; \quad Z_0 = \begin{pmatrix} \mu \\ \hline \operatorname{vec} \Sigma \end{pmatrix} ,$$

где μ и Σ - математическое ожидание и ковариационная матрица распределения генеральной совокупности.

Парринг [4] доказала, что при предположении конечности четвертого момента исходного признак-вектора имеет место сходимость по распределению, если $n \to \infty$:

$$\sqrt{n}\,(Z - Z_0) \xrightarrow{\mathcal{D}} N(0, \Psi) ,$$

где блок-матрица

$$\Psi = \begin{pmatrix} \Sigma & \vdots & \bar{M}_3' \\ \hline \bar{M}_3 & \vdots & \mathfrak{X} \end{pmatrix} ; \qquad (1)$$

$$\mathfrak{X} = \bar{M}_4 - \operatorname{vec}\Sigma\,(\operatorname{vec}\Sigma)' , \qquad (2)$$

а

$$\bar{M}_3 = E[(X_i - \mu) \otimes (X_i - \mu)' \otimes (X_i - \mu)] ;$$

$$\bar{M}_4 = E[(X_i - \mu) \otimes (X_i - \mu)' \otimes (X_i - \mu) \otimes (X_i - \mu)'] .$$

Применяя теорему Андерсона (см. [1, с. 108] о сходимости по распределению, получим следующий результат.

ТЕОРЕМА 1. Пусть $\mathbb{X} = (X_1, \dots, X_n)$ - выборка объема n с моментами $EX_i = \mu$; $DX_i = \Sigma$; $M_4 < \infty$ и функция $g(x)$ пространства $R^{p(p+1)}$ в R^q имеет частные производные второго порядка в окружности точки Z_0. Тогда, если $n \to \infty$,

$$\sqrt{n}\,[g(Z) - g(Z_0)] \xrightarrow{\mathcal{D}} N(0, \omega) ,$$

где

$$\omega = \left(\left. \frac{dg(Z)}{dZ} \right|_{Z=Z_0} \right)' \Psi \left. \frac{dg(Z)}{dZ} \right|_{Z=Z_0} ,$$

Ψ определено равенством (1), а матричная производная $\left. \frac{dg(Z)}{dZ} \right|_{Z=Z_0} \neq 0$ определена в смысле МакРаэ [5]:

$$\frac{dY}{dX} = Y \otimes \frac{d}{dX} .$$

Дальнейшие упрощения в общем случае невозможны, так как матрица ω зависит от конкретного вида функции $g(x)$. Рассмотрим,

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

какой вид имеет утверждение теоремы I в одном конкретном случае – если $g(\bar{z}) = \bar{x}'S^{-1}\bar{x}$. В этом случае (см. [2])

$$\frac{d(\bar{x}'S^{-1}\bar{x})}{d\bar{z}} = \begin{pmatrix} 2S^{-1}\bar{x} \\ -(S^{-1}\bar{x}) \otimes (S^{-1}\bar{x}) \end{pmatrix}$$

и при предположении $\mu = 0$ предельная ковариационная матрица

$$\pi = \begin{pmatrix} 2\Sigma^{-1}\mu \\ -(\Sigma^{-1}\mu) \otimes (\Sigma^{-1}\mu) \end{pmatrix}' \begin{pmatrix} \Sigma & \bar{M}_3 \\ \bar{M}_3' & \bar{M}_4 \end{pmatrix} \begin{pmatrix} 2\Sigma^{-1}\mu \\ -(\Sigma^{-1}\mu) \otimes (\Sigma^{-1}\mu) \end{pmatrix}. \quad (3)$$

Если X_i - симметричные случайные векторы, то $\bar{M}_3 = 0$ и получается

$$\pi = 4(\mu'\Sigma^{-1}\mu) + (\mu'\Sigma^{-1}) \otimes (\mu'\Sigma^{-1}) \bar{M}_4 (\Sigma^{-1}\mu) \otimes (\Sigma^{-1}\mu). \quad (4)$$

ТЕОРЕМА 2. Пусть $\mathcal{X} = (X_1, \ldots, X_n)$ - выборка объема n с моментами $EX_i = \mu \neq 0$; $DX_i = \Sigma$, $\bar{M}_4 < \infty$. Тогда для функции $\bar{x}'S^{-1}\bar{x}$ имеет место сходимость при $n \to \infty$

$$\sqrt{n}\,(\bar{x}'S^{-1}\bar{x} - \mu'\Sigma^{-1}\mu) \xrightarrow{\mathcal{D}} N(0, \pi), \quad (5)$$

где π определен равенством (3). Если X_i - симметричные случайные величины, то предельная ковариационная матрица имеет вид (4).

В частном случае, если $X_i \sim N(\mu, \Sigma)$, получим из теоремы 2 следующее следствие.

СЛЕДСТВИЕ 2.1. Если $X_i \sim N(\mu, \Sigma)$; $\mu \neq 0$, то при $n \to \infty$

$$\sqrt{n}\,(\bar{x}'S^{-1}\bar{x} - \mu'\Sigma^{-1}\mu) \xrightarrow{\mathcal{D}} N(0, \varphi),$$

где

$$\varphi = 2\mu'\Sigma^{-1}\mu\,(2 + \mu'\Sigma^{-1}\mu). \quad (6)$$

Следствие получается из теоремы с учетом , что в случае нормальной выборки $\bar{M}_3 = 0$ и

$$\bar{M}_4 = (I_{p^2} + I_{p,p})(\Sigma \otimes \Sigma) + \mathrm{vec}\,\Sigma\,(\mathrm{vec}\,\Sigma)'.$$

Приведенный пример является поучительным в нескольких аспектах. Во-первых, следует отдельно исследовать случай $\mu = 0$. Тогда

$$\left.\frac{d(\bar{x}'S^{-1}\bar{x})}{d\bar{z}}\right|_{\bar{z}=\bar{z}_0} = 0$$

и вместо сходимости к нормальному закону имеет место сходимость к закону χ^2_p.

ТЕОРЕМА 3. Пусть $\mathcal{X} = (X_1, \ldots, X_n)$ - выборка объема n с моментами $EX_i = 0$; $DX_i = \Sigma$. Если $n \to \infty$, то

$$n\,\bar{x}'S^{-1}\bar{x} \xrightarrow{\mathcal{D}} \chi^2_p. \quad (7)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Результат получим, разложив $\bar{X}'S^{-1}\bar{X}$ в ряд Тейлора, откуда вытекает, что $\bar{X}'S^{-1}\bar{X}$ и $\bar{X}\Sigma^{-1}\bar{X}$ имеют одно и то же предельное распределение, если $\mu = 0$. Но по [6] имеет место сходимость

$$n\,\bar{X}'\Sigma^{-1}\bar{X} \xrightarrow{\mathcal{D}} \chi^2_p.$$

Это значит, что в процессе $\mu \to 0$ сходимость (5) становится все хуже и хуже и на пределе не имеет места.

Второй момент, который следует учитывать, это скорость сходимости и смещенность оценки $\bar{X}'S^{-1}\bar{X}$ относительно $\mu'\Sigma^{-1}\mu$. Разлагая $\bar{X}'S^{-1}\bar{X}$ в ряд Тейлора в точке Z_0, получим для смещения следующее выражение:

$$E(\bar{X}'S^{-1}\bar{X} - \mu'\Sigma^{-1}\mu) = \frac{p}{n} + \frac{1}{\lambda_1}\,(\mu\otimes\mu)'(I_{p^2}\otimes(\mathrm{vec}\,I_p)')\times$$
$$\times\{\Sigma^{-1}\otimes\mathrm{vec}\,\Sigma^{-1}(\mathrm{vec}\,\Sigma^{-1})'\otimes I_p + (I_{pp}\otimes I_p)\,\Sigma^{-1}\otimes\mathrm{vec}\,\Sigma^{-1}(\mathrm{vec}\,\Sigma^{-1})'(I_{pp}\otimes I_p)\otimes I_p\}\times$$
$$\times\,(I_{p,p^2}\otimes I_p + I_{p,p^3})\,(\mathrm{vec}\,\Sigma\otimes\mathrm{vec}\,\Sigma) + o\left(\frac{1}{\sqrt{n}}\right).$$

Сходимость распределения $\bar{X}'S^{-1}\bar{X}$ в случае нормально распределенной выборки исследовалась методом Монте - Карло сотрудником ТГУ Ю.Вилисмяэ. При обозначении

$$y_n = \sqrt{n}\,(\bar{X}'S^{-1}\bar{X} - \mu'\Sigma^{-1}\mu)$$

на основе κ выборок с параметрами

$$\mu = \begin{pmatrix}1\\1\\1\end{pmatrix};\quad \Sigma = \begin{pmatrix}1 & 0{,}1 & 0{,}2\\ 0{,}1 & 1{,}1 & 0{,}3\\ 0{,}2 & 0{,}3 & 1{,}2\end{pmatrix};\quad \varphi = 16{,}4889$$

получены следующие результаты.

Таблица I

κ	I000	300	I00	I00
n	200	I000	2000	3000
$\widehat{E y_n}$	0,8547	0,0780	0,9027	0,5204
$\widehat{D y_n}$	I6,875	I6,097	I5,087	I5,726

Как показывает сравнение эмпирического и теоретического распределений, методом хи -квадрат можно даже после исключения смещения заметить большое различие между этими распределениями.

В итоге можно сказать, что сходимость по распределению далеко не всегда является удовлетворительным решением при построении доверительных областей. Особенно внимательным следует быть в тех случаях, когда первая производная $\frac{dg(\bar{z})}{d\bar{z}}\Big|_{\bar{z}=\bar{z}_0}$ превращается в нуль в некоторых точках $\bar{z}_0$, которые, как правило, являются параметрами распределения или содержат их.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Литература

1. Андерсон Т.В. Введение в многомерный статистический анализ. М., 1963.
2. Колло Т. Асимптотическое распределение функции $\bar{x}'\bar{s}'\bar{x}$. - Тр. / ВЦ ТГУ, 1986, вып. 53, с. 15 - 21.
3. Колло Т., Кинкар Т. Матричная производная с применением для блок-матриц. - Тр. / ВЦ ТГУ, Тарту, 1984, вып. 51, с. 96-107.
4. Парринг А.-М. Вычисление асимптотических характеристик функций выборки. - Уч. зап. / Тартуск. ун-т , 1979, т. 492, с. 86-90.
5. Mac Rae E.C. Matrix derivatives with an application to an adaptive linear decision problem. Ann. Stat., 1974, 2, 337-346.
6. Moore D.S. Chi-square tests. - In: Hogg R.V., ed. Studies in Statistics, MAA Studies in Mathematics, 19, 66-106.
7. Treat I. Matrix calculus for multivariate distributions. Acta et Commentationes Universitatis Tartuensis, 1986, vol. 733, p. 64-84.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

К ПОСЛЕДОВАТЕЛЬНОМУ ОЦЕНИВАНИЮ ПАРАМЕТРОВ СТОХАСТИЧЕСКИХ ДИНАМИЧЕСКИХ СИСТЕМ

В.В.Конев, С.М.Пергаменщиков

I. Постановка задачи

Рассмотрим задачу оценивания матрицы неизвестных параметров $A = \|a_{ij}\|$ размера $S \times P$ по наблюдениям многомерных процессов X_k и φ_k $(k = 0, 1, \ldots)$, связанных уравнением

$$X_{k+1} = A\varphi_k + \varepsilon_{k+1} . \qquad (I.1)$$

Здесь $\{\varepsilon_k\}$ – последовательность шумов, для которой выполнены условия:

1) $M\|\varepsilon_k\|^2 < \infty$;

2) $M(\varepsilon_{k+1} \mid \varphi_0, \ldots, \varphi_k, X_1, \ldots, X_k) = 0$, $M(\varepsilon_{k+1}\varepsilon'_{k+1} \mid \varphi_0, \ldots, \varphi_k, X_1, \ldots X_k) = \Gamma_k$.

Распределения ε_k неизвестны, Γ_k – неотрицательно определенная функция от наблюдений, которая может быть неизвестной, при этом однако предполагается, что известна некоторая мажоранта ее следа, причем не ограничивая общности, можно считать, $\mathrm{tr}\,\Gamma_k \le 1$. Требуется построить оценки параметров a_{ij} с заданным среднеквадратическим уклонением.

В работах [1,2] был предложен метод идентификации параметров динамических систем (I.1), который приводит к оценкам с заданной среднеквадратической точностью. Последовательные планы [1,2] дают решение гарантированного оценивания параметров устойчивого процесса авторегрессии (АР).

Если требование устойчивости процесса АР не выполняется, то, как можно показать, двухэтапная процедура [2] оказывается, вообще говоря, неработоспособной.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Основная цель данной работы – построить гарантированные оценки параметров АР, пригодные как для устойчивого, так и для неустойчивого процесса АР.

2. Построение гарантированной оценки

Для решения поставленной задачи будем использовать подход с позиций последовательного анализа. Последовательный план оценивания матрицы A будем строить на основе оценок по методу наименьших квадратов (МНК). Ориентируясь на получение гарантированной оценки параметров неустойчивого процесса АР, целесообразно в качестве исходной оценки A взять оценку МНК по наблюдениям из скользящего окна: $\varphi_\ell, \dots, \varphi_N, x_{\ell+1}, \dots, x_{N+1}$, $0 \le \ell \le N$,

$$\hat{A}_{\ell,N} = \sum_{k=\ell}^{N} x_{k+1} \varphi_k' \, G_{\ell,N}^{-1}, \quad G_{\ell,N} = \sum_{k=\ell}^{N} \varphi_k \varphi_k' . \qquad (2.1)$$

Чтобы получить гарантированную оценку матрицы A, необходимо контролировать не только дисперсионные свойства помех, но и поведение матрицы $G_{\ell,N}^{-1}$. Такой контроль можно установить, используя систему модифицированных оценок (2.1):

$$\tilde{A}_n = \sum_{k=\nu_n}^{\tau_n} \alpha_n(k) x_{k+1} \varphi_k' \, \tilde{G}^{-1}_{\nu_n, \tau_n}, \quad \tilde{G}_{\nu_n, \tau_n} = \sum_{k=\nu_n}^{\tau_n} \alpha_n(k) \varphi_k \varphi_k' . \qquad (2.2)$$

Здесь ν_n и τ_n $(0 \le \nu_n \le \tau_n)$ – неупреждающие целочисленные функции от наблюдений x_k и φ_k моменты остановки, причем τ_n определяется по формуле

$$\tau_n = \inf\{N \ge \nu_n : \operatorname{tr} R_n^{-1} \sum_{k=\nu_n}^{N} \varphi_k \varphi_k' \ge \varkappa_n\} \qquad (2.3)$$

$(\inf\{\varnothing\} = \infty)$, где R_n – некоторая симметричная положительно определенная матрица, зависящая от наблюдаемых процессов на временном интервале $[0, \nu_n]$; $\{\varkappa_n\}$ – неубывающая последовательность положительных чисел с $\lim \varkappa_n = \infty$. Весовые функции в (2.2) определяются равенством

$$\alpha_n(k) = \begin{cases} 1, & 0 \le k < \tau_n ; \\ \alpha_n, & k = \tau_n ; \end{cases}$$

где α_n находится из уравнения

$$\operatorname{tr} R_n^{-1} \sum_{k=\nu_n}^{\tau_n - 1} \varphi_k \varphi_k' + \alpha_n \operatorname{tr} R_n^{-1} \varphi_{\tau_n} \varphi_{\tau_n}' = \varkappa_n .$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Последовательный план оценивания матрицы A представляет собой пару $(N(H), A^*(H))$, в которой длительность наблюдений $N(H)$ и оценка $A^*(H)$ определяются по формулам:

$$N(H) = \tau_\sigma , \quad A^*(H) = \Big(\sum_{k=1}^{\sigma} \beta_k^2\Big)^{-1} \sum_{k=1}^{\sigma} \beta_k^2 \tilde{A}_k , \tag{2.4}$$

где

$$\sigma = \sigma(H) = \inf\Big\{ \ell \geq 1 : \sum_{k=1}^{\ell} \beta_k^2 \geq H \Big\},$$

$$\beta_n^2 = \begin{cases} 1 / \operatorname{tr}[\tilde{G}^{-1}_{\nu_n, \tau_n} R_n]^2, & \text{если } \tilde{G}_{\nu_n, \tau_n} \text{ - не вырождена;} \\ 0 , & \text{в противном случае.} \end{cases} \tag{2.5}$$

Для конечности $N(H)$ потребуем, чтобы с вероятностью единица выполнялись условия

$$\sum_{k \geq 1} \|\varphi_k\|^2 = +\infty , \quad \sum_{k \geq 1} \beta_k^2 = +\infty . \tag{2.6}$$

ТЕОРЕМА I. Пусть для процесса (I.I) последовательности γ_n и R_n выбраны так, что расходятся ряды (2.6) и, кроме того, $\operatorname{tr} R_n^{-1} \leq 1/c_n$ для некоторой детерминированной последовательности положительных чисел $\{c_n\}$, удовлетворяющих неравенству

$$\rho = \sum \alpha_n / c_n < \infty .$$

Тогда для любого $H>0$ последовательный план (2.4) обладает свойствами:

1) $N(H) < \infty$ $\quad$ п.н.;

2) $M_A \|A^*(H) - A\|^2 \leq \rho / H$.

Доказательство. Утверждение I следует из расходимости рядов (2.6). Используя (I.I) в (2.2), находим уклонение последовательной оценки (2.4):

$$A^*(H) - A = \Big(\sum_{k=1}^{\sigma} \beta_k^2\Big)^{-1} \sum_{k=1}^{\sigma} \beta_k^2 (\tilde{A}_k - A) = \tag{2.7}$$

$$= \Big(\sum_{k=1}^{\sigma} \beta_k^2\Big)^{-1} \sum_{k=1}^{\sigma} \beta_k^2 \ell_k B_k R_k^{-1/2} ,$$

где

$$\ell_k = \Big(\sum_{j=\nu_k}^{\tau_k} \alpha_k(j) \varepsilon_{j+1} \varphi_j'\Big) R_k^{-1/2},$$

$$B_k = R_k^{1/2} \tilde{G}^{-1}_{\nu_k, \tau_k} R_k^{1/2} . \tag{2.8}$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Далее воспользуемся следующим неравенством

$$M_A \|\gamma_k\|^2 \le \varkappa_k ,$$

которое доказывается по той же схеме, что и основная теорема в [I]. С помощью этого неравенства и определения момента σ получим второе утверждение теоремы I

$$M_A \|A^*(H) - A\|^2 \le M_A \Big[\Big(\sum_{k=1}^{\sigma} B_k^2 \Big)^{-1} \sum_{k=1}^{\sigma} B_k^2 \|\gamma_k B_k R_k^{-1/2}\| \Big]^2 \le$$

$$\le M_A \Big(\sum_{k=1}^{\sigma} B_k^2 \Big)^{-1} \sum_{k=1}^{\sigma} B_k^2 \|\gamma_k\|^2 \|B_k\|^2 \|R_k^{-1/2}\|^2 \le$$

$$\le \frac{1}{H} \sum_{k=1}^{\infty} M_A \|\gamma_k\|^2 \operatorname{tr} R_k^{-1} \le \frac{1}{H} \sum_{k=1}^{\infty} \varkappa_k / c_k .$$

При выводе этих неравенств учитывалось, что $B_k^2 \|B_k\|^2 \le 1$ и $\operatorname{tr} R_k^{-1} \le 1/c_k$. Теорема I доказана.

Замечание I. Как следует из теоремы I, оценка $A^*(H)$ сходится к A при $H \to \infty$, при $n \to \infty$ последовательность $A^*(H_n)$ сходится с вероятностью единица, если $\sum 1/H_n < \infty$.

Замечание 2. Последовательный план (2.4) определен с точностью до последовательностей γ_n и R_n, которые должны выбираться так, чтобы выполнялось второе условие (2.6). В случае, когда $\gamma_n = 0$ и $R_n = c_n E$, план (2.4) совпадает с последовательной процедурой, предложенной в [2]. При этом условия (2.6) выполняются, например, для устойчивого процесса авторегрессии. Если процесс авторегрессии является неустойчивым, то для таких последовательностей γ_n и R_n, как можно показать, второе условие (2.6) нарушается. Ниже установлено, что при определенном выборе γ_n, R_n утверждение теоремы I остается верным и для неустойчивого процесса авторегрессии.

3. Оценивание параметров авторегрессии

Пусть $\{y_n\}$ - процесс авторегрессии p- го порядка

$$y_{n+1} = a_1 y_n + \ldots + a_p y_{n-p+1} + \xi_{n+1}, \qquad (3.1)$$

где ξ_n - последовательность независимых, одинаково распределенных случайных величин с $M\xi_n = 0$, $M\xi_n^2 = \sigma^2$. Используя обозначения

$$x_n = (y_n, \ldots, y_{n-p+1})', \quad \varepsilon_n = (\xi_n, \ldots 0)',$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

представим процесс (3.1) в векторной форме

$$X_{n+1} = A X_n + \varepsilon_{n+1}, \qquad (3.2)$$

где

$$A = \begin{bmatrix} a_1, \ldots, a_p \\ E_{p-1}, 0 \end{bmatrix},$$

E_p - единичная матрица размера p.

Последовательный план оценивания матрицы A процесса (3.2) определяется по формулам (2.4) при $y_k = x_k$. Последовательность $\{\nu_n\}$ в этих формулах определим равенством

$$\nu_n = \inf\{m \geq 0 : g_m \leq 1/c_n\}, \qquad (3.3)$$

где

$$g_m = \begin{cases} \operatorname{tr} G_m^{-1}, & \text{если } G_m \text{ - не вырождена;} \\ +\infty, & \text{в противном случае;} \end{cases}$$

$R_n = G_{\nu_n}$, $G_k = G_{0,k}$.

Свойства этих последовательных планов оценивания матрицы так же, как и процесса (3.2), зависят от значений ее характеристических корней. Ниже $\lambda_1(A)$ и $\lambda_p(A)$ обозначают максимальный и минимальный по модулю характеристические корни матрицы A.

ТЕОРЕМА 2. Пусть для матрицы A выполнено одно из условий:

1. $|\lambda_p(A)| > 1$;
2. $|\lambda_1(A)| < 1$.

Тогда для любого $H > 0$ последовательный план оценивания матрицы A процесса (3.2) определяемый по формулам (2.4), (3.3) обладает свойствами:

1) $N(H) < \infty$ п.н.;

2) $M_A \| A^*(H) - A \|^2 \leq \rho / H$, $\rho = \sum_{n \geq 1} \alpha e_n / c_n$.

Доказательство. Покажем, что для процесса (3.2) выполнены условия теоремы 1. Расходимость первого ряда в условии (2.6) можно доказать, используя лемму Бореля-Кантелли. Для проверки второго условия (2.6) в первом случае нам потребуется следующая лемма из [3].

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ЛЕММА I. При условии $|\lambda_p(A)|>1$ для процесса (3.2) с вероятностью единица справедливы предельные соотношения

$$\lim_{N\to\infty} A^{-N} G_N (A')^{-N} = F, \quad \lim_{N\to\infty} A^{-N} X_N = Z, \quad \|Z\| < \infty,$$

$$\lim_{N\to\infty} \frac{1}{N} \log \lambda_1(G_N) = 2\log|\lambda_1(A)|, \qquad (3.4)$$

$$\lim_{N\to\infty} \frac{1}{N} \log \lambda_p(G_N) = 2\log|\lambda_p(A)|,$$

где $\lambda_p(A)$ и $\lambda_1(A)$ обозначают минимальный и максимальный характеристические корни матрицы A ; F - п.н. положительно определенная матрица.

Из леммы I следует конечность с вероятностью единица моментов ν_n и z_n для любого n. Заметим, что при этом

$$\lim_{n\to\infty} (z_n - \nu_n) = +\infty \qquad \text{п.н.} \qquad (3.5)$$

Действительно, согласно (2.3) и (3.3) имеем

$$\varkappa_n = \operatorname{tr} G_{\nu_n}^{-1} \sum_{j=\nu_n}^{z_n} \alpha_n(j) X_j X_j' = \operatorname{tr} F_{2,n} (A')^{z_n-\nu_n} F_{1,n}^{-1} A^{z_n-\nu_n} =$$

$$= \| F_{2,n}^{1/2} (A')^{z_n-\nu_n} F_{1,n}^{-1/2} \|^2 \le \operatorname{tr} F_{2,n} \operatorname{tr} F_{1,n}^{-1} \| A^{z_n-\nu_n} \|^2,$$

где

$$F_{1,n} = A^{-\nu_n} G_{\nu_n} (A')^{-\nu_n}, \quad F_{2,n} = A^{-z_n} \tilde{G}_{\nu_n, z_n} (A')^{-z_n}.$$

Поскольку $\lim \varkappa_n = \infty$, то из последнего неравенства получаем (3.5).

Используя (3.5), докажем неравенство

$$\overline{\lim_n} B_n^2 > 0 \qquad \text{п.н.} \qquad (3.6)$$

С помощью определения (2.5) предельного соотношения (3.5), и леммы I с вероятностью единица справедливо неравенство

$$\overline{\lim_{n\to\infty}} B_n^2 > \lim_{n\to\infty} 1 / \operatorname{tr}[G_{\nu_n} G_{\nu_n, z_n-1}^{-1}]^2 = \infty,$$

из которого следует (3.6) и справедливость первого утверждения теоремы 2.

Пусть выполнено условие 2, т.е. процесс АР устойчив. Покажем,

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

что с вероятностью единица расходится второй ряд (2.6). Известно [4], что в этом случае с вероятностью единица справедливо предельное соотношение

$$\lim_{N\to\infty} \frac{G_N}{N} = F \qquad \text{п.н.}, \qquad (3.7)$$

причем F - положительно определенная матрица.

Проверим условие (3.6). Имеем

$$tr\, B_n^2 = tr\left[G_{\nu_n} \tilde{G}^{-1}_{\nu_n, z_n} \right]^2 = tr\left[\frac{G_{\nu_n}}{\nu_n} \frac{\nu_n}{z_n - \nu_n} (z_n - \nu_n) \tilde{G}^{-1}_{\nu_n, z_n} \right]^2. \qquad (3.8)$$

Покажем, что отношение $\nu_n / (z_n - \nu_n)$ асимптотически ограничено. Из определения момента z_n следует, что

$$\mathcal{E}_n \le tr\, G_{\nu_n}^{-1} \sum_{j=\nu_n}^{z_n} X_j X_j' = \frac{z_n}{\nu_n} tr\, \nu_n G_{\nu_n}^{-1} \frac{1}{z_n} G_{z_n} - tr\, \nu_n G_{\nu_n}^{-1} \frac{1}{\nu_n} G_{\nu_n - 1}.$$

Отсюда, переходя к пределу, получим

$$\lim_{n\to\infty} \frac{\nu_n}{z_n - \nu_n} = 0 \qquad \text{п.н.},$$

поскольку $\lim \mathcal{E}_n = \infty$. Используя этот факт и учитывая (3.7), находим

$$\lim_{n\to\infty} \frac{1}{z_n - \nu_n} \tilde{G}_{\nu_n, z_n} = F \qquad \text{п.н.} \qquad (3.9)$$

Из (3.8) и (3.9) следует справедливость (3.6). Теорема 2 доказана.

Поскольку длительность $N(H)$ последовательной процедуры оценивания является случайной величиной, то представляет интерес изучение ее свойств. В следующих теоремах установлена точная асимптотическая граница для времени оценивания параметров авторегрессии.

При этом последовательности $\{C_n\}$ и $\{\mathcal{E}_n\}$ определяются равенствами:

$$C_n = \begin{cases} C(H), & n \le \Psi(H); \\ C_n^*, & n > \Psi(H); \end{cases} \qquad \mathcal{E}_n = \begin{cases} \mathcal{E}(H), & n \le \Psi(H); \\ \mathcal{E}_n^*, & n > \Psi(H), \end{cases} \qquad (3.10)$$

где $\Psi(H) = [H] + 1$ ($[H]$ обозначает целую часть числа H), $C(H)$ и $\mathcal{E}(H)$ связаны соотношением $C(H) = H\,\mathcal{E}(H)$, причем $\mathcal{E}(H) = O(\ln H)$ и $\lim \mathcal{E}(H) = \infty$ при $H \to \infty$; $\{C_n^*\}$ и $\{\mathcal{E}_n^*\}$ - неубывающие пос-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ледовательности положительных чисел такие, что

$$\sum_{n\geq 1} \varkappa_n^* / C_n^* < \infty .$$

Такой выбор C_n и $\varkappa_n$ обеспечивает, как можно проверить, асимптотическое совпадение последовательной оценки $A^*(H)$ с оценкой $\tilde{A}_1$, которая вычисляется по формуле (2.2), используя значения процесса (3.2) на первом интервале наблюдений $[\gamma_1, \tau_1]$.

ТЕОРЕМА 3. Пусть все характеристические корни матрицы A по модулю больше единицы. Тогда

$$\lim_{H\to\infty} \frac{N(H)}{\log_{|\lambda_1(A)|^2} H} = 1 \qquad \text{п.н.}$$

Доказательство. Обозначим через $\tilde{\tau}(H) = \tau_1$. В дальнейшем нам потребуется следующая

ЛЕММА 2. В условиях теоремы 3 выполняется предельное соотношение

$$\lim_{H\to\infty} [N(H) - \tilde{\tau}(H)] = 0 \qquad \text{п.н.} \qquad (3.11)$$

Доказательство лемм. 2. В силу выбора последовательностей C_n и $\varkappa_n$ в виде (3.10) имеем

$$N(H) = \tau_\sigma = \tilde{\tau}(H)\,\chi[\sigma \leq \psi(H)] + \tau_\sigma\,\chi[\sigma > \psi(H)],$$

где $\tilde{\tau}(H)$ совпадает с величиной τ_1, определяемой формулой (2.3) при $\varkappa_1 = \varkappa(H)$. Из этого равенства следует, что для справедливости (3.11) достаточно, чтобы

$$\lim_{H\to\infty} [\sigma(H) > \psi(H)] = \emptyset \qquad \text{п.н.} \qquad (3.12)$$

С другой стороны, поскольку по определению σ и в силу выбора C_n и $\varkappa_n$ выполняется неравенство

$$\sum_{k=1}^{\sigma} \beta_k^2 = B^2(H)\,\psi(H) + \sum_{k=\psi(H)+1}^{\sigma} \beta_k^2 \geq H,$$

($B^2(H)$ совпадает с величиной β_1^2, определяемой (2.5) при последовательностях (3.10)), то (3.12) будет установлено, если показать, что

$$\lim_{H\to\infty} B^2(H) > 1 \qquad \text{п.н.} \qquad (3.13)$$

Так как по условию теоремы $\lim \varkappa(H) = \infty$ при $H \to \infty$, то справедли-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

вость (3.13) вытекает из доказательства предельного соотношения (3.6). Лемма 2 доказана.

Докажем теперь утверждение теоремы 3. Имеем

$$\frac{N(H)}{\log_{|\lambda_p(A)|^2} H} = \frac{N(H) - \tilde{\tau}(H)}{\log_{|\lambda_p(A)|^2} H} + \mathcal{D}_1(H) + \mathcal{D}_2(H), \qquad (3.14)$$

где

$$\mathcal{D}_1(H) = (\tilde{\tau}(H) - \nu_1) / \log_{|\lambda_p(A)|^2} H, \quad \mathcal{D}_2(H) = \nu_1 / \log_{|\lambda_p(A)|^2} H.$$

Первое слагаемое в (3.14) при $H \to \infty$ сходится к нулю в силу леммы 2 и условия $|\lambda_p(A)| > 1$. Покажем справедливость предельных соотношений

$$\lim \mathcal{D}_1(H) = 0, \quad \lim \mathcal{D}_2(H) = 1, \qquad (3.15)$$

которые вместе с (3.14) влекут утверждение теоремы 3. Представим $\mathcal{D}_1(H)$ в виде

$$\mathcal{D}_1(H) = \frac{\tilde{\tau}(H) - \nu_1}{\log_{|\lambda_1(A)|^2} \varkappa(H)} \cdot \frac{\log_{|\lambda_1(A)|^2} \varkappa(H)}{\log_{|\lambda_1(A)|^2} H}, \qquad (3.16)$$

Поскольку по условию $\varkappa(H) = O(\ln H)$, то второй сомножитель в правой части (3.16) сходится к нулю при $H \to \infty$. Поэтому, чтобы доказать первое равенство в (3.15), достаточно установить ограниченность величины

$$d(H) = \frac{\tilde{\tau}(H) - \nu_1}{\log_{|\lambda_1(A)|^2} \varkappa(H)}.$$

Из определения момента остановки $\tilde{\tau}(H) = \tau_1$ следует равенство

$$\varkappa(H) = \| F_{2,1}^{1/2} (A')^{\tau_1 - \nu_1} F_{1,1}^{-1/2} \|^2,$$

из которого получаем

$$\frac{\log_{|\lambda_1(A)|^2} \varkappa(H)}{\tilde{\tau}(H) - \nu_1} = 1 + \frac{\log_{|\lambda_1(A)|^2} \tau(H)}{\tilde{\tau}(H) - \nu_1}, \qquad (3.17)$$

где

$$\tau(H) = |\lambda_1(A)|^{-2(\tau_1 - \nu_1)} \| F_{1,1}^{-1/2} A^{\tau_1 - \nu_1} F_{2,1}^{1/2} \|^2.$$

Используя разложение Жордана, можно доказать, что второе слагаемое в (3.17) сходится к нулю с вероятностью единица при $H \to \infty$. Таким образом, первое равенство в (3.15) установлено. Далее представим $\mathcal{D}_2(H)$ в виде

$$\mathcal{D}_2(H) = (\nu_1 / \log_{|\lambda_p(A)|^2} C(H)) \cdot (\log_{|\lambda_p(A)|^2} C(H) / \log_{|\lambda_p(A)|^2} H).$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

В силу условий $C(H) = H\varkappa(H)$ и $\varkappa(H) = O(\ln H)$ второй член в этом равенстве стремится к единице при $H \to \infty$. Сходимость первого сомножителя к единице в этом равенстве следует из определения момента ν_1 и предельного соотношения (3.4). Теорема 3 доказана.

ТЕОРЕМА 4. Пусть для процесса авторегрессии первого порядка (3.2) с параметром $|a|>1$ выполнено условие: плотность распределения шумов ограничена некоторой константой L. Тогда среднее время последовательного оценивания удовлетворяет неравенству

$$M_a N(H) \leq \log_{a^2} H + \Delta(a),$$

где $\Delta(a)$ - некоторая функция от a.

Доказательство. Заметим, что для процесса авторегрессии первого порядка (3.2) последовательная процедура (2.4) совпадает с одноэтапной процедурой, предложенной в [I], а время оценивания определяется формулой

$$N(H) = \inf\{K \geq 0 : \sum_{i=0}^{K} x_i^2 \geq H\}.$$

Тогда

$$M_a N(H) = 1 + \sum_{k \geq 1} P_a\{\sum_{i=0}^{k} x_i^2 < H\} \leq 1 + \sum_{k \geq 1} P_a\{x_k^2 < H\} =$$

$$= 1 + \sum_{k \geq 1} P_a\{|z_k| \leq \sqrt{H}\, a^{-k}\},$$

где

$$z_k = \frac{x_k}{a^k} = \sum_{i=0}^{k} a^{-i}\varepsilon_i .$$

Поскольку по условию теоремы плотность распределения ε_i ограничена L, то для любого $k \geq 0$ и $x \geq 0$

$$P_a\{|z_k| \leq x\} \leq L\, x . \qquad (3.18)$$

Учитывая неравенство (3.18), получим

$$M_a N(H) \leq \log_{a^2} H + 1 + \sum_{k > \log_{a^2} H} P_a\{|z_k| \leq x\} \leq \log_{a^2} H + \Delta(a),$$

где

$$\Delta(a) = 1 + L \sum_{k \geq 0} a^{-k}.$$

Теорема 4 доказана.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ТЕОРЕМА 5. Для устойчивого процесса авторегрессии (3.2) ($|\lambda_1(A)| < 1$) последовательный план оценивания (2.4) обладает свойствами:

1. $\lim\limits_{H\to\infty} \frac{N(H)}{c(H)\varkappa(H)} = \frac{1}{p} tr F^{-1}$ п.н.

2. Вектор $\sqrt{N(H)}(a^*(H)-a)$ является асимптотически нормальным с параметрами $(0, F^{-1})$, где F - предельная матрица в (3.7), вектор $a=(a_1,\dots,a_p)'$ - неизвестных параметров процесса (3.1), $a^*(H)$ - первый столбец матрицы $[A^*(H)]'$.

Доказательство. В силу выбора последовательностей $\{c_n\}$ и $\{\varkappa_n\}$ в виде (3.10) имеем

$$N(H) = \tilde{\tau}(H)\chi[\sigma \le \psi(H)] + N(H)\chi[\sigma > \psi(H)],$$

где

$$\tilde{\tau}(H) = \tau_1.$$

Поскольку $\lim \varkappa_1 = \lim \varkappa(H) = \infty$ при $H \to \infty$, то справедливо предельное равенство

$$\lim_{H\to\infty} \frac{\nu_1}{\tau_1 - \nu_1} = 0 \quad \text{п.н.,} \qquad (3.19)$$

которое доказывается подобно (3.9).

Далее из определения $B(H) = B_1$, (3.8) и (3.19) вытекает, что

$$\lim B(H) = \infty \quad \text{п.н.,}$$

а следовательно, справедливы соотношения (3.12) и (3.11). Из определения τ_1 легко получить с учетом (3.9), что

$$\lim_{H\to\infty} \tilde{\tau}(H)/\varkappa(H)\nu_1 = \frac{1}{p} \quad \text{п.н.} \qquad (3.20)$$

Асимптотическое поведение ν_1 определяется функцией $c_1 = c(H)$ при $H \to \infty$ и предельным равенством (3.7). Следовательно, используя определение ν_1, находим, что

$$tr\, G_{\nu_1}^{-1}\nu_1 \le \nu_1/c(H) < tr\, G_{\nu_1-1}^{-1}\nu_1.$$

Отсюда получаем, что

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\lim_{H\to\infty} \frac{\gamma_1}{C(H)} = tr F^{-1} \qquad \text{п.н.,} \quad (3.21)$$

Из (3.20), (3.21) и леммы 2 следует справедливость первого утверждения теоремы 5.

Докажем второе утверждение. Поскольку, как было показано выше, для устойчивого процесса авторегрессии (3.2) выполняется (3.12), то последовательная оценка $a^*(H)$ совпадает асимптотически п.н. с оценкой МНК, вычисленной на первом интервале наблюдений $[\gamma_1, \tau_1]$, которая имеет следующий вид:

$$\tilde{a}_1 = \tilde{G}^{-1}_{\gamma_1,\tau_1} \sum_{k=\gamma_1}^{\tau_1} \alpha_1(k) X_k y_{k+1},$$

а ее нормированное уклонение

$$\sqrt{\tau_1}\,(\tilde{a}_1 - a_1) = \sqrt{\tau_1}\, \tilde{G}^{-1}_{\gamma_1,\tau_1} \sum_{k=\gamma_1}^{\tau_1} \alpha_1(k) X_k \xi_{k+1} =$$

$$= \mathcal{D}(H) - d(H),$$

где

$$\mathcal{D}(H) = \tau_1 \hat{G}^{-1}_{\gamma_1,\tau_1} \frac{1}{\sqrt{\tau_1}} \sum_{k=0}^{\tau_1} \alpha_1(k) X_k \xi_{k+1};$$

$$d(H) = \tau_1 \hat{G}^{-1}_{\gamma_1,\tau_1} \sqrt{\frac{\gamma_1}{\tau_1}}\, \frac{1}{\sqrt{\gamma_1}} \sum_{k=0}^{\gamma_1-1} X_k \xi_{k+1}.$$

Вектор $\mathcal{D}(H)$ является асимптотически при $H\to\infty$ нормальным с нулевым средним и ковариацией F^{-1} (доказательство этого факта содержится в [5]), а вектор $d(H)$ стремится к нулю по вероятности в силу (3.19). Теорема 5 доказана.

ЛИТЕРАТУРА

1. Борисов В.З., Конев В.В. Последовательное оценивание параметров дискретных процессов. - А и Т, 1977, № 10, с. 58-64.
2. Конев В.В., Пергаменщиков С.М. Последовательные планы идентификации динамических систем. - А и Т, 1981, № 7, с. 84-92.
3. Lai T.L., Wei C.Z. (1983) Asymptotic properties of general autoregressive models and strong consistency of least-square estimates of their parameters - J. Mul-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

tivariate Anal. 13, 1-23.

4. Андерсон Т. Статистический анализ временных рядов. - М :Мир, 1976. - 756 с.
5. Конев В.В., Пергаменщиков С.М. Об оценивании числа наблюдений при последовательной идентификации параметров динамических систем. - А и Т, 1984, № 12, с. 56-62.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

СТРУКТУРНО УСТОЙЧИВЫЕ МНОГОФАКТОРНЫЕ МОДЕЛИ

В.И.Котюков, В.А.Калинкин

Вопросы построения статистических моделей, устойчивых к малым изменениям анализируемой выборки, наличию в ней резко выделяющихся наблюдений, весьма актуальны. Однако внимание исследователей концентрируется пока лишь на параметрической робастности (устойчивости оценок параметров) для простейших моделей, обычно используемых при анализе однородных данных /I, 2, 3/. При исследовании же сложных объектов, которое обычно осуществляется в условиях априорной неопределенности относительно вида функции распределения $F(y,x)$, значений факторов $\{y, X=(X_1,\dots,X_n)\}$, их большого количества (n), неоднородности анализируемой выборки $V=\{(y_i, x_i=(x_{i1},\dots,x_{in}));\ i=1,m\}$, в последнее время широко используются параметризованные модели сложной структуры. Например, многофакторные кусочно-линейные регрессионные и дискриминантные модели, модели кластерного анализа, и другие /4/.

При построении таких моделей на основе анализа V необходимо обеспечить в первую очередь устойчивость их структурных параметров. Потеря структурной робастности, в отличие от параметрической неустойчивости, ведет не только к более значительному ухудшению чисто статистических свойств модели (снижению качества прогноза по ней), но и лишает её интерпретируемости специалистом, что в ряде случаев не менее существенно. Заметим, что вопросы структурной устойчивости возникают и при построении многих традиционных моделей. Например, при выборе наиболее информативных признаков X_j из n исходных для линейной регрессии.

Рассматриваемый здесь подход может быть классифицирован

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

как алгоритмический метод решения задачи структурной устойчивости.

I. Определение структурной устойчивости

При изложении метода будем апеллировать к вопросам построения кусочно-линейных моделей (КЛМ) как одним из наиболее популярных параметризованных моделей сложной структуры в прикладных задачах анализа многофакторных объектов /4/. Регрессионная КЛМ k -го порядка имеет вид

$$y^* = g(X;\theta) = \sum_{\nu=1}^{k} \psi(X;\beta_\nu)\, g(X;\alpha_\nu) ,$$

где $\theta = \{k, \beta_\nu, \alpha_\nu\}$; $\psi(X;\beta_\nu)$ - параметризованная характеристическая (индикаторная) функция области $O_\nu \subset R^n$ аргумента X ;

$$\bigcup_{\nu=1}^{k} O_\nu = R^n \quad ; \quad O_\nu \cap O_q = \phi \quad \text{при } \nu \neq q ;$$

$$g(X;\alpha_\nu) = \sum_{j=1}^{n} \alpha_{\nu j} X_j + \alpha_{\nu 0} .$$

Функции $\psi(\cdot)$ обычно строятся либо с помощью предварительного использования процедур кластерного анализа объектов в R^n , либо путем реализации той или иной схемы выбора разделяющих гиперплоскостей из их исходного списка.

Две модели $g(X;\theta_1)$ и $g(X;\theta_2)$, построенные одним и тем же алгоритмом A на основе анализа соответственно выборок V_1 и V_2 , где $\theta_e = (\theta_{e1}, \ldots, \theta_{eq})$ при $e = 1,2$, будем называть с т р у к т у р н о э к в и в а л е н т н ы м и , если $\theta_1 \simeq \theta_2$, т.е. $q = q_1 = q_2$ и $\theta_{1i} \simeq \theta_{2i}$ для $\forall i$. Здесь $\simeq$ - знак "структурной" неразличимости соответствующих параметров.

Например, в задаче кластерного анализа два различных варианта таксономии будут структурно эквивалентными, если центры соответствующих таксонов метрически близки в R^n . Две частные линейные модели $g(X;\alpha_{1\nu})$ и $g(X;\alpha_{2\nu})$ эквивалентны если все соответствующие компоненты векторов $\alpha_{1\nu}$ и $\alpha_{2\nu}$ отличаются друг от друга незначительно.

Содержательная и количественная интерпретация отношения $\simeq$ априорно задается исследователем.

При исследовании устойчивости оценок параметров часто используется модель "засоренного" распределения

$$F_\varepsilon(y,x) = (1-\varepsilon) F(y,x) + \varepsilon F^*(y,x),$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

где $F^*(y,x)$ некоторое засоряющее распределение, а ε - его малая вероятностная доля /1, 2, 3/. Эта модель будет использована и нами, но в алгоритмическом аспекте.

Будем наряду с исходной выборкой V объема m, полученной согласно $F(y,x)$, рассматривать и выборку V^* объема m^*, полученную согласно $F^*(y,x)$. В качестве $F^*(y,x)$ целесообразно брать ту функцию распределения, по отношению к засоряющему эффекту которой должна быть устойчива модель $g(X;\theta)$. Пусть

$$V_\varepsilon = V \cup V^*; \quad m_\varepsilon = m + m^*; \quad \varepsilon = m^*/m_\varepsilon.$$

Модель $g(X;\theta)$ будем называть структурно ε-устойчивой, если для $g(X;\theta_\varepsilon)$ имеет место $\theta \simeq \theta_\varepsilon$.

Если имеется несколько выборок $\{V_\varepsilon\}$, то требуется выполнение $\theta \simeq \theta_\varepsilon$ для $\forall \theta_\varepsilon$.

Данное определение будет использовано в третьем параграфе при конструировании алгоритмов, обеспечивающих для моделей сложных объектов свойство структурной устойчивости.

2. Простые устойчивые модели

Здесь будут рассмотрены робастные алгоритмы линейной регрессии и кластерного анализа как базовые при построении многих сложных структурно устойчивых моделей $g(X;\theta)$ (например, КЛМ).

Для построения устойчивых линейных регрессионных моделей в настоящее время предложены различные схемы (λ-регрессия, SAD-регрессия, и другие) /1, 2, 3/. Нам же представляется целесообразным нижеследующее обобщение обычной медианной регрессии /5/, позволяющее в рамках единой оптимизационной постановки учитывать как требования устойчивости (взвешенный медианный критерий), так и интерпретируемости модели (априорные ограничения на знаки коэффициентов, на область прогнозируемых значений).

Для модели $y^* = \sum_{j=1}^{n} \alpha_j x_j + \alpha_0$ будем минимизировать критерий

$$\delta = \sum_{i=1}^{m} \omega_i |y_i - y_i^*|,$$

где ω - априорно заданный вектор весов наблюдений.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Пусть $(y_i - y_i^*) = (\beta_i^+ - \beta_i^-)$. Тогда необходимо

$$\sum_{i=1}^{m} \omega_i (\beta_i^+ + \beta_i^-) \Rightarrow \min_{\alpha,\beta}$$

при ограничениях

$$\sum_{j=0}^{n} \alpha_j \tilde{x}_{ij} + (\beta_i^+ - \beta_i^-) = y_i ; \quad i = \overline{1,m} ;$$

$$\beta_i^+ \geq 0 ; \quad \beta_i^- \geq 0 ; \quad i = \overline{1,m} ;$$

$$\alpha_j \geq 0 ; \quad j = \overline{1,n} ;$$

$$y_{i\,min} \leq \sum_{j=0}^{n} \alpha_j \tilde{x}_{ij} \leq y_{i\,max} ; \quad i = \overline{1,m} .$$

Здесь $\tilde{x}_{i0} = 1$ при $i = \overline{1,m}$; при априорном требовании $\alpha_j \geq 0$ имеем $\tilde{x}_{ij} = x_{ij}$, а при $\alpha_j \leq 0$ — $\tilde{x}_{ij} = -x_{ij}$. Это задача линейного программирования.

Необходимо заметить, что учёт априорных ограничений $\{\alpha_j \geq 0, \alpha_e \leq 0, y_{i\,min}, y_{i\,max}\}$ сам по себе повышает статистическую устойчивость модели. Последняя может быть еще повышена за счет итерационной корректировки весов ω.

Для задач кластерного анализа также целесообразно использовать "медианную" идеологию. Будем исходную выборку V разбивать на K непересекающихся групп объектов таким образом, чтобы минимизировать некоторую оценку среднего внутригруппового различия наблюдений в R^n.

Нормализуем исходные данные: $\tilde{x}_{ij} = x_{ij} / s_j$, где $s_j = \underset{i}{med}\{|x_{ij} - \underset{i}{med}\{x_i\}|\}$. Здесь $med\{\cdot\}$ — медианный оператор. Пусть $\{\omega_i\}$ — априорные веса реализаций, а $\{\tilde{\omega}_j\}$ — признаков.

Введем в R^n медианную метрику, для которой справедливы известные три аксиомы метрического пространства

$$\rho(\tilde{x}_i, \tilde{x}_e) = \underset{j}{med}\{\tilde{\omega}_j |\tilde{x}_{ij} - \tilde{x}_{ej}|\} .$$

Допустим, что V разбито некоторым образом на K таксонов $\{V_1, \ldots, V_K\}$: $\cup V_\nu = V$; $V_\nu \cap V_\mu = \emptyset$ при $\nu \neq \mu$. Тогда качество такого разбиения будем определять согласно

$$T = \underset{i}{med}\{\omega_i \sum_{\nu=1}^{K} \psi_\nu(\tilde{x}_i) \cdot \rho(\tilde{x}_i, \tilde{x}_\nu^\mu)\} ,$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

где

$$\psi_\nu(\tilde{x}_i)=\{1 \text{ при } \tilde{x}_i\in V_\nu;\ 0 \text{ при } \tilde{x}_i\notin V_\nu\};$$

$$\tilde{x}_\nu^{\mu}=(\tilde{x}_{\nu 1}^{\mu},\ldots,\tilde{x}_{\nu n}^{\mu});\quad \tilde{x}_{\nu j}^{\mu}=\underset{\tilde{x}_i\in V_\nu}{med}\{\tilde{x}_{ij}\}.$$

Для целей минимизации T по разбиению $\{V_\nu\}$ модифицируем известный итерационный алгоритм "K - средних" /4/:

а) задается некоторое начальное разбиение $\{V_\nu\}$;

б) для каждого V_ν определяется $\tilde{x}_\nu^{\mu}$;

в) m исходных объектов перераспределяются заново по K группам согласно принципу

$$\tilde{x}_i\in V_\nu,\ \text{если } \rho(\tilde{x}_i,\tilde{x}_\nu^{\mu})=\min_{q=1,K}\{\rho(\tilde{x}_i,\tilde{x}_q^{\mu})\};$$

г) если реализации $\{\tilde{x}_i\}$ иначе разбились на группы, то пункты б) и в) повторяются; в противном случае работа алгоритма закончена.

Процесс сходится к локальному минимуму T за конечное число шагов.

3. Построение структурно устойчивых моделей

Основная особенность предлагаемых структурно устойчивых алгоритмов заключается в том, что они одновременно анализируют как выборку V, так и V_ε (см. § I). Предполагается, что алгоритм A есть некоторая многошаговая процедура построения модели $g(X;\Theta)$. Таковыми являются многие методы построения КЛМ, выбирающие разделяющие границы для областей $\{O_\nu\}$ в заданном исходном списке возможных границ, методы выбора информативных признаков X_j из n исходных, рассмотренная выше процедура кластерного анализа, и другие.

Допустим на t-м шаге работы алгоритма A по анализу V и одновременно V_ε получена модель $g(X;\Theta^t)$, являющаяся структурно ε-устойчивой, т.е. $\Theta^t\simeq\Theta_\varepsilon^t$. Тогда на основе $\{g(X;\Theta^t),g(X;\Theta_\varepsilon^t)\}$ алгоритмом A делается $(t+1)$-й шаг, в результате которого получаются $\{g(X;\Theta^{t+1}),\ g(X;\Theta_\varepsilon^{t+1})\}$. Если $\Theta^{t+1}\simeq\Theta_\varepsilon^{t+1}$, то работа алгоритма продолжается. В противном случае за окончательное решение принимается $g(X;\Theta^t)$. Если алгоритм A на каждом t-м шаге сохраняет несколько решений, то оконча-

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

тельно на t-м шаге запоминаются лишь те из них, которые структурно ε-устойчивы.

Литература

I. Устойчивые статистические методы оценки данных/ Под ред. Р.Л.Лонера, Г.Н.Уилкинсона. - М.: Машиностроение, 1984. - 329 с.

2. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания. - М.: Финансы и статистика, 1980. - 208 с.

3. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Исследование зависимостей. - М.: Финансы и статистика, 1985. - 487 с.

4. Котюков В.И. Многофакторные кусочно-линейные модели. - М.: Финансы и статистика, 1984. - 216 с.

5. Киселев Н.И. Линейное программирование в задачах регрессии. - Заводская лаборатория, 1981, № 5, с. 33-37.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

О СРЕДНЕМ И ДИСПЕРСИИ ФУНКЦИЙ ОТ СТАТИСТИК

Г.М.Кошкин

Известна теорема о среднем и дисперсии некоторой функции от выборочных моментов /I, с.388/; в /2/ приведена без доказательства теорема I, содержащая аналогичный результат для более широкого класса двумерных статистик (t_n, m_n), которые могут, в частности, являться оценкой параметра (θ_1, θ_2). В данной работе приведено с доказательством обобщение теоремы I из /2/ на случай многомерной векторной статистики $(t_{1n}, \dots, t_{sn})$, которая может быть оценкой векторного параметра $(\theta_1, \dots, \theta_s)$.

Введём следующие обозначения: $x_1, \dots, x_n$, $x_i \in R^1$, $i = \overline{1,n}$, - случайная повторная выборка, которой соответствует выборочное вероятностное пространство $(R^n, \mathcal{B}^n, P)$ /3/; $t_n = (t_{1n}, \dots, t_{sn})$ - векторная статистика с компонентами $t_{jn} = t_j(x_1, \dots, x_n)$, $j = \overline{1,s}$, а $\tau_n = (\tau_{1n}, \dots, \tau_{sn})$, $\tau_{jn} = \mathsf{E}\, t_{jn}$ - её математическое ожидание; $\|B(t_n)\|$ - ковариационная матрица $\begin{bmatrix} cov(t_{1n}, t_{1n}) & \cdots & cov(t_{1n}, t_{sn}) \\ \cdots & \cdots & \cdots \\ cov(t_{sn}, t_{1n}) & \cdots & cov(t_{sn}, t_{sn}) \end{bmatrix}$;

$M_\nu(t_{jn}) = \mathsf{E}(t_{jn} - \tau_{jn})^\nu$ - центральный момент порядка ν случайной величины t_{jn}; $d(n)$, $q(n)$ - положительные числовые последовательности, неограниченно возрастающие с ростом n; C_i, $i = 0,1,2,\dots$ - неотрицательные постоянные; $\int = \int_{R^1}$.

ТЕОРЕМА I. Пусть: $A_1)$ $M_{2i}(t_{jn}) = O(1/d^i(n))$, $i = 2,3,\dots$; $j = \overline{1,s}$; $A_2)$ $\lim\limits_{n\to\infty} d(n)\|B(t_n)\| = \begin{bmatrix} b_{11} & \cdots & b_{1s} \\ \cdots & \cdots & \cdots \\ b_{s1} & \cdots & b_{ss} \end{bmatrix}$,

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$0 < b_{jj} < \infty$, $j = \overline{1,s}$; A_3) вещественная функция $H(z): R^s \to R^1$ в некоторой окрестности точки τ_n (τ_{jn}, $j = \overline{1,s}$, - ограничены) непрерывна и имеет непрерывные первые и вторые производные по аргументам z_j, причем $\lim\limits_{n\to\infty} \left.\frac{\partial H(z)}{\partial z_j}\right|_{\tau_n} = H_j$, $j = \overline{1,s}$; A_4) для всевозможных значений $x_1, \dots, x_n$ функции $|H(t_n)|, |t_{jn}|$, $j = \overline{1,s}$, меньше величины $C_0 q(n)$, $q(n) = d^{\gamma}(n)$, $\gamma \geq 0$.

Тогда математическое ожидание и дисперсия случайной величины $H(t_n)$ выражаются формулами:

$$EH(t_n) = H(\tau_n) + O\left(\frac{1}{d(n)}\right); \tag{1}$$

$$DH(t_n) = \frac{1}{d(n)} \sum_{j,p=1}^{s} H_j H_p b_{jp} + o\left(\frac{1}{d(n)}\right). \tag{2}$$

Доказательство. Следуя Крамеру /1/ (см. с.388-390), обозначим через $\bar{Z}$ множество всех точек в R^n, для которого выполняются s неравенств $|t_{jn} - \tau_{jn}| < \varepsilon$, $\varepsilon > 0$, $j = \overline{1,s}$, а через $\bar{Z}^*$ - дополнительное множество. По теореме Чебышева

$$P(\bar{Z}^*) < \frac{C_1}{\varepsilon^{2\kappa} d^{\kappa}(n)}, \quad P(\bar{Z}) > 1 - \frac{C_1}{\varepsilon^{2\kappa} d^{\kappa}(n)}. \tag{3}$$

Привлекая первое неравенство (3) и условие A_4), выбираем такое $\kappa_1 > 1$, чтобы $q(n)/d^{\kappa_1}(n) = o(1/d(n))$, и имеем

$$EH(t_n) = \int_{\bar{Z}} H(t_n)\,dP + \int_{\bar{Z}^*} H(t_n)\,dP = \int_{\bar{Z}} H(t_n)\,dP + o\left(\frac{1}{d(n)}\right).$$

При достаточно малом ε для каждой точки множества $\bar{Z}$ справедливо

$$\int_{\bar{Z}} H(t_n)\,dP = H(\tau_n)P(\bar{Z}) + \sum_{j=1}^{s} (H_j + o(1)) \int_{\bar{Z}} (t_{jn} - \tau_{jn})\,dP + \int_{\bar{Z}} R\,dP, \tag{4}$$

где $R = \frac{1}{2} \sum_{j,p=1}^{s} H'_{jp} (t_{jn} - \tau_{jn})(t_{pn} - \tau_{pn})$, H'_{jp} - значения вторых производных $\frac{\partial^2 H(z)}{\partial z_j \partial z_p}$ в некоторой промежуточной точке между t_n и τ_n.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Рассмотрим слагаемые в правой части равенства (4). Так как $1 > P(\bar{Z}) > 1 - \frac{C_1}{\varepsilon^{2\kappa} d^{\kappa}(n)}$, то

$$H(\tau_n)P(\bar{Z}) = H(\tau_n) + O(1/d^{\kappa}(n)).$$

В силу неравенства Шварца, условия A_2) и первого неравенства (3)

$$\int_{\bar{Z}}(t_{jn} - \tau_{jn})\,dP = \int_{R^n}(t_{jn} - \tau_{jn})\,dP - \int_{Z^*}(t_{jn} - \tau_{jn})\,dP =$$

$$= -\int_{Z^*}(t_{jn} - \tau_{jn})\,dP \le \sqrt{Dt_n \cdot P(Z^*)} = O\left(\frac{1}{d^{\frac{\kappa_1+1}{2}}(n)}\right).$$

Производные H'_{jp} при достаточно малом ε ограничены ввиду условия A_3), поэтому, учитывая A_2), получаем, что $\int_{\bar{Z}} R\,dP = O(1/d(n))$. Соотношение (I) доказано.

Докажем теперь соотношение (2). Принимая во внимание (I), имеем

$$DH(t_n) = \int_{R^n}\left(H(t_n) - EH(t_n)\right)^2 dP =$$

$$= \int_{R^n}\left(H(t_n) - H(\tau_n)\right)^2 dP + O\left(\frac{1}{d^2(n)}\right) = I_1 + I_2 + o\left(\frac{1}{d(n)}\right),$$

где

$$I_1 = \int_{\bar{Z}}\left(H(t_n) - H(\tau_n)\right)^2 dP, \qquad I_2 = \int_{Z^*}\left(H(t_n) - H(\tau_n)\right)^2 dP.$$

Аналогично, как при доказательстве (I),

$$I_2 \le 2\left\{\frac{C_0^2 q^2(n) C_1}{\varepsilon^{2\kappa} d^{\kappa}(n)} + \frac{H^2(\tau_n) C_1}{\varepsilon^{2\kappa} d^{\kappa}(n)}\right\} = O\left(\frac{q^2(n)}{d^{\kappa}(n)}\right). \qquad (5)$$

Выбирая такое $\kappa_2 > 1$, чтобы $q^2(n)/d^{\kappa_2}(n) = O(1/d(n))$, получаем

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$DH(t_n) = I_1 + o\left(\frac{1}{d(n)}\right).$$

С учётом разложения функции $H(t_n)$ в окрестности точки τ_n представим I_1 следующим образом:

$$I_1 = \int_{\bar{z}} \left[\sum_{i=1}^{s} (H_i + o(1))(t_{in} - \tau_{in}) + R\right]^2 dP. \qquad (6)$$

Очевидно, что

$$\int_{\bar{z}} (t_{jn} - \tau_{jn})(t_{pn} - \tau_{pn})\,dP = cov(t_{jn}, t_{pn}) - \int_{\bar{z}^*} (t_{jn} - \tau_{jn})(t_{pn} - \tau_{pn})\,dP. \qquad (7)$$

В силу неравенства Шварца по аналогии с (5)

$$\left|\int_{\bar{z}^*} (t_{jn} - \tau_{jn})(t_{pn} - \tau_{pn})\,dP\right| \le$$

$$\le \left[\int_{\bar{z}^*} (t_{jn} - \tau_{jn})^2 dP \int_{\bar{z}^*} (t_{pn} - \tau_{pn})^2 dP\right]^{\frac{1}{2}} = O\left(\frac{q^2(n)}{d^{\tau_2}(n)}\right). \qquad (8)$$

Теперь из (7), (8) и A_2) следует

$$\int_{\bar{z}} (t_{jn} - \tau_{jn})(t_{pn} - \tau_{pn})\,dP = \frac{b_{jp}}{d(n)} + o\left(\frac{1}{d(n)}\right). \qquad (9)$$

Привлекая неравенство

$$\left(\sum_{i=1}^{n} a_i\right)^2 \le n \sum_{i=1}^{n} a_i^2 \qquad (10)$$

и неравенство Шварца, получаем

$$ER^2 \le \frac{s^2}{4} \sum_{j,p=1}^{s} (H'_{jp})^2 E\left\{(t_{jn} - \tau_{jn})^2 (t_{pn} - \tau_{pn})^2\right\} \le$$

$$\le \frac{s^2}{4} \sum_{j,p=1}^{s} (H'_{jp})^2 \left[E(t_{jn} - \tau_{jn})^4 E(t_{pn} - \tau_{pn})^4\right]^{\frac{1}{2}} = O\left(\frac{1}{d^2(n)}\right); \qquad (11)$$

$$\left|\int_{\bar{z}} (t_{jn} - \tau_{jn}) R\,dP\right| \le \sqrt{Dt_n \cdot ER^2} = O\left(\frac{1}{d^{3/2}(n)}\right). \qquad (12)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Из (6), (9), (II), (I2) немедленно следует (2). Теорема I доказана.

Замечание. Если в условии A_4) теоремы I дополнительно $\lim_{n\to\infty} q^2(n)/d(n) = 0$, то ослабляется условие A_1) (достаточно ограничиться случаем $i = 2$).

Обозначим $H_j(\tau_n) = \left.\frac{\partial H(z)}{\partial z_j}\right|_{\tau_n}$, $H_j = \lim_{n\to\infty} H_j(\tau_n)$.

Докажем аналог теоремы I при несколько других условиях.

ТЕОРЕМА 2. Пусть: $Б_1$) $M_4(t_{jn}) = O\left(\frac{1}{d^2(n)}\right)$;

$Б_2$) $\lim_{n\to\infty} d(n)\|B(t_n)\| = \begin{bmatrix} b_{11} \cdots b_{1s} \\ \cdots \cdots \cdots \\ b_{s1} \cdots b_{ss} \end{bmatrix}$, $0 < b_{ii} < \infty$, $i = \overline{1,s}$;

$Б_3$) на R^s вещественная функция $H(z)$: $R^s \to R^1$ дифференцируема и удовлетворяет условиям

$|H_j(z_1, \ldots, z_j', \ldots, z_s) - H_j(z_1, \ldots, z_j'', \ldots, z_s)| < C_j|z_j' - z_j''|$,

$j = \overline{1,s}$; $Б_4$) $\lim_{n\to\infty} \tau_n = \tau$, $\tau_{jn} < \infty$, $\tau_j < \infty$, $j = \overline{1,s}$.

Тогда справедливы утверждения теоремы I.

Доказательство. Запишем разложение на R^s

$$H(t_n) = H(\tau_n) + \sum_{j=1}^{s} H_j(\tau_n)(t_{jn} - \tau_{jn}) + Q, \qquad (I3)$$

где

$$Q = \sum_{j=1}^{s} \left[H_j(\tau_{1n}, \ldots, \tau_{jn} + \theta_j(t_{jn} - \tau_{jn}), \ldots, \tau_{sn}) - H_j(\tau_n)\right](t_{jn} - \tau_{jn}), \quad 0 < \theta_j < 1, \quad j = \overline{1,s}.$$

Усредняя левую и правую часть равенства (I3) с учётом условий $Б_3$) и $Б_2$), приходим к (I).

Из (I) и (I3) в силу $Б_2$), $Б_3$) и $Б_4$) имеем

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$DH(t_n) = E\left\{\sum_{j=1}^{s} H_j(\tau_n)(t_{jn}-\tau_{jn}) + Q + O\left(\frac{1}{d(n)}\right)\right\}^2 =$$

$$= \frac{1}{d(n)} \sum_{i,j=1}^{s} H_i H_j b_{ij} + o\left(\frac{1}{d(n)}\right) + E\left[Q + O\left(\frac{1}{d(n)}\right)\right]^2 +$$

$$+ 2E\left\{\left[\sum_{j=1}^{s} H_j(\tau_n)(t_{jn}-\tau_{jn})\right]\left[Q + O\left(\frac{1}{d(n)}\right)\right]\right\}. \qquad (14)$$

Но согласно $Б_3$) и $Б_2$)

$$E|Q| \le \sum_{j=1}^{s} C_j \theta_j E(t_{jn}-\tau_{jn})^2 = O\left(\frac{1}{d(n)}\right),$$

согласно $Б_1$) и (10)

$$EQ^2 \le E\left(\sum_{j=1}^{s} C_j \theta_j (t_{jn}-\tau_{jn})^2\right)^2 \le$$

$$\le sC_j^2\theta_j^2 \sum_{j=1}^{s} E(t_{jn}-\tau_{jn})^4 = O\left(\frac{1}{d^2(n)}\right),$$

следовательно,

$$E\left[Q + O\left(\frac{1}{d(n)}\right)\right]^2 = O\left(\frac{1}{d^2(n)}\right). \qquad (15)$$

Далее, воспользовавшись неравенством для абсолютных моментов /4, с.190/ и неравенством Шварца, имеем

$$\left|E\left[\sum_{j=1}^{s} H_j(\tau_n)(t_{jn}-\tau_{jn})\cdot Q\right]\right| \le \sum_{j,k=1}^{s} |H_j(\tau_n)| C_k \theta_k \cdot$$

$$\cdot E\left|(t_{jn}-\tau_{jn})(t_{kn}-\tau_{kn})^2\right| \le \sum_{j=1}^{s} |H_j(\tau_n)| \theta_j C_j \left(E(t_{jn}-\tau_{jn})^4\right)^{\frac{3}{4}} +$$

$$+ \sum_{j\ne k}^{s} |H_j(\tau_n)| \theta_j C_k \left[E(t_{jn}-\tau_{jn})^2 E(t_{kn}-\tau_{kn})^4\right]^{\frac{1}{2}} =$$

$$= O(1/d^{3/2}(n)). \qquad (16)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Теперь из соотношений (I4)-(I6) следует (2). Теорема 2 доказана.

Литература

I. Крамер Г. Математические методы статистики. - 2-е изд. - М.: Мир, I975.- 648 с.

2. Кошкин Г.М. Об одном методе устранения смещения оценок. - Теория вероятн. и ее примен., I987, т. XXXII, вып. I, с. I47-I49.

3. Боровков А.А. Математическая статистика. Оценка параметров, проверка гипотез. - М.: Наука, I984.- 472 с.

4. Гнеденко Б.В. Курс теории вероятностей. - 5-е изд. - М.: Наука, I969.- 400 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

СРЕДНЕКВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ НЕПАРАМЕТРИЧЕСКИХ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК С ВРЕМЕННЫМ СДВИГОМ

И.Л.Кожевникова

Асимптотическое изучение статистических характеристик спектров многомерных временных рядов [I] имеет большое практическое значение [2] . Введение непараметрических оценок взаимных спектров с временным сдвигом [I] дает возможность вычислять спектральные оценки для выборок большой длины, разбивая их на отдельные отрезки, возможно пересекающиеся [3] . За счет осреднения оценок по отрезкам, на которые разбита вся выборка, достигается уменьшение дисперсии оценки. Именно эта процедура позволила успешно применить оценки спектральной плотности с временным сдвигом в одномерном случае к исследованию датчиков псевдослучайных чисел [4] . Следует отметить, что такие оценки являются робастными относительно возмущений в соседних частотах [3] .

Для выборки из многомерной стационарной гауссовской последовательности $X(t)=\{X_1(t),\dots,X_n(t)\}$, $t=0,\pm1,\dots,N-1$, $N=(T-1)L+M+1$, рассматривались оценки амплитудного спектра, фазового спектра, квадрата модуля спектра когерентности [I] , определяемые соответственно следующими выражениями:

$$\bar{A}_{xy}(\lambda)=\left[\bar{L}^2_{xy}(\lambda)+\bar{Q}^2_{xy}(\lambda)\right]^{1/2};$$

$$\bar{F}_{xy}(\lambda)=\operatorname{arctg}\left(-\frac{\bar{Q}_{xy}(\lambda)}{\bar{L}_{xy}(\lambda)}\right);$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$|\bar{K}_{ij}(\lambda)|^2 = \frac{\bar{L}_{ij}^2(\lambda) + \bar{Q}_{ij}^2(\lambda)}{\bar{f}_{ii}(\lambda)\bar{f}_{jj}(\lambda)}, \quad i \neq j, \; i,j = 1, \ldots, n. \tag{1}$$

Функции L, T, M являются целочисленными функциями объема выборки N и удовлетворяют условиям

$$\frac{M}{N} \to 0, \quad \frac{L}{N} \to 0, \quad \frac{LT}{N} \to 1, \tag{2}$$

если $T \to \infty$, $L \to \infty$, $M \to \infty$ при $N \to \infty$.

Последовательность $X(t)$ принадлежит классу $\mathcal{H}(\lambda, \alpha_{ij}, i,j = 1, \ldots, n)$ [1, 3]. Оценка взаимной спектральной плотности [1] имеет вид

$$\bar{f}_{ij}(\lambda) = \frac{1}{T} \sum_{k=0}^{T} W_{Mi}^{Lk}(\lambda) \overline{W_{Mj}^{Lk}(\lambda)}, \quad i,j = 1, \ldots, n, \tag{3}$$

$$W_{Mi}^{Q}(\lambda) = \sum_{t=-\infty}^{\infty} a_M(t-Q) X_i(t) e^{i\lambda t}.$$

Функция $a_M(t)$, называемая окном данных, – неотрицательная функция, $t = 0, \pm 1, \ldots$, равная 0 вне отрезка $[0, M-1]$. Обозначим

$$\varphi_M^Q(x) = \sum_{t=-\infty}^{\infty} a_M(t-Q) e^{itx}, \quad \varphi_M^Q(x) = \varphi_M(x) e^{iQx} \tag{4}$$

Если $\sum_{t=-\infty}^{\infty} a_M^2(t) = 1/2\pi$, то $\int_{-\pi}^{\pi} |\varphi_M(x)|^2 = 1$.

Функции $\bar{L}_{ij}(\lambda) = \mathrm{Re}\, \bar{f}_{ij}(\lambda)$ и $\bar{Q}_{ij}(\lambda) = \mathrm{Im}\, \bar{f}_{ij}(\lambda)$ – есть оценки коспектра и квадратурного спектра [1]. Следуя [1], будем обозначать для удобства каждую из функций (2) через $g_i(Y_1, \ldots, Y_k)$, $i = 1, 2, 3$, в которых $Y_1 = \bar{L}_{ij}(\lambda)$, $Y_2 = \bar{Q}_{ij}(\lambda)$, $Y_3 = \bar{f}_{ii}(\lambda)$, $Y_4 = \bar{f}_{jj}(\lambda)$, причем каждая из функций $g_i(Y_1, \ldots, Y_k)$ является нелинейной функцией аргументов Y_1, Y_2, Y_3 и Y_4. Рассмотрим разложение Тейлора $g_i(Y_1, \ldots, Y_k)$ в окрестности точки $\{EY_1, \ldots, EY_k\}$:

$$g_i(Y_1, \ldots, Y_k) = g_i(EY_1, \ldots, EY_k) + \tag{5}$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$+\sum_{j=1}^{k}\left.\frac{\partial g_i}{\partial y_j}\right|_{EY_1,\dots,EY_k}(y_j-EY_j)+R,\quad i=1,2,3.$$

Линейным приближением оценки $g_i(Y_1,\dots,Y_k)$ будем называть выражение вида

$$g_i^{(1)}(Y_1,\dots,Y_k)=g_i(EY_1,\dots,EY_k)+$$

$$+\sum_{j=1}^{k}\left.\frac{\partial g_i}{\partial y_j}\right|_{EY_1,\dots,EY_k}(y_j-EY_j).$$

Остаточный член (5) определяется

$$R=\frac{1}{2}\sum_{l,m=1}^{k}\left.\frac{\partial^2 g_i}{\partial y_l\,\partial y_m}\right|_{\theta_1,\dots,\theta_k}(y_l-EY_l)(y_m-EY_m),\qquad(6)$$

для которого частные производные вычисляются в точке $\{EY_1+\theta_1E(Y_1-EY_1),\dots,EY_k+\theta_kE(Y_k-EY_k)\}$, $0<\theta_m<1$, $m=1,\dots,k$.

Сформулируем лемму [I], которая понадобится при доказательстве теоремы.

ЛЕММА. Пусть $X(t)=\{X_1(t),\dots,X_n(t)\}$, $t=0,\pm1,\dots$, – многомерная стационарная последовательность, удовлетворяющая условиям: $EX_r(t)=0$, $|EX_r(t)X_j(t+u)|<c$, $0<c<\infty$, $|f_{rj}(\lambda)|<c_1$, $0<c_1<\infty$, $r,j=1,\dots,n$.

Тогда четвертные семиинварианты $s_4(\xi_\nu-E\xi_\nu)$, $\nu=1,2$, $\xi_1=\mathrm{Re}\,Tf_{rj}(\lambda)$, $\xi_2=\mathrm{Im}\,Tf_{rj}(\lambda)$ при $N\to\infty$ имеют порядок

$$s_4(\xi_\nu-E\xi_\nu)=O(T\max|\varphi_m(x)|^6),\quad \nu=1,2,$$

где функции $f_{rj}^{x}(\lambda)$ и $|\varphi_m(x)|^2$ определены формулами (3) и (4), а функции k, τ, M удовлетворяют условиям (2).

Существенное ослабление рассматриваемых ранее условий (см. теорему 3.5 [I]) было предложено И.Г.Журбенко и реализовано в приведенной ниже теореме.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

ТЕОРЕМА. Пусть $X(t)=\{X_1(t),\dots,X_n(t)\}\in\mathcal{X}(\lambda,\alpha_{rj},\ r,j=1,\dots,n),\ t=0,\pm1,\dots,$ - многомерная стационарная гауссовская последовательность. Пусть, кроме того, для последовательности весовых функций $\Phi_N(x)=|\varphi_M(x)|^2,\ M=M(N)$, при $N\to\infty$ соблюдаются условия

$$\max_x|\varphi_M(x)|^2=O\left(\int_{-\pi}^{\pi}|\varphi_M(x)|^4dx\right); \tag{7}$$

$$L^2\int_{-\pi}^{\pi}|\varphi_M(x)|^4dx=o(1), \tag{8}$$

а функции L,T,M удовлетворяют условиям (2). Тогда среднеквадратическое отклонение нелинейных оценок взаимных спектров от их линейных приближений при $N\to\infty$ имеет порядок

$$E\left(g_i(y_1,\dots,y_k)-g_i^{(\lambda)}(y_1,\dots,y_k)\right)^2=O\left(\int_{-\pi}^{\pi}|\varphi_M(x)|^4dx/N\right)^2,$$

если

$$\left(\int_{-\pi}^{\pi}|\varphi_M(x)|^4dx\right)^2/N\to0. \tag{9}$$

Доказательство. Для доказательства теоремы необходимо оценить порядок четвертого момента остаточного члена (6). Семиинвариант четвертого порядка $S_4(\xi),\ E\xi=0,$ выражается [5] следующим образом:

$$S_4(\xi)=m_4(\xi)-3m_2^2(\xi).$$

Из леммы следует, что порядок $S_4((\xi_\nu-E\xi_\nu)/T),\ \nu=1,2,$ определяется величиной $(\max_x|\varphi_M(x)|^6/T^3)$.

Так как функция $|\varphi_M(x)|^2>0$, то $\max_x|\varphi_M(x)|^6=(\max_x|\varphi_M(x)|^2)^3$. Отсюда и из (7) следует, что порядок $S_4((\xi_\nu-E\xi_\nu)/T),\ \nu=1,2,$ равен

$$\frac{1}{T^3}\left(\max_x|\varphi_M(x)|^6\right)=\frac{1}{T^3}O\left(\int_{-\pi}^{\pi}|\varphi_M(x)|^4dx\right)^3=$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$= \frac{L^3}{N^3} O\left(\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right)^3, \qquad (10)$$

так как $N \approx L T$. Поэтому [I]:

$$E\left(\bar{L}_{xy}(\lambda) - E\bar{L}_{xy}(\lambda)\right)^2 = O\left(\frac{1}{N}\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right). \qquad (11)$$

Из выражений (8), (10), и (11) следует, что порядок $S_4(\xi_v - E\xi_v), v = 1,2$, меньше порядка квадрата второго момента оценки коспектра, т.к

$$\frac{\frac{L^3}{N^3} O\left(\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right)^3}{O\left(\frac{1}{N}\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right)^2} = \frac{L^2}{T} O\left(\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right) = o(1).$$

Таким образом, четвертый момент оценки коспектра, входящий в четвертый момент остаточного члена, характеризующего отклонение нелинейной оценки взаимного спектра от ее линейного приближения, равен

$$O\left(\frac{1}{N}\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right)^2. \qquad (12)$$

Четвертый момент остаточного члена (6) зависит и от квадратов вторых моментов оценок $\bar{Q}_{xy}(\lambda)$, $\bar{f}_{zz}(\lambda)$ и $\bar{f}_{yy}(\lambda)$, а также смешанных моментов четвертого порядка. Что касается членов, содержащих четвертые моменты $\bar{Q}_{xy}(\lambda)$, $\bar{f}_{zz}(\lambda)$ и $\bar{f}_{yy}(\lambda)$, они оцениваются аналогичным образом и имеют порядок (12). Смешанные четвертые моменты упомянутых оценок оцениваются с помощью неравенства Коши-Буняковского и также имеют порядок (12). Остается сравнить порядок

$$O\left(\frac{1}{N}\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right)^2$$

с порядком наибольшего остаточного члена дисперсии линейного приближения нелинейной оценки взаимного спектра, который равен [I]:

$$O\left(\frac{\left(\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right)^{1-\beta_{xc}}}{N}\right). \qquad (13)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Сравнивая (I2) и (I3), получим, что при $N \to \infty$

$$O\left(\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx / N\right)^2 : O\left(\left(\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right)^{1-\beta_{K^2}} / N\right) =$$

$$= O\left(\left(\int_{-\pi}^{\pi} |\varphi_M(x)|^4 dx\right)^{1+\beta_{K^2}} / N\right) = o(1),$$

так как выполняется (9) и $0 < \beta_{K^2} < 1$.

Теорема доказана.

Предположение о гауссовости принято для упрощения доказательства, и результат может быть обобщен на широкие классы негауссовских последовательностей.

Литература

I. Журбенко И.Г., Кожевникова И.А., Оценки взаимных спектров многомерных гауссовских последовательностей.-Известия АН СССР, серия математическая, I985, **вып.6, с.I229-I245.**

2. Отнес Р., Эноксон Л., Прикладной анализ временных рядов.- -М.: Мир, I982,- 430 с.

3. Журбенко И.Г., Спектральный анализ временных рядов.-М.:Изд-во МГУ, I982,- I68 с.

4. Журбенко И.Г., Кожевникова И.А., Клиндухова О.В., Выявление периодических составляющих в последовательности псевдослучайных чисел.-Кибернетика, I984, **№4, с.89-96.**

5. Ширяев А.Н., Вероятность.-М.: Наука, I980,- 576 с.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

МЕТОДЫ ОПТИМАЛЬНОЙ И АДАПТИВНОЙ ФИЛЬТРАЦИИ В СЛУЧАЕ ПОМЕХ, СОДЕРЖАЩИХ ВЫБРОСЫ

Э.А.Клёкис

При решении реальных задач оценивания сигналов часто встречается ситуация, когда небольшая доля измерений содержит грубые ошибки или выбросы [I]. Традиционные методы фильтрации (фильтр Калмана) являются чувствительными к резковыделяющимся наблюдениям, появляется проблема обеспечения устойчивости фильтров к отклонениям свойств наблюдений от предполагаемых. Возникает задача робастной фильтрации.

Применение минимаксного подхода [2] к решению задачи робастной фильтрации сталкивается с трудностями синтеза фильтра для наименее благоприятного распределения помех из заданного класса. С использованием этого подхода возможно получение только приближенных фильтров [I, 3].

В настоящей работе предлагается оптимальный фильтр для оценивания фазовых координат динамической системы, когда плотность вероятности помехи наблюдения является смесью двух гауссовских плотностей. Для случая, когда отсутствует априорная информация о вероятностных характеристиках выбросов, приведен алгоритм фильтрации, основанный на обнаружении выбросов и коррекции коэффициента усиления фильтра. Представлены результаты сравнения эффективности фильтров методом статистического моделирования, коротко упоминается о применении к решению реальных задач.

Рассмотрим нестационарную динамическую систему с дискретным временем, описываемую уравнениями:

$$x_{k+1} = \Phi_k x_k + \xi_k; \quad z_k = H_k x_k + \varrho_k, \tag{I}$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$z_k = (1-\gamma_k)\zeta_k + \gamma_k \chi_k \quad (k = 1,2,\dots), \qquad (2)$$

где $x_k \in R^n$ – вектор фазовых координат; $z_k \in R^m$ – вектор наблюдений; Φ_k, H_k – известные матрицы; $\{\xi_k\}, \{\zeta_k\}, \{\chi_k\}$ – последовательности некоррелированных гауссовских случайных величин $\xi_k \in R^n$, $\zeta_k \in R^m$, $\chi_k \in R^m$, имеющих нулевые средние значения и ковариационные матрицы соответственно $\{Q_k\}, \{R_k^{(0)}\}, \{R_k^{(1)}\}$, причем $\|R_k^{(1)}\| \gg \|R_k^{(0)}\|$; $\{\gamma_k\}$ – последовательность независимых, одинаково распределенных случайных величин, принимающих 2 значения $\gamma_k \in \{0,1\}$; x_0 – гауссовская случайная величина со средним значением $\bar{x}_0$ и ковариационной матрицей P_0. Последовательности $\{\xi_k\}$, $\{\zeta_k\}, \{\chi_k\}$ и x_0 взаимонезависимы, а также задана вероятность $P(\gamma_k = 1) = \varepsilon \ll 1$. Таким образом, в нормальном режиме работы погрешности измерительной системы моделируют последовательность $\{\zeta_k\}$, а $\{\chi_k\}$ является моделью редких выбросов. Задача состоит в последовательном оценивании x_k по наблюдениям $Z_1^k \equiv \{z_1, \dots, z_k\}$ при $k = 1,2,\dots$.

Основой для получения оптимальной оценки фазовых координат x_k является апостериорная плотность вероятности. Введем обозначения: $w_k(x_k) \equiv f(x_k | Z_1^k)$ – апостериорная плотность вероятности;

$$w_k^{(j)}(x_k) \equiv f(x_k | \gamma_k = j, Z_1^k), \quad w_k^{(ij)}(x_k) \equiv f(x_k | \gamma_k = j, \gamma_{k-1} = i, Z_1^k)$$

– условные апостериорные плотности вероятности; $W_k^{(j)} \equiv P(\gamma_k = j | Z_1^k)$ – апостериорная вероятность значения j случайной величины γ_k; $N[x|m,C]$ – гауссовская плотность вероятности случайной величины x, определяемая математическим ожиданием m и ковариационной матрицей C.

ТЕОРЕМА. Пусть выполнены вышесделанные предположения о динамической системе и возмущениях. Тогда апостериорная плотность вероятности $w_k(x_k)$ определяется выражениями:

$$w_k(x_k) = \sum_{j=0}^{1} W_k^{(j)} w_k^{(j)}(x_k), \quad w_k^{(j)}(x_k) = \frac{\sum_{i=0}^{1} \psi_{k-1}^{(ij)} W_{k-1}^{(i)} w_k^{(ij)}(x_k)}{\sum_{\alpha=0}^{1} \psi_{k-1}^{(ij)} W_{k-1}^{(\alpha)}}; \qquad (3)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$w_k^{(ij)}(x_k) = \left(\psi_{k-1}^{(ij)}\right)^{-1} N\left[z_k | H_k x_k, R_k^{(j)}\right] \int N\left[x_k | \Phi_{k-1} x_{k-1}, Q_{k-1}\right] w_{k-1}^{(i)}(x_{k-1}) dx_{k-1}; \quad (4)$$

$$\psi_{k-1}^{(ij)} = \int N\left[z_k | H_k x_k, R_k^{(j)}\right] \int N\left[x_k | \Phi_{k-1}, Q_{k-1}\right] w_{k-1}^{(i)}(x_{k-1}) dx_{k-1} dx_k; \quad (5)$$

$$W_k^{(0)} = \frac{(1-\varepsilon)\sum_{i=0}^{1} \psi_{k-1}^{(i0)} W_{k-1}^{(i)}}{\sum_{\beta=0}^{1} (-1)^{\beta} (1-\varepsilon-\beta) \sum_{\alpha=0}^{1} \psi_{k-1}^{(\alpha\beta)} W_{k-1}^{(\alpha)}}, \quad W_k^{(1)} = 1 - W_k^{(0)}; \quad (6)$$

$$\hat{x}_0^{(j)} = \bar{x}_0, \quad p_0^{(j)} = p_0, \quad W_0^{(0)} = 1-\varepsilon \quad (i,j = 0,1). \quad (7)$$

Апостериорные плотности вероятности $w_k^{(j)}(x_k)$ не имеют конечномерного параметрического представления, поэтому точная реализация вычислений согласно (3)–(7) невозможна. Представим алгоритм фильтрации в гауссовском приближении. Обозначим $x_k \equiv M(x_k | Z_1^k)$, $\hat{x}_k^{(j)} \equiv M(x_k | \vartheta_k = j, Z_1^k)$, $\hat{x}_k^{(ij)} \equiv M(x_k | \vartheta_k = j, \vartheta_{k-1} = i, Z_1^k)$ – соответственно безусловное, условное по $\vartheta_k = j$ и условное по $\vartheta_k = j$, $\vartheta_{k-1} = i$ апостериорное математическое ожидание x_k. Им соответствуют обозначения для ковариационных матриц: p_k, $p_k^{(j)}$, $p_k^{(ij)}$. Тогда алгоритм фильтрации имеет вид

$$\hat{x}_k = \sum_{j=0}^{1} W_k^{(j)} \hat{x}_k^{(j)}, \quad p_k = \sum_{j=0}^{1} W_k^{(j)} \left\{ p_k^{(j)} + \left[\hat{x}_k^{(j)} - \hat{x}_k\right]\left[\hat{x}_k^{(j)} - \hat{x}_k\right]^T \right\}; \quad (8)$$

$$\hat{x}_k^{(j)} = \sum_{i=0}^{1} \varkappa_k^{(ij)} \hat{x}_k^{(ij)}, \quad p_k^{(j)} = \sum_{i=0}^{1} \varkappa_k^{(ij)} \left\{ p_k^{(ij)} + \left[\hat{x}_k^{(ij)} - \hat{x}_k^{(j)}\right]\left[\hat{x}_k^{(ij)} - \hat{x}_k^{(j)}\right]^T \right\}; \quad (9)$$

$$\varkappa_k^{(ij)} = \frac{N\left[\nu_k^{(i)} | 0, V_k^{(ij)}\right] W_{k-1}^{(i)}}{\sum_{\alpha=0}^{1} N\left[\nu_k^{(\alpha)} | 0, V_k^{(\alpha j)}\right] W_{k-1}^{(\alpha)}}, \quad W_k^{(i0)} = \frac{(1-\varepsilon)\sum_{i=0}^{1} N\left[\nu_k^{(i)} | 0, V_k^{(i0)}\right] W_{k-1}^{(i)}}{\sum_{\beta=0}^{1} (-1)^{\beta} (1-\varepsilon-\beta) \sum_{\alpha=0}^{1} N\left[\nu_k^{(\alpha)} | 0, V_k^{(\alpha\beta)}\right] W_{k-1}^{(\alpha)}}; \quad (10)$$

$$\hat{x}_k^{(ij)} = \Phi_{k-1} \hat{x}_{k-1}^{(i)} + K_k^{(ij)} \nu_k^{(i)}, \quad p_k^{(ij)} = (I - K_k^{(ij)} H_k) \bar{p}_k^{(i)}; \quad (11)$$

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

$$\nu_k^{(i)} = z_k - H_k \Phi_{k-1} \hat{x}_{k-1}^{(i)}, \quad V_k^{(ij)} = H_k \bar{P}_k^{(i)} H_k^T + R_k^{(j)}; \tag{I2}$$

$$\bar{P}_k^{(i)} = \Phi_{k-1} P_{k-1}^{(i)} \Phi_{k-1}^T + Q_{k-1}, \quad K_k^{(ij)} = \bar{P}_k^{(i)} H_k^T (V_k^{(ij)})^{-1} \quad (i,j = 0,1). \tag{I3}$$

Для реализации этого алгоритма необходимо синтезировать четыре фильтра калмановского типа, взвешенная сумма выходных сигналов которых определяет искомую оценку.

Далее предположим, что отсутствует априорная информация об источнике выбросов (неизвестна вероятность ε и ковариационные матрицы $\{R_k^{(j)}\}$). В этом случае построить фильтр (8)–(I3) уже невозможно. Зададим другую модель помех наблюдения:

$$z_k = \xi_k + \gamma_k B \Delta_k, \tag{I4}$$

где $\Delta_k \in R^s$ – неизвестный вектор, представляющий амплитуду выброса; $B \in R^{m \times s}$ – матрица, состоящая из нулей и единиц и определяющая компоненты вектора z_k, искажаемые выбросами. Для модели помехи вида (I4) адаптивный алгоритм фильтрации, основанный на обнаружении выбросов (оценивании γ_k) и изменении структуры фильтра в зависимости от результатов обнаружения, определяется выражениями:

$$\hat{x}_k = \Phi_{k-1} \hat{x}_{k-1} + K_k [I - \hat{\gamma}_k B L_k] \nu_k, \quad \nu_k = z_k - H_k \Phi_{k-1} \hat{x}_{k-1}; \tag{I5}$$

$$P_k = (I - K_k H_k) \bar{P}_k + \hat{\gamma}_k K_k B D_k B^T K_k^T, \qquad K_k = \bar{P}_k H_k^T V_k^{-1}; \tag{I6}$$

$$\bar{P}_k = \Phi_{k-1} P_{k-1} \Phi_{k-1}^T + Q_{k-1}, \quad V_k = H_k \bar{P}_k H_k^T + R_k^{(0)}; \tag{I7}$$

$$D_k = (B^T V_k^{-1} B)^{-1}, \qquad L_k = D_k B^T V_k^{-1}; \tag{I8}$$

$$\hat{\gamma}_k = \begin{cases} 0, & \ell(\nu_k) \le h \\ 1 & \ell(\nu_k) > h \end{cases}, \quad \ell(\nu_k) = \nu_k^T V_k^{-1} B D_k B^T V_k^{-1} \nu_k. \tag{I9}$$

В зависимости от результатов обнаружения выброса меняется вес

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

соответствующих компонент наблюдений при формировании оценки (I5), а также выражение для ковариационной матрицы P_k, значение которой в случае принятия решения о наличии выброса ($\hat{\gamma}_k = 1$) увеличивается. Решающее правило для обнаружения выброса основано на вычислении статистики обобщенного отношения правдоподобия $\ell(\nu_k)$ и сравнении ее с порогом h.

Фильтр (I5)-(I9) является более простым по сравнению с (8)-(I3) и не требует априорной информации о характеристиках выбросов, однако он имеет тот недостаток, что его структура жестко зависит от результатов обнаружения выброса и, следовательно, в случае ложных тревог появляются дополнительные погрешности оценивания. В то же время фильтр (8)-(I3) формирует оценку $\hat{x}_k$ как взвешенную сумму оценок, основанных на различных гипотезах о наблюдениях, и поэтому в нем более гибко учитываются текущие характеристики наблюдений.

При статистическом моделировании одномерные наблюдения генерировались согласно моделям (I), (2). Обозначим σ_α^2, $\alpha = \overline{0,3}$ выборочные среднеквадратические ошибки фильтрации усредненные по I000 реализаций в момент $k = 20$ для следующих фильтров: σ_0^2 – оптимального фильтра при точно известной $\{\gamma_k\}$; σ_1^2 – фильтра (8)-(I3); σ_2^2 – фильтра (I5)-(I9); σ_3^2 – оптимального линейного фильтра, синтезированного при дисперсии $R = (1-\varepsilon)R^{(0)} + \varepsilon R^{(1)}$. Значения показателя $\varrho_\alpha = \sigma_0^2/\sigma_\alpha^2$, $\alpha = \overline{1,3}$ при $\phi = 0{,}9$, $R^{(0)} = I$, $\beta = 0{,}I$, $h = I6$ приведены в таблице.

ε	$R^{(1)}$	ϱ_1	ϱ_2	ϱ_3
0,2	9	0,92I	0,804	0,760
	I00	0,897	0,873	0,499

Фильтр вида (I5)-(I9) использован при оценивании состояния электроэнергетической системы.

Литература

I. Ершов А.А. Стабильные методы оценки параметров (обзор). – Автоматика и телемеханика, I978, № 7, с. 68-73.
2. Хьюбер П. Робастность в статистике. – М.: Мир, I984. – 304 с.
3. Masreliez C.J. Approximate non-Gaussian filtering with linear state and observation relations. – IEEE Trans. Automat. Contr. 1975, vol. 20, N 1, p. 107-110.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

МАТЕРИАЛЫ

VI Всесоюзной школы-семинара по
непараметрическим и робастным методам
статистики в кибернетике

Часть I

Подписано к печати 3/04-87 КЗ 06138 Формат 60x84 1/16
Бумага типографская №3. Печать офсетная. Печ. л. 14,81 ;
усл. печ. л. 13,82; уч.-изд. л. 12,5. Тираж 200экз. Заказ 111
Цена 1р. 90к.

Ротапринт ТГУ, 634029, Томск, ул. Никитина, 4.

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

Digital Library (repository)
of Tomsk State University
http://vital.lib.tsu.ru

1-660936