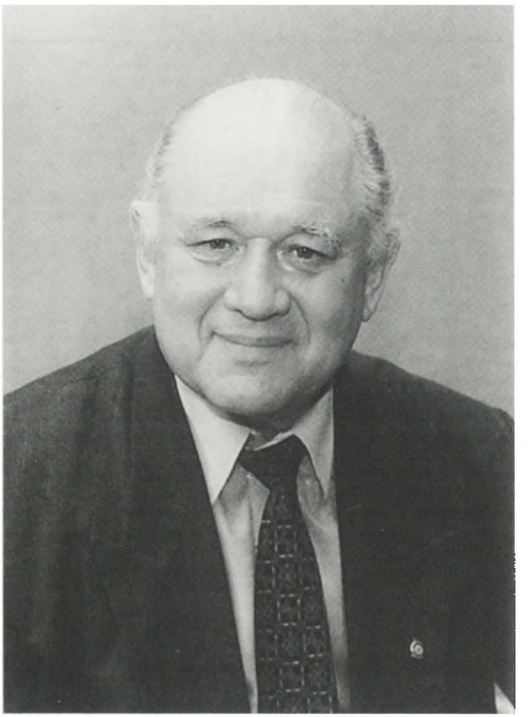
ТРУДЫ

X ЮБИЛЕЙНОГО СИМПОЗИУМА

ПО НЕПАРАМЕТРИЧЕСКИМ И РОБАСТНЫМ СТАТИСТИЧЕСКИМ МЕТОДАМ В КИБЕРНЕТИКЕ



Профессор Томского госуниверситета
Феликс Петрович Тарасенко

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ТРУДЫ

X ЮБИЛЕЙНОГО СИМПОЗИУМА

ПО НЕПАРАМЕТРИЧЕСКИМ И РОБАСТНЫМ СТАТИСТИЧЕСКИМ МЕТОДАМ В КИБЕРНЕТИКЕ

ТОМСК - 2004

УДК 519.62
ББК 32.973
Т 782

Т 782 **Труды X юбилейного симпозиума по непараметрическим и робастным статистическим методам в кибернетике.** — Томск: Изд-во НТЛ, 2004. — 224 с.

Издание посвящено 70-летию заслуженного деятеля науки и техники РФ, профессора Томского государственного университета, бессменному руководителю оргкомитета проводимых симпозиумов – Феликсу Петровичу Тарасенко. В труды симпозиума включены работы, написанные на основе отобранных оргкомитетом докладов, посвященных задачам математической статистики, теории вероятностей и их применению в решении проблем финансовой и актуарной математики, автоматического управления, теории информации, распознавания образов и других задач кибернетики.

Для научных сотрудников, а также студентов старших курсов соответствующих специальностей.

УДК 519.62
ББК 32.973

© Томский госуниверситет, ФПМК, 2004

Содержание

Слово юбиляра 5

I. Непараметрические методы

Васильев В.А. Об оценивании распределения шумов управляемого процесса авторегрессии 8

Воробейчиков С.Э., Кабанова Т.В. Об обнаружении момента изменения среднего случайной последовательности 15

Дмитриев Ю.Г., Зенкова Ж.Н. Об оценивании симметричного распределения по цензурированной выборке 22

Дмитриев Ю.Г., Тарасенко Ф.П. Информационные подходы к непараметрическому оцениванию функционалов от распределений 30

Koshkin G.M. and Tarasenko F.P. Continuous-Discrete Stochastic Objects: Nonparametric Identification and Control 34

Красноштанов А.П. Непараметрическая идентификация многосвязных систем при неполной информации 48

Лемешко Б.Ю., Постовалов С.Н. Рекомендации ГОСТ Р 50.1-033-2001 и ГОСТ Р 50.1-037-2002 как итог исследований вопросов применения критериев согласия 55

Медведев Г.А. Непараметрические методы оценивания нейтральных к риску вероятностей по данным рыночных наблюдений 67

Неделько В.М. К задаче оценивания качества решающей функции 78

Пестунов И.А. О некоторых аспектах реализации классификатора Розенблатта–Парзена в условиях нестационарности и малых выборок 86

Тарасенко П.Ф. Проверка гипотез для индикаторного анализа линейных моделей 93

II. Робастные методы

Домбровский В.В., Ляшенко Е.А. Робастное управление дискретными системами со случайными параметрами и мультипликативными шумами 109

Лемешко Б.Ю., Помадин С.С. Корреляционный анализ многомерных случайных величин при нарушении предположений о нормальности 114

Медведев А.Г. Робастная версия GMM-оценки 128

Смагин В.И., Иванова Е.В. Синтез робастных регуляторов для систем со случайными скачкообразными параметрами 140
Харин Ю.С., Акинфин А.В. Робастность в планировании регрессионных экспериментов при наличии функциональных искажений 145
Шуленин В.П. Робастные статистические процедуры 159
III. Статистические приложения
Демин Н.С., Рожкова С.В., Рожкова О.В. Оценивание и распознавание в стохастических системах при непрерывно-дискретных наблюдениях со скользящей памятью 173
Dmitriev Yu.G., Tarasenko F.P., Ustinov Yu.K. On Statistical Structure of Growing Systems 179
Домбровский В.В., Гальперин В.А. Управление инвестиционным портфелем в непрерывном времени при квадратической функции риска 185
Перемитина Т.О., Полищук Ю.М. Комплексный анализ многомерных данных на основе метода главных компонент с применением ГИС 192
Поддубный В.В. Выделение из таблиц сводных данных таблиц сопряженности признаков для их трехмерного графического отображения и обработки в пакете STATISTICA 197
Сущенко С.П., Сущенко М.С. О быстродействии многоуровневой памяти неблокирующего типа 209
Шестернева О.В. Об использовании уравнения Винера–Хопфа для идентификации линейных динамических систем в условиях пассивного эксперимента 215

СЛОВО ЮБИЛЯРА

Многоуважаемые читатели!

Перед вами сборник научных статей, посвященный мне в связи с моим 70-летием.

Невозможно придумать более приятный подарок ко дню рождения человека, вовлеченного в научную и педагогическую деятельность. Один мой коллега, профессор М. Хупер (M. Hooper) из Глазго, написал мне, что хотя ему нечего представить в этот сборник в связи с его специфически обуженной тематической направленностью, сама идея его посвящения моему юбилею чрезвычайно привлекательна. И я с ним согласен. Выражаю глубокую признательность моим друзьям, коллегам, сотрудникам, по инициативе и трудами которых появилось данное издание. Особая благодарность профессору Ю.Г. Дмитриеву, заведующему кафедрой теоретической кибернетики Томского государственного университета, организовавшему все это мероприятие.

Но все-таки основное значение и смысл данного сборника не в его связанности с моим юбилеем, а в том, что он продолжает многолетнюю (более 30 лет) традицию сибирских статистиков. С 1972 г., а затем регулярно, через каждые 2-3 года, проводится школа-семинар (конференция, симпозиум) по непараметрическим и робастным статистическим методам и их приложениям. Долгое время это мероприятие называлось "Всесоюзная школа-семинар по ...", так как по тогдашним правилам только школе-семинару дозволялось длиться более трех дней. Мы проводили наши встречи в загородных домах отдыха, и это позволяло не только вести продуктивные научные дискуссии, но и общаться в приятной неформальной обстановке, использовать разнообразные виды отдыха.

География представительства участников охватывала всю страну (поэтому после 1991 г. мы стали называться не Всесоюзной, а Международной конференцией), но проводили встречи неизменно в сибирских краях: Томске, Красноярске, Иркутске, Шушенском, Дивногорске, Железногорске... Оргкомитет постоянно возглавляли я и проф. А.В. Медведев из Красноярска. Наши встречи именовались школой-семинаром не только по упомянутой выше организационной причине. Хотя в них нередко принимали участие видные корифеи (Ю.И. Алимов, Б.Р. Левин, Г.А. Медведев, Ю.И. Неймарк, Р.З. Хасьминский, Я.З. Цыпкин и др.), все это было адресовано научной молодежи. Теперь уже многие участники школы сами стали солидными учеными, заработали международный авторитет.

По известным причинам в последние годы стало сложнее собираться в прежнем составе, и интервалы между встречами удлинились. Надеюсь, вскоре состоится наша следующая очная встреча, а пока мне кажется удачной идея поддержания традиции путем проведения очередной конференции в виртуальной форме. Результатом воплощения этой идеи и является настоящий сборник; надеюсь, читатели оценят его высокий научный уровень.

В расчете на то, что сборник будут читать многие, хочу сказать несколько слов о его содержании. Сравнительно недавно возникшие и продолжающие бурно развиваться разделы статистики – непараметрические и робастные методы – являются ответом на потребности пользователей статистической обработки данных. Дело в том, что требования к точности и надежности выводов возрастают не только в "точных" естественных науках, но и в традиционно "качественных", гуманитарных. Это резко расширяет число пользователей методов математической статистики. При этом круг возникающих задач расширяется настолько, что он выходит за пределы множества задач классической статистики. Дело в том, что любая статистическая процедура основана на определенных предположениях о природе обрабатываемых данных. Большинство классических методов построено в предположении об известности распределения обрабатываемых наблюдений (обычно – нормального), их статистической независимости и т. д. Во многих практических задачах данные не отвечают таким предположениям и использование алгоритмов, основанных на неверных посылках, дает выводы гораздо худшего качества, чем ожидалось. В попытках исправить такое положение и возникли непараметрические и робастные методы.

Непараметрическая статистика принципиально отказывается от известности функции распределения, ограничиваясь доступными сведениями о ее непрерывности или дискретности, симметричности, обладании какими-то моментами и т.п. Удивительно, что иногда непараметрические процедуры мало отличаются от параметрических по качеству решений.

Робастная ветвь статистики несколько более консервативна: в ней считается, что истинное распределение находится в окрестности некоторого известного, и синтезируются алгоритмы, менее чувствительные к отклонениям истинного распределения от предполагаемого, нежели параметрические, – ценой некоторого снижения качества на точном распределении и ограничения окрестности с гарантированным качеством.

Эти ветви анализа данных и были неизменным центром внимания сибирских специалистов в прикладной статистике, а три десятилетия

школы-семинара являются проявлением признания нашей принадлежности к числу лидеров в этой области. Публикация данного сборника пусть косвенно, но подтверждает это еще раз.

В заключение чересчур, быть может, пространного предисловия, и пользуясь привилегией юбиляра, скажу несколько слов о себе. Грех мне жаловаться на пролетевшие годы. Я уже давно перевалил за среднюю продолжительность жизни россиянина, а статистика еще утверждает, что чем дольше человек живет, тем больше вероятность ему дожить до ста лет. Так что каждому следующему году можно только радоваться. И в прошедшие годы мне, в основном, везло. (Иногда я даже думаю, что в именах есть некий судьбоносный смысл: Феликс по латыни значит "счастливый", – стучу по дереву.) Мне повезло учиться и работать в Томском госуниверситете – одном из лучших университетов мира. Я оказался в числе основателей и лидеров томской кибернетической школы. Научная и педагогическая деятельность позволила мне побывать в пятнадцати странах (с особым удовольствием вспоминаю почти двухлетнее преподавание в Дар-эс-Салаамском университете – в качестве эксперта ЮНЕСКО. Могу похвастать восхождением на Килиманджаро). Почти полторы сотни научных публикаций; несколько книг; более трех десятков кандидатов наук, некоторые из них стали докторами и добились бóльших научных успехов, чем я. Три сына и дочь, семь внуков, любимая и любящая жена; добрые друзья; дом, построенный в деревне собственноручно. Замечательный дружный коллектив, в котором я продолжаю работать. Что еще нужно для счастья?!

Еще раз сердечно благодарю всех, принявших участие в создании данного сборника и согласившихся посвятить его мне.

Ф. Тарасенко

I. Непараметрические методы

В.А. Васильев

Об оценивании распределения шумов управляемого процесса авторегрессии

Решается задача оценивания статистических характеристик распределения шумов управляемого процесса авторегрессии первого порядка с неизвестным параметром динамики. Рассматриваются квадратичные критерии качества управления, оптимизирующие скорость минимизации дисперсии объекта. Результаты могут быть использованы, в частности, при построении законов управления для более сильных критериев качества.

1. Введение

В последнее время достигнут значительный прогресс в построении регуляторов с обратной связью для стохастических систем с неизвестными параметрами. Процедуры адаптивного управления предполагают использование различных методов оценивания неизвестных параметров объекта. Обычно используются такие методы, как метод наименьших квадратов (МНК), метод стохастической аппроксимации (МСА), метод максимального правдоподобия (ММП) и др. Наиболее удобным с практической точки зрения является МНК, поскольку он не требует априорной информации о параметрах системы и распределении ее шумов. В то же время МСА и ММП могут давать оценки с более сильными статистическими свойствами, но при этом требуют знания либо самого распределения помех объекта, либо функционалов от него.

Использование того или иного метода оценивания параметров наблюдаемого процесса тесно связано с выбором функционала качества при формировании закона адаптивного управления. Как правило, функционал качества строится на основе квадратичной метрики, вследствие чего решается задача минимизации дисперсии объекта. При этом в адаптивной постановке возможна только асимптотическая минимизация с различной скоростью сходимости. Наиболее сильный критерий качества, дающий оптимальную скорость сходимости дисперсии авторегрессионного объекта и оценок его параметров динамики, представлен в работе [1]. Закон адаптивного управления в [1] строился на основе оценок МСА параметров процесса, что потребовало знания дисперсии его

шумов и логарифмической производной их распределения. При этом видно, что одновременная адаптация управления к неизвестным параметрам и характеристикам шумов приводит к весьма сложной для аналитического исследования задаче. Поэтому вполне естественно предварительно решить задачу оценивания характеристик шумов объекта и использовать эти оценки в алгоритме оптимального управления [1]. Эта задача может быть решена только при использовании какого-то другого управления, не зависящего от неизвестных характеристик. При этом желательно, чтобы критерий качества не был бы сильно ослаблен и позволял решать по смыслу ту же задачу. Закон управления с такими свойствами был предложен в работе [2], где рассмотрены два различных критерия качества, которые так же, как и в [1], учитывают скорость сходимости дисперсии авторегрессионного объекта, однако в несколько более слабом смысле. В [2] показано, что построенный закон управления является оптимальным в широком классе управляющих воздействий и распределений шумов процесса. При построении закона управления использовались последовательные оценки наименьших квадратов параметра динамики объекта.

Оценки МНК неизвестных регрессионных параметров при построении закона управления использовались также в [3], где рассмотрен квадратический критерий, учитывающий скорость потраекторного сближения объекта и его шума.

Другие известные результаты в области адаптивного управления линейными стохастическими системами не учитывают, как правило, скорость сходимости основных характеристик объекта к своему оптимуму (см., например, [4] и библиографию к [3] и [4]).

В данной работе рассматривается задача непараметрического оценивания ряда статистических характеристик шумов процесса авторегрессии первого порядка с управлением и неизвестным параметром динамики. При этом предполагается, что управление строится согласно процедуре, предложенной в [2], что позволяет, с одной стороны, решать самостоятельную задачу адаптивного управления в смысле критериев работы [2] и в то же время дает принципиальную возможность оптимального управления объектом в смысле более сильного критерия работы [1].

Использование непараметрического подхода при решении статистических задач позволяет, как известно из [5], получать статистические выводы при низком уровне априорной информации о неизвестном распределении и решать реальные практические задачи.

Непараметрические оценки, рассматриваемые в данной работе, строятся по существенно зависимой выборке, но, тем не менее, сохраняют асимптотические свойства классических оценок.

2. Постановка задачи. Построение оптимального управления

Рассмотрим линейный случайный процесс, удовлетворяющий уравнению

$$x_n = \lambda x_{n-1} + u_n + \xi_n, \ n \geq 1, \tag{1}$$

где $x = \{x_n\}$, $u = \{u_n\}$ и $\xi = \{\xi_n\}$ - выходная, входная и шумовая последовательности соответственно; $x_0 = 0$; λ - неизвестный параметр, такой, что $|\lambda| \leq L$; $0 < L < \infty$ - известная постоянная. Шум ξ представляет собой процесс независимых одинаково распределенных случайных величин (н.о.р.с.в.) с неизвестной функцией распределения $F(\cdot)$ и неизвестной дисперсией σ^2.

Задача состоит в построении и исследовании асимптотических свойств непараметрических оценок распределения $F(\cdot)$ и его статистических характеристик.

В дальнейшем будет рассмотрена управляющая последовательность u, которая минимизирует критерии вида

$$J_1(u, \mathcal{P}) = \sup_{\mathcal{P}} \overline{\lim_{N\to\infty}} \frac{1}{\ln N} \sum_{n=1}^{N} M(x_n - \xi_n)^2; \tag{2}$$

$$J_2(u, \mathcal{P}) = \sup_{\mathcal{P}} \overline{\lim_{N\to\infty}} \frac{1}{N} \sum_{n=1}^{N} nM(x_n - \xi_n)^2. \tag{3}$$

Здесь $\mathcal{P}$ - класс функций распределения шумов $F(\cdot)$, для определения которого нам потребуются следующие предположения на шумы ξ :

A_1) ξ - последовательность н.о.р.с.в. с $E\xi_1 = 0$, $E\xi_1^2 = \sigma^2$, $E\xi_1^{2m} < \infty$ для некоторого $m \geq 1$;

A_2) ξ_1 имеет симметричную плотность $f(x) = f(-x)$, невозрастающую на $[0, +\infty)$;

A_3) $f(x)$ абсолютно непрерывна и

$$I(f) = \int_{-\infty}^{\infty} \frac{[f'(t)]^2}{f(t)} dt < \infty.$$

Рассмотрим следующие классы распределений шумов ξ:

– $\mathcal{P}_m$ – класс распределений шумов, удовлетворяющих A_1);
– $\mathcal{P}_\infty = \bigcap\limits_{m\geq 1} \mathcal{P}_m$;
– $\overline{\mathcal{P}}_m$ – класс распределений шумов, удовлетворяющих A_1), A_2);
– $\overline{\mathcal{P}}_\infty$ – подкласс распределений из $\mathcal{P}_\infty$, удовлетворяющих A_2), A_3).

Введем класс допустимых управляющих воздействий.

Определение 1. Будем говорить, что управляющая последовательность $u = \{u_n\}$ допустима, если

1) для всех $n = 1, 2, \ldots$ существует некоторый полином $p_n(z)$, такой, что $|u_n| \leq p_n(\|x^n\|)$, $x^n = (x_1, x_2, \ldots, x_n)$, $\|\cdot\|$ – евклидова норма;

2) $\lim\limits_{n\to\infty} Mx_n^2 = \sigma^2$.

Обозначим через U класс всех допустимых управляющих последовательностей. Принимая во внимание квадратический тип критериев (2), (3), естественно рассматривать линейное управление вида

$$u_n^* = -\lambda_{n-1}^* x_{n-1}, \tag{4}$$

где

$$\lambda_n^* = \mathrm{proj}_{[-L,L]}\lambda^*(k-1), \quad \tau(k-1) \leq n < \tau(k),\ k \geq 1, \tag{5}$$

$$\lambda^*(k) = \frac{\sum\limits_{n=1}^{\tau(k)} \alpha(n,k) x_{n-1}(x_n - u_{n-1}^*)}{\sum\limits_{n=1}^{\tau(k)} \alpha(n,k) x_{n-1}^2} =$$

$$= k^{-\varphi} \sum_{n=1}^{\tau(k)} \alpha(n,k) x_{n-1}(x_n - u_{n-1}^*),$$

$$\tau(k) = \inf\{t \geq 1 : \sum_{n=1}^{t} x_{n-1}^2 \geq k^\varphi\},\ \varphi > 0,$$

$$\alpha(n,k) = \begin{cases} 1, & n < \tau(k), \\ \alpha_k, & n = \tau(k), \end{cases} \quad \alpha_k = \left(k^\varphi - \sum_{j=1}^{\tau(k)-1} x_{j-1}^2\right) / x_{\tau(k)-1}^2.$$

Можно показать, что управление $u^* = \{u_n^*\}$ допустимо, то есть $u^* \in U$, если распределение $F(\cdot)$ помех входит в класс $\mathcal{P}_2$.

Теорема 1. *Рассмотрим систему (1) при $F \in \mathcal{P}_2$, и пусть закон управления u^* определен в (4), (5) при $\varphi > 4$. Тогда*

1) $\boldsymbol{J}_1(u^*, \mathcal{P}_2) = J_2(u^*, \mathcal{P}_2) = \sigma^2$;

2) *управление* u^* *оптимально в классе* U *в смысле критериев (2), (3) при* $F \in \overline{\mathcal{P}}_\infty$, *то есть*

$$J_k(u^*, \overline{\mathcal{P}}_\infty) = \inf_U J_k(u, \overline{\mathcal{P}}_\infty) = \sigma^2, \ k = 1, 2;$$

3) *при* $F \in \overline{\mathcal{P}}_m$ *для некоторого* $m \geq 2$ *последовательность оценок (5) параметра* λ *и управляемый процесс (1) удовлетворяют предельным соотношениям:*

$a)$ $\lim\limits_{n\to\infty} n^\gamma |\lambda_n - \lambda| = 0$ *a.s.*;

$b)$ $\lim\limits_{n\to\infty} n^\gamma |x_n - \xi_n| = 0$ *a.s.*,

где $\gamma = 0$, *если* $m = 2$ *и* $\gamma < 1/2 - 1/m$, *если* $m > 2$;

$c)$ $\sup\limits_{\overline{\mathcal{P}}_m} \sum\limits_{n=1}^{N} n^{s\gamma} M|x_n - \xi_n|^s = O(1), \quad N \to \infty,$

где $2 < s \leq 3$, *если* $m = 2$ *и* $s = m$, *если* $m > 2$.

Теорема 1 показывает, что управление u^* оптимизирует функционалы (2) и (3) как в широком классе распределений шумов объекта, так и в классе допустимых управлений. При этом найдены скорость сходимости оценок параметра динамики и скорость потраекторного сближения объекта управления и его шумового процесса. Равенство c) в третьем утверждении теоремы 1 устанавливает скорость сходимости также и для моментов отклонений реализации наблюдаемого процесса от помех

3. Непараметрическое оценивание распределения помех объекта

Из теоремы 1 видно, что построенный алгоритм управления обеспечивает асимптотические свойства процесса x, достаточные для решения целого класса задач по непараметрическому оцениванию распределений шумов ξ объекта x, а также основных характеристик этого распределения (см. [6–11]).

В качестве примера рассмотрим непараметрические оценки дисперсии σ^2 помех ξ процесса (1) и оценки логарифмической производной плотности $f(\cdot)$ их распределения. Эти характеристики необходимы для построения неулучшаемого по скорости сходимости алгоритма оптимального управления, предложенного Немировским и Цыпкиным [1].

Заметим, что вследствие асимптотической близости наблюдаемого процесса x и шумового процесса ξ при использовании управления u^*

величины x_n являются естественными оценками шумов ξ_n, и оценка дисперсии σ^2 может быть построена по наблюдениям x (см. [6]):

$$s_n = \frac{1}{n}\sum_{k=1}^{n} x_k^2,\ n \geq 1, \tag{6}$$

а непараметрическая оценка логарифмической производной $T(t) = \dfrac{f'(t)}{f(t)}$ плотности $f(t)$ распределения помех ξ имеет вид (см. [11])

$$\tilde{T}_n(t) = \frac{T_n(t)}{(1+\delta_n|T_n(t)|^q)^\rho}, \tag{7}$$

где $T_n(t) = \hat{f}'(t)/\hat{f}(t)$ – оценка подстановки, $\delta = \{\delta_n\}_{n\geq 1}$ – вспомогательная последовательность положительных чисел, постоянные ρ и q удовлетворяют соотношениям $\rho q \geq 1$, $\rho > 0$, $q > 0$. Здесь в качестве $\hat{f}'(t)$ и $\hat{f}(t)$ используются непараметрические оценки производной $f'(t)$ и плотности $f(t)$ вида

$$\hat{f}_n^{(\alpha)}(t) = \frac{1}{nh_n^{1+\alpha}}\sum_{k=1}^{n} K^{(\alpha)}\left(\frac{t-x_k}{h_n}\right),$$

где $K(z)$ – ядерная функция, $h = \{h_n\}_{n\geq 1}$ – некоторая последовательность положительных чисел, $\alpha = 0; 1$.

Как показано в работе [11], такая структура оценки отношения $T(t)$ позволяет избежать неограниченного роста оценки и ее характеристик вследствие возможной близости к нулю знаменателя.

Дадим вспомогательные определения.

Определение 2. Плотность распределения $f(t)$ принадлежит $\mathcal{N}$, если $\sup\limits_t f(t) \leq C$, функция $f(t)$ непрерывна вместе с производной порядка $\nu+1$ включительно на R^1, $\nu \geq 2$, причем производная порядка $\nu+1$ удовлетворяет условию Липшица с показателем $0 < \mu \leq 1$.

Обозначим $\mathcal{H}$ как множество монотонно убывающих последовательностей $\boldsymbol{h} = \{h_n\}_{n\geq 1}$ положительных чисел, таких, что

$$\lim_{n\to\infty}(h_n + (nh_n^3)^{-1}) = 0.$$

Определение 3. Для некоторого четного $\nu \geq 2$ ядро $K(z)$ принадлежит Σ_ν, если $K(z)$ – финитная, непрерывно дифференцируемая функция:

$$K(z) = K(-z),\ \int |z|^\nu\,|K(z)|\,dz < \infty,$$

$$\int K(z)dz = 1, \quad \int z^j K(z)dz = 0, \quad \int z^\nu K(z)dz \neq 0, \quad j = \overline{1, \nu - 1}.$$

Сформулируем теоремы, отражающие свойства оценок (6) и (7).

Теорема 2. *Пусть процесс авторегрессии* $\{x_n\}$, *задаваемый уравнением (1), устойчив,* $Mx_0^8 < \infty$, *а шумы* $\{\xi_n\}$ *имеют распределение* $F(\cdot) \in \overline{\mathcal{P}}_4$ *и образуют последовательность независимых одинаково распределенных случайных величин. Тогда оценки* s_n *дисперсии* σ^2, *заданные формулой (6), при* $n \to \infty$ *обладают свойством равномерной сходимости в среднеквадратическом смысле:*

$$\sup_{\lambda \in [-L, L]} |nM_\lambda (s_n - \sigma^2)^2 - \sigma^2| = o(1).$$

Обозначим

$$D = f^{-1}(t) \int [K'(z)]^2 dz, \quad \omega = \frac{1}{\nu!} f^{-2}(t) f^{(\nu+1)}(t) \int z^\nu K(z)dz.$$

Теорема 3. *Пусть процесс авторегрессии* $\{x_n\}$, *задаваемый уравнением (1), устойчив,* $Mx_0^{8(\nu+1)} < \infty$, *а шумы* $\{\xi_n\}$ *имеют распределение* $F(\cdot) \in \overline{\mathcal{P}}_{4(\nu+1)}$ *и образуют последовательность независимых одинаково распределенных случайных величин. Пусть, кроме того, плотность* $f(t) \in \mathcal{N}$, $\nu \geq 2$, $h \in \mathcal{H}$ *и ядро* $K \in \Sigma_\nu$. *Тогда для оценки* $\tilde{T}_n$, *определенной в (7) при* $q > 4$ *и* $\delta_n = O(h_n^{2\nu} + (nh_n^3)^{-1})$, *главная часть ее среднеквадратического отклонения определяется равенством*

$$\sup_{\lambda \in [-L, L]} \left| M_\lambda [\tilde{T}_n(t) - T(t)]^2 - \frac{f(t)D}{nh_n^3} - \omega h_n^{2\nu} \right| = o(h_n^{2\nu} + (nh_n^3)^{-1}).$$

Проблемы оценивания характеристик распределения помех авторегрессии без управления рассматривались в работах [6] и [12]. При этом в [6] предложены оценки функции распределения, функционалов от него и производных плотности этого распределения с заданным среднеквадратическим качеством, а в работе [12] исследованы асимптотические оценки функционалов интегрального типа.

СПИСОК ЛИТЕРАТУРЫ

1. *Немировский А. С., Цыпкин Я. З.* Об оптимальных алгоритмах адаптивного управления // Автоматика и телемеханика. 1984. № 12. С. 64-77.
2. *Konev V. V., Vasil'iev V. V.* On optimal adaptive control of a linear stochastic process // Proc. 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics, August 24-29, 1997. Berlin, Germany. Vol. 5. P. 87-91.

3. *Guo L.* The logarithm law on self-tuning regulators // 12th World Congress International Federation of Automatic Control, 18-23 July 1993. Sydney, Australia, 1993. Preprints of Papers. Vol. 1. P. 227-232.
4. *Piunovskiy A.B.* Optimal control of stochastic sequences in problems with constraints. – Kluwer Academic Publishers, 1996.
5. *Тарасенко Ф. П.* Непараметрическая статистика. – Томск: Изд-во Том. ун-та, 1976.
6. *Васильев В. А.* Об идентификации динамических систем авторегрессионного типа // Автоматика и телемеханика. 1997. № 12. С. 107-119.
7. *Васильев В. А., Кошкин Г. М.* Оценивание предельной плотности распределения и ее производных по наблюдениям с ослабевающей зависимостью // Проблемы передачи информации. 1997. Т. 33. Вып. 2. С. 66-80.
8. *Васильев В. А., Кошкин Г. М.* Оценивание функций от плотности распределения по зависимым наблюдениям // Проблемы передачи информации. 1997. Т. 33. Вып. 4. С. 45-60.
9. *Васильев В. А., Кошкин Г. М.* Непараметрическая идентификация авторегрессий // Теория вероятностей и ее применения. 1998. Т. 43. Вып. 3. С. 577-588.
10. *Koshkin G. M., Vasil'iev V. A.* Nonparametric estimation of derivatives of a multivariate density from dependent observations // Mathematical Methods of Statistics. 1998. Vol. 7. No 4. P. 361-400.
11. *Васильев В. А., Кошкин Г. М.* Непараметрическое оценивание отношений производных многомерной плотности распределения по зависимым наблюдениям // Сибирский математический журнал. 2000. Т. 41. № 2. С. 284-303.
12. *Васильев В. А.* Об оценивании распределения возмущений процесса авторегрессии // Матем. статистика и ее приложения. – Томск: Изд-во Том. ун-та, 1986. Вып. 10. С. 9–24.

С.Э. Воробейчиков, Т.В. Кабанова

Об обнаружении момента изменения среднего случайной последовательности [1]

Рассматривается задача обнаружения момента изменения среднего в последовательности независимых случайных величин с неизвестной функцией распределения. Предлагается последовательная процедура, позволяющая обнаруживать как увеличение, так и уменьшение среднего. Находятся характеристики процедуры, и устанавливается ее оптимальность.

[1] Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант РФФИ-ДФГ є 02-01-04001

1. Введение

В прикладных задачах, связанных с обработкой экспериментальных данных, часто возникает необходимость обнаруживать моменты изменения свойств наблюдаемых процессов (моменты разладки). В случае, когда необходимо обнаружить момент разладки как можно быстрее, применяются последовательные алгоритмы. Достаточно хорошо изученной является задача обнаружения изменения функции распределения в последовательности независимых случайных величин в предположении, что известны начальная и конечная модели процесса [1–3]. Лорденом [3] было установлено, что оптимальной в классе последовательных процедур обнаружения (в смысле минимума среднего времени запаздывания при заданном среднем времени между ложными тревогами) является алгоритм кумулятивных сумм (АКС), предложенный Пейджем и представляющий собой многократно возобновляемую последовательную процедуру Вальда классификации двух простых гипотез.

В последнее время рассматриваются более сложные ситуации, когда распределение наблюдений до или(и) после момента разладки неизвестны. В [4, 5] установлены асимптотические соотношения, связывающие время запаздывания в обнаружении разладки и среднее время между ложными тревогами. В [6] была предложена последовательная непараметрическая процедура обнаружения увеличения среднего значения наблюдений, для которой найдены основные характеристики.

В данной работе на основе результатов [6] предлагается новая процедура, позволяющая обнаруживать как увеличение, так и уменьшение среднего в наблюдениях. Получены формулы для определения среднего времени между ложными тревогами и среднего времени запаздывания в обнаружении разладки. При построении процедуры используется метод обнаружения разладки в многоканальной системе, предложенный в [7].

2. Постановка задачи. Построение процедуры обнаружения и исследование ее свойств

Пусть наблюдается последовательность независимых случайных величин $\{X_i\}$ с неизвестной функцией распределения $F(x)$. В момент разладки θ происходит изменение среднего значения наблюдений на неизвестную величину a. Требуется по наблюдениям $\{X_i\}$ обнаружить момент разладки θ.

В [6] было предложено для обнаружения момента увеличения среднего в наблюдениях ($a > 0$) использовать АКС на основе величин $y_i =$

$\operatorname{sign}(x_i - x_{i-k})$, $k \geq 1$. Для любого распределения наблюдений x_i статистики y_i имеют нулевое среднее, если $i < \theta$ или $i - k \geq \theta$. Для $i \leq \theta < i + k$ среднее значение y_i положительно. Таким образом, среднее значение последовательности y_i отличается от нуля только на интервале длины k.

Основными характеристиками для последовательных процедур обнаружения являются среднее время между соседними моментами ложных тревог T_0 и среднее время запаздывания в обнаружении разладки T_1. При этом для оптимальных процедур величина T_1 растет логарифмически с ростом T_0. Для выполнения такого соотношения необходимо, чтобы среднее значение слагаемых в АКС до разладки было отрицательным, а после разладки – положительным.

Так как среднее значение y_i равно нулю до момента разладки, будем из каждого слагаемого вычитать некоторую положительную величину δ. Эта величина определяет чувствительность процедуры к изменению среднего. Уменьшение δ позволяет обнаруживать более слабые изменения. Значение δ будем выбирать рациональным, чтобы упростить расчет характеристик процедуры.

Определим последовательность случайных величин $z_i = ly_i - m$. Здесь m и l – натуральные числа, причем $m < l$. Величины z_i могут принимать значения n и $(-N)$, где $n = l - m$, $N = l + m$. Среднее значение этих величин до и после разладки определяется равенствами

$$\mathrm{M}_0\{z_i\} = -m < 0, \quad \mathrm{M}_1\{z_i\} = l\mathrm{M}_1\{y_i\} - m,$$

M_0 обозначает усреднение по распределению $F(x)$, а M_1 – по распределению $F(x - a)$. К величинам z_i применяется алгоритм кумулятивных сумм

$$S_i = \max(N, z_i + S_{i-1}), \quad S_0 = N. \tag{1}$$

Решение о наличии разладки принимается, если кумулятивная сумма S_i достигнет целочисленного порога $h > N$.

В случае, когда параметр сдвига a меньше нуля (это соответствует уменьшению среднего значения наблюдений после разладки), процедуру обнаружения можно построить аналогичным образом, заменяя величины x_i на $-x_i$.

В общем случае параметр сдвига a может быть как положительным, так и отрицательным. Одновременное использование двух АКС, один из которых ориентирован на обнаружение увеличения среднего, а другой – на уменьшение, приводит к значительным трудностям при нахождении характеристик. Предлагается процедура, в которой два АКС применяются поочередно. Переход от одного из них к другому происходит в

моменты времени, когда соответствующая кумулятивная сумма достигает крайних значений: N или h. Такой подход использовался ранее в [7] при решении задачи поиска сигнала в многоканальной системе.

Рассмотрим два АКС для обнаружения положительного и отрицательного сдвигов. Для первого из них в качестве входной последовательности будем использовать последовательность статистик $z_i^1 = l \times$ $\times \operatorname{sign}(x_i - x_{i-k}) - m$, для другого $z_i^2 = l \operatorname{sign}(x_{i-k} - x_i) - m$. Процедуру кумулятивных сумм (1) будем применять к последовательностям z_i^j, изменяя номер j последовательности z_i^j в определенные ниже моменты времени.

Будем считать последовательность, в которой после разладки среднее значение наблюдений увеличивается, первой, а последовательность, в которой среднее уменьшается, – второй. Обозначим через z_i статистики z_i^j, используемые в текущий момент времени i в АКС; P_0 – распределение величин z_i до момента разладки, а через P_1 и P_2 – распределения этих величин после разладки соответственно в первой и второй последовательностях. Тогда до момента разладки для обеих последовательностей распределения z_i одинаковы:

$$P_0(z_i = n) = P_0(z_i = -N) = 1/2.$$

После разладки распределения z_i изменяются:

$$P_1(z_i = n) = P_2(z_i = -N) = p;$$
$$P_1(z_i = -N) = P_2(z_i = n) = q, \quad p > 1/2.$$

Если в текущий момент времени i кумулятивная сумма S_i достигает верхнего порога h, то принимается решение о наличии разладки. Если сумма S_i становится меньше N, то ее значение полагается равным N и величины z_i выбираются далее из другой последовательности.

Перейдем к исследованию свойств процедуры обнаружения. Обозначим:

$T_0(j)$ – среднее число шагов, оставшихся для принятия решения о наличии разладки, если текущее значение $S_i = j$ и разладка еще не произошла;

$T_1(j)$ – среднее число шагов, оставшихся для принятия решения о наличии разладки, если разладка произошла, текущее значение $S_i = j$ и наблюдается первая последовательность;

$T_2(j)$ – среднее число шагов, оставшихся для принятия решения о наличии разладки, если разладка произошла, текущее значение $S_i = j$ и наблюдается вторая последовательность.

Тогда среднее время между ложными тревогами T равно

$$T = T_0(N),$$

а запаздывания для первой и второй последовательностей –

$$T_1 = T_1(N), \qquad T_2 = T_2(N).$$

Величины $T_k(j)$, $k = 0, 1, 2$, находятся из уравнений

$$\begin{aligned} T_0(j) &= 1 + \frac{1}{2}T_0(j+n) + \frac{1}{2}T_0(j-N); \\ T_1(j) &= 1 + pT_1(j+n) + qT_1(j-N); \\ T_2(j) &= 1 + qT_2(j+n) + pT_2(j-N) \end{aligned} \tag{2}$$

с граничными условиями

$$\begin{gathered} T_0(k) = 0, \quad k = \overline{h, h+n-1}; \quad T_0(k) = T_0(N), \quad k = \overline{0, N-1}; \\ T_1(k) = 0, \quad k = \overline{h, h+n-1}; \; T_1(k) = D(N) + T_1(N), \quad k = \overline{0, N-1}; \\ T_2(k) = T_1(N), \; k = \overline{0, N-1}, \quad k = \overline{h, h+n-1}. \end{gathered} \tag{3}$$

Здесь $D(j)$ – среднее время до окончания наблюдения второй последовательности, если текущее значение S_i равно j. Эта величина находится из уравнения

$$D(j) = 1 + qD(j+n) + pD(j-N)$$

с нулевыми граничными условиями. Общее решение для $D(j)$ имеет вид

$$D(j) = B_0 j + B_1 \nu_1^j + \ldots + B_{N+n} \nu_{N+n}^j, \quad B_0 = \frac{1}{Np - nq}, \tag{4}$$

где ν_i – корни характеристического полинома

$$P_2(\nu) = q\nu^{N+n} - \nu^N + p. \tag{5}$$

Решения уравнений (2) определяются формулами

$$T_0(j) = C_0 j + C_1 \lambda_1^j + \ldots + C_{N+n} \lambda_{N+n}^j, \quad C_0 = \frac{1}{\frac{1}{2}N - \frac{1}{2}n}; \tag{6}$$

$$T_1(j) = A_0 j + A_1 \mu_1^j + \ldots + A_{N+n} \mu_{N+n}^j, \quad A_0 = \frac{1}{Nq - Np}; \tag{7}$$

$$T_2(j) = B_0 j + b_1 \nu_1^j + \ldots + b_{N+n} \nu_{N+n}^j, \tag{8}$$

где λ_i и μ_i – соответственно корни полиномов

$$P_0(\lambda) = \lambda^{N+n} - 2\lambda^N + 1, \quad P_1(\mu) = p\mu^{N+n} - \mu^N + q. \tag{9}$$

Получим асимптотически точные (при $h \to \infty$) формулы для $T_0(N)$, $T_1(N)$ и $T_2(N)$.

Величина $T_0(N)$, определяемая соотношением (2), зависит от корней λ_i полинома $P_0(\lambda)$. В [6] с использованием теоремы Руше было установлено, что $P_0(\lambda)$ имеет только два положительных корня $\lambda_1 = 1$ и $\lambda_{N+1} = \widetilde{\lambda} > 1$, а остальные корни λ_i лежат в областях

$$|\lambda_k| < 1, \quad k = \overline{2, N}; \quad |\lambda_k| > \widetilde{\lambda}, \quad k = \overline{N+2, N+n}. \tag{10}$$

При этом все корни $P_0(\lambda)$ являются простыми. Величины C_i в (6) находятся из граничных условий (3):

$$T = T_0(N) = C_0\left[\sum_{k=2}^{N} \frac{1}{1-\lambda_k} \frac{W_\lambda(2, N; 1, k)}{W_\lambda(2, N)} - \right.$$

$$\left. - \sum_{k=N+2}^{N+n} \frac{1}{1-\lambda_k} \frac{W_\lambda(N+2, N+n; 1; k)}{W_\lambda(N+2, N+n)} - h\right] +$$

$$+\widetilde{C}\left[(1-\widetilde{\lambda}^h) - (1-\widetilde{\lambda}) \sum_{k=2}^{N} \frac{1}{1-\lambda_k} \frac{W_\lambda(2, N; N+2, k)}{W_\lambda(2, N)} + \right] + o(1),$$

где

$$\widetilde{C} = \frac{C_0\left[N - \sum\limits_{k=2}^{N} \frac{1-\lambda_k^N}{1-\lambda_k} \frac{W_\lambda(2, N; 1, k)}{W_\lambda(2, N)}\right]}{\left[(1-\widetilde{\lambda}^N) - (1-\widetilde{\lambda}) \sum\limits_{k=2}^{N} \frac{1-\lambda_k^N}{1-\lambda_k} \frac{W_\lambda(2, N; N+1, k)}{W_\lambda(2, N)}\right]} + o\left(\frac{1}{\widetilde{\lambda}^h}\right).$$

Здесь через $W_x(i, j)$ и $W_x(i, j, m, k)$ обозначены определители Вандермонда, составленные из элементов

$$W_x(i, j) = W(x_i, ..., x_j),$$
$$W_x(i, j; m, k) = W(x_i, ..., x_{k-1}; x_m, x_{k+1}, ..., x_j).$$

При нахождении $T_1(N)$, $T_2(N)$ и $D(N)$ используется тот факт, что корни ν_i и μ_i полиномов $P_2(\nu)$ и $P_1(\mu)$, определенных соотношениями

(5) и (9), расположены следующим образом: $\nu_1 = 1$, $\nu_{N+1} = \widetilde{\nu} > 1$,

$$|\nu_k| < 1, \quad k = \overline{2,N}; \quad |\nu_k| > \widetilde{\nu}, \quad k = \overline{N+2, N+n}; \tag{11}$$

$$\mu_1 = 1, \quad mu_{N+1} = \widetilde{\mu} < 1, \quad |\mu_k| < \widetilde{\mu}, \quad k = \overline{2,N}; \tag{12}$$

$$|\mu_k| > 1, \quad k = \overline{N+2, N+n}.$$

Учитывая (4), (7) и (8) и решая систему (3) для определения неизвестных постоянных, получаем

$$D = D(N) = B_0\left[N - \sum_{k=2}^{N} \frac{1-\nu_k^N}{1-\nu_k} \frac{W_\nu(2,N;1,k)}{W_\nu(2,N)}\right] + o(1); \tag{13}$$

$$T = T_1(0) = -A_0\left[h + \sum_{k=N+2}^{N+n} \frac{1}{1-\mu_k} \frac{W_\mu(N+2,N+n;1,k)}{W_\mu(N+2,N+n)} - \right.$$

$$\left. - \sum_{k=2}^{N} \frac{1}{1-\mu_k} \frac{W_\mu(2,N;1,k)}{W_\mu(2,N)}\right] +$$

$$+\widetilde{A}\left[1 - (1-\widetilde{\mu}) \sum_{k=2}^{N} \frac{1-\mu_k^h}{1-\mu_k} \frac{W_\mu(2,N;N+1,k)}{W_\mu(2,N)}\right] + o(1), \tag{14}$$

где

$$\widetilde{A} = \frac{D + A_0\left[(N) - \sum_{k=2}^{N} \frac{1-\mu_k^N}{1-\mu_k} \frac{W_\mu(2,N;1,k)}{W_\mu(2,N)}\right]}{\left[(1-\widetilde{\mu}^N) - (1-\widetilde{\mu}) \sum_{k=2}^{N} \frac{1-\mu_k^N}{1-\mu_k} \frac{W_\mu(2,N;N+1,k)}{W_\mu(2,N)}\right]} + o(1). \tag{15}$$

Величины $T_1(N)$ и $T_2(N)$ определяются равенствами

$$\begin{aligned} T_1 &= T_1(N) = T - D, \\ T_2 &= T_2(N) = T_1(N) + D = T. \end{aligned} \tag{16}$$

Таким образом, среднее время между ложными тревогами T растет экспоненциально при $h \to \infty$, а T_1 и T_2 — линейно по h:

$$\begin{aligned} T &= L_1\widetilde{\lambda}^h + L_2 h + L_3 + o(1); \\ T_1 &= Kh + K_1 + o(1); \\ T_2 &= Kh + K_2 + o(1). \end{aligned} \tag{17}$$

Процедуры обнаружения, обладающие такими свойствами, принято называть [4] оптимальными.

СПИСОК ЛИТЕРАТУРЫ

1. *Page E.S.* Continuous inspection schemes // Biometrika. 1954. No. 1. P. 141-145.
2. *Ширяев А.Н.* Статистический последовательный анализ. М.: Наука, 1976.
3. *Lorden G.* Procedures for reacting to a change in distribution // Annals Math. Statistics. 1971. Vol. 42. No. 6. P. 1897-1908.
4. *Дарховский Б.С., Бродский Б.Е.* Непараметрический метод скорейшего обнаружения изменения среднего случайной последовательности // Теория вероятностей и ее применение. 1987. Т. 32. № 4. С. 703-711.
5. *Yakir B.* On the average run length to false alarm in surveillance problems which possess an invariance structure // Annals Math. Statistics. 1998. Vol. 26. No. 3. P. 1198-1214.
6. *Воробейчиков С.Э.* Об обнаружении изменения среднего в последовательности случайных величин // Автоматика и телемеханика. 1998. № 3. С. 50-56.
7. *Воробейчиков С.Э., Конев В.В.* Последовательная процедура обнаружения разладки в многоканальной системе // Радиоэлектроника. 1990. Т. 35. № 10. С. 2104-2111.

Ю.Г. Дмитриев, Ж.Н. Зенкова

Об оценивании симметричного распределения по цензурированной выборке

В работе рассматривается проблема учета дополнительной информации при построении статистических процедур на основе цензурированных данных. Строится оценка функции распределения случайной величины с учетом симметрии относительно известного центра α, исследуются свойства полученной оценки.

Привлечение дополнительной информации для улучшения качества статистических процедур рассматривалось в ряде работ [1–6]. Однако существующие методы основаны на полной выборке. Зачастую данные представлены в виде урезанных, цензурированных выборок [8]. Для такого класса задач процедуры привлечения априорной информации требуют своей разработки.

В данной работе строится модифицированная эмпирическая функция распределения (м.э.ф.р.) с учетом априорной информации о симметрии функции распределения (ф.р.) относительно известного центра α для цензурированной выборки (ц.в.).

Здесь исследовалось время наработки до отказа невосстанавливаемого элемента, рассматривались прогрессивно цензурированные выборки (п.ц.в.), моменты цензурирования T_j, $j = \overline{1,m}$, – неслучайные, $T_0 = 0$, $T_m = T$.

Пусть $(X, I) = \{(X_1, I_1), (X_2, I_2), ..., (X_N, I_N)\}$ – п.ц.в. объема N, где для $i = \overline{1,N}$

$$I_i = \begin{cases} 0, & \text{если } X_i \text{ – наработка до отказа,} \\ 1, & \text{если } X_i \text{ – условная наработка} \\ & \text{или наработка до цензурирования.} \end{cases}$$

Пусть нам известно, что $F(t)$, $t \in [0,T]$, – ф.р. случайной величины (с.в.) τ – времени наработки до отказа – является симметричной относительно известного центра симметрии α, т. е. $F(t)$ удовлетворяет условию

$$F(t) = 1 - F(2\alpha - t + 0),\ t \in [0,T]. \tag{1}$$

Здесь $\alpha = T/2$.

Для $i = \overline{1,m}$ введем обозначения:

r_i – число полных наработок в интервале $(T_{i-1}, T_i]$, $r_0 = 0$;

n_i – число условных наработок или наработок до цензурирования в интервале $(T_{i-1}, T_i]$, $n_0 = 0$;

$N_i = N - \sum_{j=0}^{i-1}(r_i + n_i)$ — число не отказавших изделий в момент T_i;

$N_i^* = N_{i-1}^* - \dfrac{n_{i-1}}{1 - \widehat{F}_{i-1}}$, $N_0^* = N$, где для $j = \overline{1,m}$

$$\widehat{F}_j = \begin{cases} \widehat{F}_{j-1} + \dfrac{r_j}{N_j^*} = \sum\limits_{s=0}^{j-1} \dfrac{r_s}{N_s^*}, & N_s^* > 0, \\ \widehat{F}_{j-1}, & N_j^* \le 0,\ N_s^* > 0,\ s = \overline{1, j-1} \end{cases}$$

– оценка ф.р. $F(t)$ в точке T_j, $\widehat{F}_0 = 0$.

Тогда эмпирическая ф.р. (э.ф.р.), построенная на основе ц.в., имеет

вид [8, с. 45]

$$F_N(t)=\begin{cases} 0, & \text{если } t\le 0,\\ \dfrac{l_1}{N_1}, & \text{если } t\in(0,T_1],\ l_1=\overline{0,r_1},\\ \left.\begin{array}{ll}\widehat{F}_1+(1-\widehat{F}_1)\dfrac{l_2}{N_2}, & N_2>0,\\ \widehat{F}_1, & N_2\le 0,\end{array}\right\}, & \\ & \text{если } t\in(T_1,T_2],\ l_2=\overline{0,r_2},\\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ \left.\begin{array}{ll}\widehat{F}_{m-1}+(1-\widehat{F}_{m-1})\dfrac{l_m}{N_m}, & N_m>0,\\ \widehat{F}_{m-1}, & N_m\le 0,\end{array}\right\}, & \\ & \text{если } t\in(T_{m-1},T_m],\ l_m=\overline{0,r_m}. \end{cases} \tag{2}$$

Заметим, что э.ф.р. для полной выборки $F_N^*(t)$ может быть представлена как частный случай формулы (2), если рассмотреть $m=1$, $T_m=+\infty$.

Необходимо построить э.ф.р. с.в. τ по п.ц.в. (X,I) с учетом свойства (1).

Допустим, что ц.в. состоит из условных наработок и наработок до отказа. М.э.ф.р. $F_N^s(t)$ для выборки (X,I) определяется формулой [7]

$$F_N^s(t)=\frac{F_N(t)+1-F_N(2\alpha-t+0)}{2}, \tag{3}$$

где $F_N(t)$ – э.ф.р., построенная по формуле (2). М.э.ф.р. $F_N^s(t)$ является симметричной [2].

Рассмотрим п.ц.в., построенные по следующей схеме цензурирования: для $j=\overline{1,m}$ величина n_j — случайная и численно равна доле g_j, $0<g_j<1$, от числа исправных изделий в конце интервала $(T_{j-1},T_j]$.

При данном способе цензурирования $F_N(t)$ является несмещенной оценкой $F(t)$ [8], поэтому

$$\mathrm{M}F_N^s(t)=F(t).$$

Рассмотрим дисперсию м.э.ф.р.

$$\mathrm{D}F_N^s(t)=\frac{1}{4}\mathrm{D}F_N(t)+\frac{1}{4}\mathrm{D}F_N(2\alpha-t+0)-$$

$$-\frac{1}{2}\mathbf{cov}\left(F_N(t), F_N(2\alpha - t + 0)\right).$$

Здесь $\mathbf{cov}(\xi, \eta)$ – ковариация двух с.в. ξ и η, $\mathbf{M}\xi$ – математическое ожидание с.в. ξ, $\mathbf{D}\xi$ – дисперсия с.в. ξ.

Очевидно, что $\mathbf{D}F_N^s(t) = \mathbf{D}F_N^s(2\alpha - t + 0)$.

Обозначим

$$F_{\min}(t) = \min\left\{F(t), F(2\alpha - t + 0)\right\};$$

$$F_{\max}(t) = \max\left\{F(t), F(2\alpha - t + 0)\right\};$$

$$\widehat{F}_{\min}(t) = \min\left\{F_N(t), F_N(2\alpha - t + 0)\right\};$$

$$\widehat{F}_{\max}(t) = \max\left\{F_N(t), F_N(2\alpha - t + 0)\right\}.$$

Очевидно, что в силу (1)

$$F_{\min}(t) = 1 - F_{\max}(t). \tag{4}$$

Тогда [9]

$$\mathbf{cov}\left(F_N(t), F_N(2\alpha - t + 0)\right) = \frac{1 - F_{\max}(t)}{1 - F_{\min}(t)}\mathbf{D}\widehat{F}_{\min}(t) =$$

$$= \frac{F_{\min}(t)}{1 - F_{\min}(t)}\mathbf{D}\widehat{F}_{\min}(t),$$

поэтому

$$\mathbf{D}F_N^s(t) = \frac{1}{4}\left(1 - \frac{2F_{\min}(t)}{1 - F_{\min}(t)}\right)\mathbf{D}\widehat{F}_{\min}(t) + \frac{1}{4}\mathbf{D}\widehat{F}_{\max}(t). \tag{5}$$

В [8] получена следующая формула для дисперсии $\mathbf{D}F_N(t)$ в точках T_j, $j = \overline{0, m}$

$$\mathbf{D}\widehat{F}_j = \frac{F_j(1 - F_j)}{N} + \sum_{i=1}^{j}\frac{(F_j - F_i)(1 - F_j)g_j}{N(1 - F_i)\prod\limits_{s=1}^{i}(1 - g_s)}. \tag{6}$$

Дисперсия $F_N(t)$ возрастает с ростом количества интервалов цензурирования.

Выразим в формуле (5) $\mathbf{D}\widehat{F}_{\max}(t)$ через $\mathbf{D}\widehat{F}_{\min}(t)$.

Для простоты рассмотрим случай

$$\widehat{F}_{\min}(t) = \widehat{F}_k,\ \widehat{F}_{\max}(t) = \widehat{F}_{k+j}, \tag{7}$$

где $\widehat{F}_k$, $\widehat{F}_{k+j}$ – м.э.ф.р. в точках T_k и T_{k+j} соответственно, $k = \overline{0, m-1}$, $j > 0$, $k + j \le m$.

Пользуясь тем, что $F_N^s(t)$ является несмещенной оценкой э.ф.р. и $N_i > 0$ для $i = \overline{1, m}$ при данном способе цензурирования , получим

$$\mathbf{D}\widehat{F}_{k+j} = \mathbf{M}\widehat{F}_{k+j}^2 - F_{k+j}^2 = \mathbf{M}\left(\widehat{F}_{k+j-1} + \left(1 - \widehat{F}_{k+j-1}\right)\frac{r_{k+j}}{N_{k+j}}\right)^2 - F_{k+j}^2.$$

Используя свойства условного математического ожидания, а также тот факт, что с.в. r_{k+j} имеет биномиальное распределение

$$\mathbf{Bi}\left(N_{k+j}, \frac{F_{k+j} - F_{k+j-1}}{1 - F_{k+j-1}}\right),$$

получим

$$\mathbf{D}\widehat{F}_{k+j} = \mathbf{M}\widehat{F}_{k+j-1}^2 + 2\mathbf{M}\left\{\widehat{F}_{k+j-1}\mathbf{M}\left\{\frac{r_{k+j}}{N_{k+j}}\middle| r_{k+j-1}\right\}\right\} -$$

$$-2\mathbf{M}\left\{\widehat{F}_{k+j-1}^2\mathbf{M}\left\{\frac{r_{k+j}}{N_{k+j}}\middle| r_{k+j-1}\right\}\right\} +$$

$$+\mathbf{M}\left\{\left(1 - \widehat{F}_{k+j-1}\right)^2\mathbf{M}\left\{\frac{r_{k+j}^2}{N_{k+j}^2}\middle| r_{k+j-1}\right\}\right\} - F_{k+j}^2 =$$

$$= \mathbf{M}\widehat{F}_{k+j-1}^2 + 2F_{k+j-1}\frac{F_{k+j} - F_{k+j-1}}{1 - F_{k+j-1}} - 2\frac{F_{k+j} - F_{k+j-1}}{1 - F_{k+j-1}}\mathbf{M}\widehat{F}_{k+j-1}^2 +$$

$$+\mathbf{M}\left\{\left(1 - \widehat{F}_{k+j-1}\right)^2\mathbf{M}\left\{\frac{r_{k+j}^2}{N_{k+j}^2}\middle| r_{k+j-1}\right\}\right\} - F_{k+j}^2.$$

Т.к. с.в. r_{k+j} имеет биномиальное распределение, то [9]

$$\mathbf{M}r_{k+j}^2 = \frac{(F_{k+j} - F_{k+j-1})(1 - F_{k+j})}{(1 - F_{k+j-1})^2}N_{k+j} + \left(\frac{F_{k+j} - F_{k+j-1}}{1 - F_{k+j-1}}\right)^2 N_{k+j}^2,$$

поэтому

$$\mathbf{D}\widehat{F}_{k+j} = \left(\frac{1 - F_{k+j}}{1 - F_{k+j-1}}\right)^2\mathbf{D}\widehat{F}_{k+j-1} +$$

$$+\frac{(F_{k+j}-F_{k+j-1})(1-F_{k+j})}{(1-F_{k+j-1})^2}\mathbf{M}\frac{(1-\widehat{F}_{k+j-1})^2}{N_{k+j}}.$$

Для достаточно больших N [8]

$$\mathbf{M}\left\{\frac{1}{N_{k+j}}\middle| r_{k+j-1}\right\}\approx\frac{1}{\mathbf{M}\{N_{k+j}|\,r_{k+j-1}\}}=$$

$$=\frac{1}{N(1-F_{k+j-1})\prod\limits_{s=1}^{k+j-1}(1-g_s)},$$

поэтому, пренебрегая слагаемым порядка $o\left(\frac{1}{N}\right)$, получим

$$\mathbf{D}\widehat{F}_{k+j}=\left(\frac{1-F_{k+j}}{1-F_{k+j-1}}\right)^2\mathbf{D}\widehat{F}_{k+j-1}+\frac{(F_{k+j}-F_{k+j-1})(1-F_{k+j})}{N(1-F_{k+j-1})\prod\limits_{s=1}^{k+j-1}(1-g_s)}.$$

Аналогично для $s=\overline{0,m-1}$, $t\in[T_s;T_{s+1})$ можно получить

$$\mathbf{D}F_N(t)=\left(\frac{1-F(t)}{1-F_s}\right)^2\mathbf{D}\widehat{F}_s+\frac{(F(t)-F_s)(1-F(t))}{N(1-F_s)\prod\limits_{i=1}^{s}(1-g_i)}.$$

Повторив данную процедуру $j-1$ раз, получим

$$\mathbf{D}\widehat{F}_{k+j}=\left(\frac{1-F_{k+j}}{1-F_k}\right)^2\mathbf{D}\widehat{F}_k+$$

$$+\sum_{i=0}^{j-1}\left(\frac{1-F_{k+j}}{1-F_{k+j-i}}\right)^2\frac{\left(F_{k+j-i}-F_{k+j-(i+1)}\right)\left(1-F_{k+j-i}\right)}{N\left(1-F_{k+j-(i+1)}\right)\prod\limits_{s=1}^{k+j-(i+1)}\left(1-g_s\right)}.$$

Подставим полученное выражение в формулу (5), учитывая (4), (6) и (7).

$$\mathbf{D}F_N^s(T_k)=\frac{1}{4}\left(1-\frac{2F_k}{1-F_k}\right)\mathbf{D}\widehat{F}_k+\frac{1}{4}\mathbf{D}\widehat{F}_{k+j}=\left(\frac{\frac{1}{2}-F_k}{1-F_k}\right)^2\mathbf{D}\widehat{F}_k+$$

$$+\frac{1}{4}\sum_{i=0}^{j-1}\left(\frac{F_k}{1-F_{k+j-i}}\right)^2\frac{\left(F_{k+j-i}-F_{k+j-(i+1)}\right)\left(1-F_{k+j-i}\right)}{N\left(1-F_{k+j-(i+1)}\right)\prod\limits_{s=1}^{k+j-(i+1)}\left(1-g_s\right)})=$$

$$= \left(\frac{\frac{1}{2} - F_k}{1 - F_k}\right)^2 \cdot \frac{F_k(1 - F_k)}{N} + \frac{\left(\frac{1}{2} - F_k\right)}{1 - F_k} \sum_{i=1}^{k} \frac{(F_k - F_i)\left(\frac{1}{2} - F_k\right) g_i}{N(1 - F_i) \prod_{s=1}^{i}(1 - g_s)} +$$

$$+ \frac{1}{4} \sum_{i=0}^{j-1} \left(\frac{F_k}{1 - F_{k+j-i}}\right)^2 \frac{\left(F_{k+j-i} - F_{k+j-(i+1)}\right)\left(1 - F_{k+j-i}\right)}{N\left(1 - F_{k+j-(i+1)}\right) \prod_{s=1}^{k+j-(i+1)} (1 - g_s)}.$$

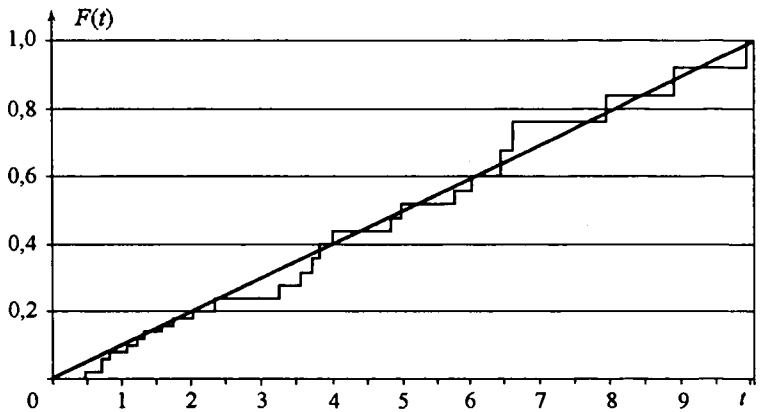


Рис. 1. Графики функций $F_N(t)$ и $F(t) = 0,1t$, $t \in [0,10]$ ч, при $N = 50$, $T_1 = 2$ ч, $T_2 = 6$ ч.

Построим м.э.ф.р. на основе ц.в., приведенной в качестве примера в [8, с. 48]. Рассматривалась полная выборка объемом $N = 50$ наработок изделий до отказа из равномерного в [0,10] ч распределения, моменты цензурирования $T_1 = 2$ ч, $T_2 = 6$ ч. Ц.в. представлена в виде ряда: 0,5; 0,715; 0,727; 0,842; 1,09; 1,21; 1,31; 1,57; 1,73; 2; 2*(20); 2,32; 3,28; 3,59; 3,72; 3,83; 4; 4,8; 5; 5,75; 6; 6*(5); 6,36; 6,59; 7,9; 8,9; 9,9. В этом ряду звездочкой помечены неполные наработки, а числом в скобках – их количество. По формуле (2) строится э.ф.р. (см. рис. 1).

Равномерная ф.р. $F(t) = 0,1t$, $t \in [0,10]$ ч, является симметричной относительно центра $\alpha = 5$ ч, поэтому можно использовать формулу (3) для нахождения м.э.ф.р. (см. рис. 2).

Для данного случая ($T_1 = 2$ ч, $T_2 = 6$ ч.) приведены графики дисперсий оценок ф.р. (см. рис. 3). Очевидно, что значения дисперсии э.ф.р., построенной по формуле (3), удалось уменьшить путем модификации э.ф.р.

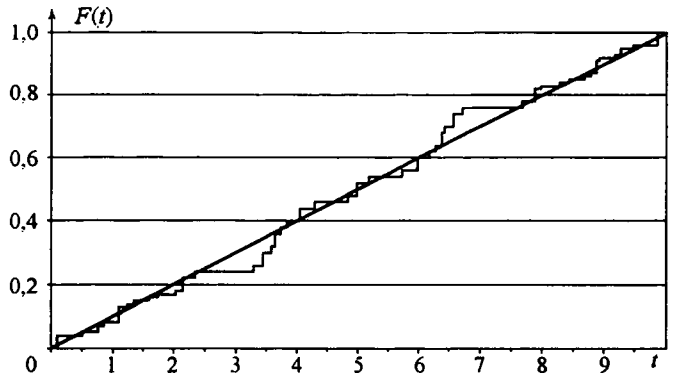


Рис. 2. Графики функций $F_N^s(t)$ и $F(t) = 0,1t$, $t \in [0, 10]$ ч, при $N = 50$, $T_1 = 2$ ч, $T_2 = 6$ ч.

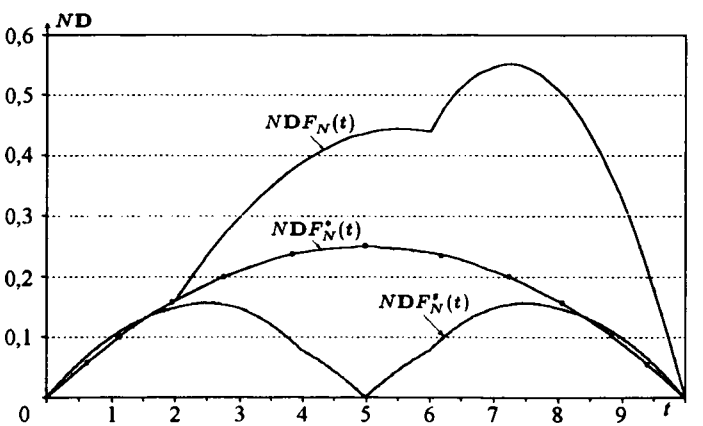


Рис. 3. Графики зависимостей дисперсий э.ф.р. для полной выборки ($N\mathbf{D}F_N^*(t)$), э.ф.р. для п.ц.в. справа ($N\mathbf{D}F_N(t)$) и м.э.ф.р. ($N\mathbf{D}F_N^s(t)$) при $N = 50$, $g_1 = g_2 = 0,5$.

СПИСОК ЛИТЕРАТУРЫ

1. *Пугачёв В.Н.* Комбинированные методы определения вероятностных характеристик. М.: Советское радио, 1973.
2. *Дмитриев Ю.Г., Устинов Ю.К.* Статистическое оценивание распределений вероятностей с использованием дополнительной информации. Томск: Изд-во Том. ун-та, 1988. 194 с.
3. *Дмитриев Ю.Г.* Статистическая обработка данных с использованием априорной информации. Дисс. ... докт. физ.-мат. наук. 2000.
4. *Дмитриев Ю.Г., Зенкова Ж.Н.* Модифицированная с учетом S_α^p-симметрии эмпирическая функция распределения // Материалы Всероссийской научно практической конференции "Информационные технологии и математическое моделирование". Томск: Твердыня, 2002. С. 368-370.
5. *Дмитриев Ю.Г., Зенкова Ж.Н.* Критерии согласия Колмогорова для симметричных распределений // Вестник ТГУ. Приложение. № 6, сент. 2003. С. 265-269.
6. *Зенкова Ж.Н.* Модифицированный критерий согласия Колмогорова для неравноплечно симметричных распределений // Обозрение прикладной и промышленной математики. Т. 9, Вып. 3, окт. 2002. М., 2002. С. 612-613.
7. *Зенкова Ж.Н.* Оценка функции распределения по прогрессивно цензурированной выборке с учетом симметрии // Обозрение прикладной и промышленной математики. М.: 2003. Т. 10. Вып. 3. С. 651-652.
8. *Анализ* надежности технических систем по цензурированным выборкам / В.М. Скрипник, А.Е. Назин, Ю.Г. Приходько, Ю.Н. Благовещенский. М.: Радио и связь, 1988. 184 с.
9. *Скрипник В.М., Назин А.Е.* Оценка надежности технических систем по цензурированным выборкам. Минск: Наука и техника, 1981. 144 с.

Ю.Г. Дмитриев, Ф.П. Тарасенко

Информационные подходы к непараметрическому оцениванию функционалов от распределений [1]

Обсуждается проблема применения информационных подходов к непараметрическому оцениванию функционалов от распределения при априорных условиях на класс допустимых распределений.

[1] Данная работа опубликована в журнале "Вестник ТГУ". Приложение № 6. Сентябрь 2003. С. 270-272

1. Введение

Теория информации настолько тесно связана с теорией вероятностей, что в некоторых последних учебниках первая излагается как раздел второй [1]. Связь теории информации с математической статистикой тоже несомненна, но проявляется в разных формах благодаря разнообразию задач статистики (оценивание параметров, проверка гипотез, выявление закономерностей, и пр.) и их аспектов (синтез и анализ процедур). Применение информационных подходов к проблемам статистики оказалось настолько эффективным, что А.Н. Колмогоров четко сформулировал мысль о том, что "навыки мысли и аналитический аппарат теории информации должны, по-видимому, привести к заметной *перестройке* здания математической статистики" [2]. Однако выполнение этой программы проходит медленно и еще далеко от завершения.

Можно отметить следующие направления теоретико-информационных приложений в статистике.

а) Использование в синтезе индивидуальных информационных мер; т.е. построение статистических процедур на основе прямого вычисления количества информации, содержащегося в конкретной выборке об интересующей нас характеристике распределения. Например, если принимать решения в пользу одного из двух конкурирующих известных распределений на основе разницы между индивидуальными количествами информации в выборке о каждой из них, то автоматически получается [4], что тестовая статистика есть просто логарифм отношения правдоподобия.

б) Использование усредненных информационных мер при анализе качества процедур. Например, оказалось [5], что мощность любого теста согласия зависит только от расхождения Кульбака между (приведенными в [0,1]) конкурирующими распределениями, но не от их вида.

Существенным и широко известным прорывом в построении статистики на основе информационных мер стала монография С. Кульбака [3]. Однако он ограничился рассмотрением параметрических задач, простых гипотез и линейных процедур, так что имеется широкий простор для дальнейших продвижений.

Похоже, наиболее кардинальный и продуктивный подход состоит в том, чтобы в основу синтеза процедур положить не оптимизацию привычных критериев качества (типа среднеквадратических отклонений или вероятностей ошибок), а прямое использование информационных мер (впоследствии оказывается, что в некоторых конкретных случаях порождаются либо те же оптимальные процедуры, либо асимптотически эквивалентные, но обычно применяемые в более общих случаях).

Приведем некоторые результаты такого подхода к проблеме оценивания функционалов от распределения вероятностей.

2. Модифицированные эмпирические распределения и информационные статистики

Рассмотрим на измеримом пространстве $(\mathcal{X}, \mathcal{B})$ распределения P, для которых при заданных функциях $\psi_1, \ldots, \psi_m$, отображающих $\mathcal{X}$ в R, известны значения $\beta_1, \ldots, \beta_m$ математических ожиданий

$$b_j(P) = \int_{\mathcal{X}} \psi_j(x) P(dx) = \beta_j, \quad j = \overline{1,m}. \tag{1}$$

Класс распределений, удовлетворяющих условию (1), обозначим через $\mathcal{P}_\beta^\psi$, где $\psi = (\psi_1, \ldots, \psi_m)$, $\beta = (\beta_1, \ldots, \beta_m)$.

Рассмотрим информационное расстояние Кульбака–Лейблера

$$d_{\mathrm{KL}}(P;Q) = \int_{\mathcal{X}} \ln \frac{dP}{dQ}(x) P(dx), \tag{2}$$

где распределение P абсолютно непрерывно относительно Q на $(\mathcal{X}, \mathcal{B})$. Формула (2) задает расстояние Кульбака–Лейблера, которое в [3, с. 16] называется информационной мерой направленного расхождения.

Пусть $X = (X_1, \ldots, X_N)$ – выборка объема N из распределения P,

$$P_N(B) = N^{-1} \sum_{i=1}^{N} I_{X_i}(B)$$

– эмпирическое распределение (э.р.). Здесь $I_x(B)$, $B \in \mathcal{B}$ – единичная мера, сосредоточенная в точке x. В качестве оценки распределения P при условии $P \in \mathcal{P}_\beta^\psi$ возьмем d_{KL}-проекцию э.р. P_N в априорный класс $\mathcal{P}_\beta^\psi$ (см. [6]):

$$P_N^\psi(A) = \sum_{i=1}^{N} I_{X_i}(A) \frac{\exp\{\tau_N^\circ \psi(X_i)\}}{\sum\limits_{j=1}^{N} \exp\{\tau_N^\circ \psi(X_j)\}}, \quad A \in B, \tag{3}$$

где под умножением векторов понимается их скалярное произведение а $\tau_N^\circ = (\tau_{1N}^\circ, \ldots, \tau_{mN}^\circ)$ есть решение системы нелинейных уравнений

$$\beta = \sum_{i=1}^{N} \psi(X_i) \left[\sum_{j=1}^{N} \exp\{\tau^\circ(\psi(X_j) - \psi(X_i))\} \right]^{-1}. \tag{4}$$

Таким образом,

$$d_{\mathrm{KL}}(P_N^{\psi}; P_N) = \inf_{P \in \mathcal{P}_{\beta}^{\psi}} d_{\mathrm{KL}}(P; P_N).$$

Определим *информационные статистики*:

$$d_{\mathrm{KL}}(P_N^{\psi}; P_N) = \tau_N^{\circ}\beta - \ln K_{P_N}(\tau_N^{\circ}); \tag{5}$$

$$d_{\mathrm{KL}}(P_N; P_N^{\psi}) = -\tau_N^{\circ} \int_{\mathcal{X}} \psi(x) P_N(dx) + \ln K_{P_N}(\tau_N^{\circ}); \tag{6}$$

$$\mathcal{I}(P_N^{\psi}; P_N) = \tau_N^{\circ} \int_{\mathcal{X}} (\beta - \psi(x)) P_N(dx), \tag{7}$$

где $K_{P_N}(\tau_N^{\circ}) = \int_{\mathcal{X}} e^{\tau_N^{\circ}\psi(x)} P_N(dx)$. Статистика (5) является мерой направленного расхождения между распределением P_N^{ψ}, в котором наряду со статистической содержится и априорная информация, и распределением P_N, содержащим только статистическую информацию, сосредоточенную в реализовавшейся выборке. Чем больше значение $d_{\mathrm{KL}}(P_N^{\psi}; P_N)$, тем меньше в реализовавшейся выборке X объема N информации о принадлежности P классу $\mathcal{P}_{\beta}^{\psi}$. С другой стороны, если $d_{\mathrm{KL}}(P_N^{\psi}; P_N) = 0$, то в данной реализации выборки содержится вся информация об априорном классе (это означает, что $\int_{\mathcal{X}} \psi dP_N = \beta$).

Статистика (6) также является мерой направленного расхождения, но с направлением, противоположным (5). Она характеризует *количество информации*, содержащейся в априорном классе $\mathcal{P}_{\beta}^{\psi}$ (точнее в распределении P_N^{ψ}) о выборке X, соответствующей э.р. P_N.

Статистика (7) есть суммарное расхождение между проекцией э.р. P_N^{ψ} и э.р. P_N и характеризует суммарную информацию, содержащуюся как в выборке X о $\mathcal{P}_{\beta}^{\psi}$, так и в априорном классе $\mathcal{P}_{\beta}^{\psi}$ о X. Если $P_N \in \mathcal{P}_{\beta}^{\psi}$, то τ_N° есть нулевой вектор и значения всех трех информационных статистик равны нулю.

Асимптотическое распределение оценки $P_N^{\psi}(A)$ определяется следующей теоремой.

Теорема 1. *Если выборка* $X \in P_0$ *и* $P_0 \in \mathcal{P}_{\beta}^{\psi}$, $\int_{\mathcal{X}} \psi_j^2 P_0(dx) < \infty$, $j = \overline{1,m}$, *матрица ковариаций*

$$V_{P_0} = \Big\| \int_{\mathcal{X}} \psi_j \psi_l dP_0 - \int_{\mathcal{X}} \psi_j dP_0 \cdot \int_{\mathcal{X}} \psi_l dP_0 \Big\|, \quad j,l = \overline{1,m}, \tag{8}$$

невырождена, тогда при $N \to \infty$

1) $\mathcal{L}\{\sqrt{N}(P_N^{\psi}(A) - P_0(A))\} \to \mathcal{N}(0, \sigma_{\psi}^2(A)),\ A \in \mathcal{B}$,

где $\mathcal{N}(0, \sigma_{\psi}^2(A))$ *– нормальное распределение с нулевым математическим ожиданием и дисперсией*

$$\sigma_{\psi}^2(A) = P_0(A) - P_0^2(A) - \left[\int_A (\psi(x) - \beta) P_0(dx) V_{P_0}^{-1} \int_A (\psi(x) - \beta)^T P_0(dx)\right];$$

2) *распределения статистик*

$$N\mathcal{I}_{P_N}^{\psi},\ 2N d_{\mathrm{KL}}(P_N^{\psi}; P_N) \quad \text{и} \quad 2N d_{\mathrm{KL}}(P_N; P_N^{\psi})$$

сходятся к распределению χ_m^2 *(*χ*-квадрат с* m *степенями свободы).*

СПИСОК ЛИТЕРАТУРЫ

1. *Боровков А. А.* Теория вероятностей. М.: Наука, 1976. 352 с.
2. *Колмогоров А. Н.* Предисловие редактора перевода к книге [3]. С. 5-6.
3. *Кульбак С.* Теория информации и статистика. М.: Наука, 1967. 408 с.
4. *Тарасенко Ф. П.* Об информационном подходе к методам приема // Сб. "Научная сессия НТОРЭС, посвященная Дню Радио". М.: Трудрезервиздат, 1958. С. 4.
5. *Жирова Л. П., Тарасенко Ф. П., Шуленин В. П.* Об экспериментальном сравнении критериев согласия // Труды СФТИ. Томск: Изд-во Том. ун-та, 1973. Вып. 63. С. 169-179.
6. *Дмитриев Ю. Г., Устинов Ю. К.* Статистическое оценивание распределений вероятностей с использованием дополнительной информации. Томск: Изд-во Том. ун-та, 1988. 194 с.

G.M. Koshkin and F.P. Tarasenko

Continuous-Discrete Stochastic Objects: Nonparametric Identification and Control[1]

A memoryless stochastic object is considered, part of inputs and outputs of which are observable and measurable continuous stationary processes, and others are discrete sequences of weakly dependent random variables. Inner feedbacks are supposed to be absent, so we can present the object by a system of equations between simultaneous inputs and outputs. Under

[1]Research was supported by the Russian Foundation of Basic Research (grant No. 98-07-03194)

nonparametric uncertainty of distributions a model of the object is built, which is the Rosenblatt–Parzen regression estimates of the functions based on observations of the processes and sequences.

The conditions of mean-square-sense convergence are found for the estimates in case of uniform strong mixing property for processes and sequences. A problem of control of outputs by certain inputs is solved with the proposed algorithms. The algorithms are easily feasible with digital and analog computers.

1. Introduction

This paper is based on the results published in a preprint of 8-th IFAC/IFORS Symposium on Identification and System Parameter Estimation, that was held in Beijing, China, August 27-31, 1988 [1].

Let us consider a memoryless stochastic object with p inputs

$$Z(t) = \left(Z^{[1]}(t), \ldots, Z^{[p]}(t)\right)$$

and s outputs

$$Y(t) = \left(Y^{[1]}(t), \ldots, Y^{[s]}(t)\right).$$

Inner feedbacks are supposed to be absent, so the object can be described by a system of s equations

$$Y^{[i]}(t) = F_i(Z(t)), \qquad i = \overline{1,s}, \tag{1}$$

where $F_i(\cdot)$, $i = \overline{1,s}$ are unknown continuous functions in a region Ω.

Let some of inputs $p' < p$ and outputs $s' < s$ be observable and measurable stationary in a narrow sense continuous in time processes

$$Z'(t) = \left(Z^{[1]}(t), \ldots, Z^{[p']}(t)\right),\ Y'(t) = \left(Y^{[1]}(t), \ldots, Y^{[s']}(t)\right),\ t \in [0, \infty],$$

and let the other $p'' = p - p'$ inputs and $s'' = s - s'$ outputs be observable sequences of identically distributed random variables

$$Z_i'' = \left(Z_i^{[p'+1]}, \ldots, Z_i^{[p]}\right), \quad Y_i'' = \left(Y_i^{[s'+1]}, \ldots, Y_i^{[s]}\right) \quad i = 1, \ldots, n.$$

As it is well known [2], a mathematical model of such object, which minimizes mean-square-error (MSE) between outputs of the object and of the model of it, is a system of a regression equations:

$$R_i(z) = \mathbf{E}\left(Y^{[i]}(t)|z\right) = \int y^{[i]} f_i\left(y^{[i]}, z\right) dy^{[i]} \,/\, \varphi(z), \qquad i = \overline{1,s}, \tag{2}$$

where $z \in \Omega$, $\varphi(z) = \dfrac{\partial^p}{\partial z}\mathbf{P}\{Z(t) < z\}$ is a probability density function for inputs,

$$f_i\left(y^{[i]}, z\right) = \frac{\partial^{p+1}}{\partial y^{[i]}\partial z}\mathbf{P}\left\{Y^{[i]} < y^{[i]}, Z(t) < z\right\}$$

is a density function for i-th output and all inputs, which is defined separately for continuous and discrete outputs:

$$f_i\left(y^{[i]}, z\right) = \frac{\partial^{p+1}}{\partial y^{[i]}\partial z}\mathbf{P}\left\{Y^{[i]}(t) < y^{[i]}, Z'(t) < z', Z''_1 < z''\right\}, \quad \text{if} \quad i \in \mathcal{L}';$$

$$f_i\left(y^{[i]}, z\right) = \frac{\partial^{p+1}}{\partial y^{[i]}\partial z}\mathbf{P}\left\{Y_1^{[i]} < y^{[i]}, Z'(t) < z', Z''_1 < z''\right\}, \quad \text{if} \quad i \in \mathcal{L}'';$$

$$\mathcal{L}' = \{1, 2, \ldots, s'\}, \quad \mathcal{L}'' = \{s'+1, s'+2, \ldots, s\}.$$

Here and later in this paper the vector operators supposed to act on components of the vector separately, i.e.

$$\partial z = \partial z^{[1]} \cdots \partial z^{[p]}, \quad (Z(t) < z) = \left(Z^{[1]}(t) < z^{[1]}, \ldots, Z^{[p]}(t) < z^{[p]}\right).$$

Then, let us introduce a denotation

$$\left(\frac{z - Z(t,j)}{h_{Tn}}\right) = \left(\frac{z' - Z'(t)}{h(T)}, \frac{z'' - Z''_j}{h(n)}\right).$$

2. A Weak Dependence Concept

The notation of the weak dependence between two random variables is defined in the following way. Let $X_t \in \mathbf{R}^l$, $t = 0, 1, 2, \ldots$, be a sequences of l-dimensional random variables that are defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, and let $\mathcal{F}_{s,t} = \sigma\{X_u,\ s \le u \le t\}$ be a σ-algebra of events generated by random values of $\{(X_{i_1}, \ldots, X_{i_k}) \in \Gamma\}$, where $s \le i_1 < i_2 < \ldots < i_k \le t$; and Γ is k-dimensional Borelian set. The past of the sequence $\{X_t\}$ is defined by σ-algebra $\mathcal{F}_{0,t}$; its future – by $\mathcal{F}_{t+\tau,\infty}$. The dependence between the two algebras may be measured by a number

$$|\ \mathbf{P}(AB) - \mathbf{P}(A)\mathbf{P}(B)0\ |,$$

where A is an event from $\mathcal{F}_{o,t}$, B is an event from $\mathcal{F}_{t+\tau,\infty}$, $\tau > 0$. If a value τ is increased (i.e. the past and the future moments become more distanted from each other) then for many stationary sequences the value $|\mathbf{P}(AB) - \mathbf{P}(A)\mathbf{P}(B)|$ decreases tending to zero. Such sequencies are called

the weakly depended ones (all definitions and notations introduced above are applicable to the continuous stationary processes also, see [3]).

A specific case of the weak dependence, which will be used in the next Section, is the dependence of the uniformly strong mixing (USM) type:

$$\varphi(\tau) = \sup_{A\in\mathcal{F}_{0,t}, B\in\mathcal{F}_{t+\tau,\infty}} \left| \frac{\mathbf{P}(AB) - \mathbf{P}(A)\mathbf{P}(B)0}{\mathbf{P}(A)} \right| \underset{\tau\to\infty}{\longrightarrow} 0,$$

and $\varphi(\tau)$ is called a coefficient of USM.

Statement (Theorem 17.2.3, [3]). *Let a stationary process X_t be satisfying the USM condition. Let a random variable ξ be measurable on $\mathcal{F}_{o,t}$ and random variable η be measurable on $\mathcal{F}_{t+\tau,\infty}$. If $\mathbf{E}|\xi|^p < \infty$, $\mathbf{E}|\eta|^q < \infty$, $p, q > 1$, $1/p + 1/q = 1$, then*

$$|\mathbf{E}\xi\eta - \mathbf{E}\xi\mathbf{E}\eta| \le 2(\varphi(\tau))^{1/p}\mathbf{E}^{1/p}|\xi|^p\mathbf{E}^{1/q}|\eta|^q.$$

3. Identification

Under conditions of nonparametric a priori uncertainty, equations (2) can not be written down explicitly. Instead, they can be estimated from observations of inputs and outputs. In our case, when part of them are continuous and others are discrete in time, the estimates of equation (2) are proposed to be

$$R_{Tni}(z) = \left[R^{(1)}_{Tni}(z), \quad i \in \mathcal{L}'; \quad R^{(2)}_{Tni}(z), \quad i \in \mathcal{L}'' \right], \tag{3}$$

$$R^{(1)}_{Tni}(z) = a^{(1)}_{Tni}(z) \,/\, \varphi_{Tn}(z), \tag{4}$$

$$R^{(2)}_{Tni}(z) = a^{(2)}_{Tni}(z) \,/\, \varphi_{Tn}(z), \tag{5}$$

where

$$\varphi_{Tn}(z) = \frac{1}{Tnh^{p'}(T)h^{p''}(n)} \sum_{j=1}^{n} \int_0^T K\left(\frac{z - Z(t,j)}{h_{Tn}}\right) dt,$$

$$a^{(1)}_{Tni}(z) = \frac{1}{Tnh^{p'}(T)h^{p''}(n)} \int_0^T Y^{[i]}(t) \sum_{j=1}^{n} K\left(\frac{z - Z(t,j)}{h_{Tn}}\right) dt,$$

$$a^{(2)}_{Tni}(z) = \frac{1}{Tnh^{p'}(T)h^{p''}(n)} \sum_{j=1}^{n} Y^{[i]}_j \int_0^T K\left(\frac{z - Z(t,j)}{h_{Tn}}\right) dt.$$

Let us mention that considered in [4,5] estimates are the specific cases of the above-given estimates (4) and (5), when the inputs and outputs are all continuous either all are discrete.

Now, let $\mathcal{N}(f_1, f_2)$ be a set of all points of continuity of functions $f_1(x)$, $f_2(x)$.

Assumption 1. Let $h(T) > 0$ be a such real function that

$$\lim_{T\to\infty} h(T) = 0$$

Assumption 2. Let $h(T) > 0$ be such that

$$\lim_{T,n\to\infty}\left[h(T) + \frac{1}{Th^{p'}(T)} + \frac{1}{nh^{p''}(n)}\right] = 0.$$

Assumption 3. Let a kernel $K(u)$ be a measurable in Borel sense, and $\int\limits_{\mathbf{R}^p} |K(u)|\,du < \infty$, $\lim\limits_{\|u\|\to\infty} \|u\|^p |K(u)| = 0$.

Let us denote $\psi_i(y) = \dfrac{d}{dy}\mathbf{P}\left\{Y^{[i]}(t) < y\right\}$ which is the density function for the i-th component $Y^{[i]}(t)$ of stationary process $Y(t)$; and $\mathbf{E}_\psi Y^{[i]}$ is the expectation under the true density ψ_i.

Lemma 1 (cf. Lemma 1 [6]). *Let Assumptions* 1 *and* 3 *to be hold. If for any Borel-measured function* $g(\cdot)$ $\mathbf{E}_\psi|g(Y(t))| < \infty$, *then*

$$\lim_{T,n\to\infty}\int\limits_{\mathbf{R}^p} \Phi_i(x - uh(T))K(u)\,du = \Phi_i(x)\int\limits_{\mathbf{R}^p} K(u)\,du,$$

at any point $x \in \mathcal{N}_p(\Phi_i)$ *where* $\Phi_i(x) = \int g(y)f_2(y,x)\,dy$.

P r o o f of this lemma is identical to the proof of the lemma 2.3.4 [7] under the condition (ii), if $h_n = h(T)$, $l = p$ or of the lemma 4 [8] under the condition 9), if $h_n = h(T)$.

Remark 1. Lemma 1 is also valid in the case of replacement of T by n taking its values from the set of pure numbers.

Assumption 4. Let Assumption 3 be hold and, besides that,

$$\int_{\mathbf{R}^p} K(u)\,du = 1, \quad \sup_{u\in\mathbf{R}^p} |K(u)| = L < \infty.$$

Let also $a_i(z) = \int yf_i(y,z)\,dy$.

Lemma 2 (asymptotic unbiasedness). *Let Assumptions* 1–4 *be hold. Then, at each point* $z \in \mathcal{N}(a_i)$

(i) $\lim\limits_{T,n\to\infty} a^{(1)}_{Tni}(z) = a_i(z)$ *if* $\mathbf{E}_\psi |Y^{[i]}(t)| < \infty$;

(ii) $\lim\limits_{T,n\to\infty} a^{(2)}_{Tni}(z) = a_i(z)$ *if* $\mathbf{E}_\psi |Y^{[i]}_1| < \infty$.

P r o o f. Taking into account a measurability of the process $\{Y^{[i]}(t), Z'(t)\}$ and the fact that elements $\{Z''_j\}$ of the sequence are identically distributed, we have from the definition of an expectation

$$\mathbf{E}_f a^{(1)}_{Tni}(z) = \frac{1}{Th^{p'}(T)h^{p''}(T)} \int\limits_0^T dt \int\limits_{\mathbf{R}^{p'}} \int\limits_{\mathbf{R}^{p''}} \int\limits_{\mathbf{R}^1} Y^{[i]}(t) \times$$

$$\times K\left(\frac{z - Z(t,1)}{h_{Tn}}\right) f_i\left(Y^{[i]}(t), Z'(t,1)\right) dY^{[i]}\, dZ''\, dZ'(t) =$$

$$= \frac{1}{T} \int\limits_0^T dt \int\limits_{\mathbf{R}^p} K(u)\, du \int\limits_{\mathbf{R}^1} y f_i(y, z - uh_{Tn})\, dy = \int\limits_{\mathbf{R}^p} a_i(z - uh_{Tn})\, dy. \quad (6)$$

The last integral in (6) is bounded by a value $L\mathbf{E}_\psi |Y^{[i]}(t)| < \infty$, so the Fubini theorem enables us to make in (6): 1) permutation of $\mathbf{E}_\psi$ and integral $\int\limits_0^T$; 2) alternate a succession of integrations $\mathbf{R}^1$, $\mathbf{R}^{p'}$, $\mathbf{R}^{p''}$.

By letting $g(y) \equiv 1$ in Lemma 1 and applying it to the last integral (6), and taking into account the Remark 1, the statement (i) of Lemma 2 is proved; the proof of the statement (ii) can be made similarly.

Assumption 5. Let the stationary measurable process $\{Y^{[i]}(t), Z'(t)\}$ satisfy the USM condition with the USM-coefficient $\varphi_1(\tau)$ and

$$\lim_{T\to\infty} \int\limits_0^T \sqrt{\varphi_1(\tau)}\, d\tau < \infty.$$

Assumption 6. Let the sequence of vectors $\left(Y^{[i]}_j, X''_j\right)$, $j = \overline{1,n}$ satisfy the USM-condition with the coefficient $\varphi_2(\tau)$ and

$$\sum_{\tau\geq 2} \sqrt{\varphi_2(\tau - 1)} < \infty.$$

Let $a_i^2(z) = \int y^2 f_i(y,z)\, dy$.

Lemma 3 (an order of variance). *Let the Assumptions* 2, 4–6 *be hold. If* $\mathbf{E}_\psi \left(Y^{[i]}(t)\right)^2 < \infty$, $i \in \mathcal{L}'$, $\mathbf{E}_\psi \left(Y_1^{[i]}\right)^2 < \infty$, $i \in \mathcal{L}''$ *then for* $j = 1, 2$ *under* $T, n \to \infty$

$$\mathbf{D}_f a_{Tni}^{(j)}(z) \leq$$

$$\leq \frac{4a_i^2(z)}{Th^{p'}(T)nh^{p''}(n)} \int\limits_{\mathbf{R}^p} K^2(u)\,du \int\limits_0^T \varphi_1(\tau)\,d\tau \left[1 + 2\sum_{\mu=2}^{n} \sqrt{\varphi_2(\mu - 1)}\right].$$

P r o o f. Let us introduce the following denotations:

$$Q_j^{[i]}(t) = \left[h^{p'}(T)h^{p''}(n)\right]^{-1} Y^{[i]}(t) K\left[\frac{z - Z(t,j)}{h_{Tn}}\right] -$$

$$-\mathbf{E}_f\left\{\left[h^{p'}(T)h^{p''}(n)\right]^{-1} Y^{[i]}(t) K\left[\frac{z - Z(t,j)}{h_{Tn}}\right]\right\},$$

$$R_{jk,i}(t,s) = \mathbf{E}_f\left[Q_j^{[i]}(t) Q_k^{[i]}(s)\right].$$

According to the definition of variance,

$$\mathbf{D}_f a_{Tni}^{(1)}(z) \leq \frac{1}{T^2n^2} \int\limits_0^T \int\limits_0^T \left[n\,|R_{11,i}(t,s)| + 2\sum_{\mu=2}^{n} n\,|R_{1\mu,i}(t,s)|\right] dt\,ds =$$

$$= \frac{1}{T^2n^2} \int\limits_0^T \int\limits_0^T \left[\sum_{j=1}^{n} R_{jj,i}(t,s) + \sum_{j \neq k} R_{jk,i}(t,s)\right] dt\,ds =$$

$$= \frac{1}{T^2n^2} \int\limits_0^T \int\limits_0^T \left[\sum_{j=1}^{n} R_{jj,i}(t,s) + 2\sum_{\mu=2}^{n} (n - \mu + 1)|R_{1\mu,i}(t,s)|\right] dt\,ds. \quad (7)$$

By virtue of properties of a stationary random process [9] and by the property of being identically distributed for a sequence of random variables,

we have

$$\mathbf{D}_f a^{(1)}_{Tni}(z) \le \frac{1}{T^2 n^2} \int_0^T \int_0^T \left[n|R_{11,i}(t,s)| + 2\sum_{\mu=2}^{n} n|R_{1\mu,i}(t,s)| \right] dt\, ds =$$

$$= \frac{2}{T^2 n} \int_0^T \left[(T-\tau)|R_{11,i}(0,\tau)| + 2\sum_{\mu=2}^{n} (T-\tau)|R_{1\mu,i}(0,\tau)| \right] d\tau \le$$

$$\le \frac{2}{Tn} \int_0^T \left[|R_{11,i}(0,\tau)| + 2\sum_{\mu=2}^{n} |R_{1\mu,i}(0,\tau)| \right] d\tau. \tag{8}$$

Similarly to obtaining of (2.3.7) in [7], it is not difficult to show that

$$\lim_{T,n\to\infty} h^{p'}(T) h^{p''}(n) \mathbf{E}_f \left(Q_j^{[i]}(\tau) \right)^2 =$$

$$= \lim_{T,n\to\infty} h^{p'}(T) h^{p''}(n) \int_{\mathbf{R}^{p+1}} y^2 K^2\left(\frac{z-Z}{h_{Tn}} \right) f_i(y,Z)\, dy\, dz =$$

$$= \lim_{T,n\to\infty} \int_{\mathbf{R}^p} a_i^2(z - t h_{Tn}) K^2(t)\, dt = a_i^2(z) \int_{\mathbf{R}^p} K^2(u)\, du = L\, a_i^2(z). \tag{9}$$

Now, applying the Statement to the last integral in (8), we obtain:

$$\int_0^T \left[|R_{11,i}(0,\tau)| + 2\sum_{\mu=2}^{n} |R_{1\mu,i}(0,\tau)| \right] d\tau \le$$

$$\le 2 \int_0^T \left[\sqrt{\varphi_1(\tau)} \sqrt{\mathbf{E}_f \left(Q_1^{[i]}(0) \right)^2 \mathbf{E}_f \left(Q_1^{[i]}(\tau) \right)^2} + \right.$$

$$\left. + 2\sum_{\mu=2}^{n} \sqrt{\varphi_1(\tau)\varphi_2(\mu-1)} \sqrt{\mathbf{E}_f \left(Q_1^{[i]}(0) \right)^2 \mathbf{E}_f \left(Q_\mu^{[i]}(\tau) \right)^2} \right] d\tau.$$

Then, using (9), under large enough T and n, we have:

$$\mathbf{D}_f a^{(1)}_{Tni}(z) \le \frac{4L\, a_i^2(z)}{Tnh^{p'}(T)h^{p''}(n)} \int_0^T \sqrt{\varphi_1(\tau)}\, d\tau \cdot \left[1 + 2\sum_{\mu=2}^{n} \sqrt{\varphi_2(\mu-1)} \right].$$

An upper bound for the variance $a^{(2)}_{Tni}(z)$ can be found similarly. So, the Lemma is proved.

Remark 2. Assumptions 5 and 6 are hold also for Markovian processes which satisfy:

1) the Doeblin condition [3]:

$$\varphi_1(\tau) \le C_1 e^{-C_2\tau}, \quad \varphi_2(\mu) \le C_3 e^{-C_4\tau}, \quad C_i > 0, \quad i = \overline{1,4};$$

2) the condition G_2 [10] with

$$\varphi_1(\tau) = a_1^\tau, \quad \varphi_2(\mu) = a_2^\mu, \quad 0 < a_i < 1, \quad i = 1,2.$$

Theorem 1 (convergence in mean-square-sense). *Let the Assumptions* 2, 4–6 *be hold. If* $\mathbf{E}_\psi \left(Y^{[i]}(t)\right)^2 < \infty$, $i \in \mathcal{L}'$, $\mathbf{E}_\psi \left(Y_1^{[i]}\right)^2 < \infty$, $i \in \mathcal{L}''$, *then at each point* $z \in \mathcal{N}(a_i, a_i^2)$

$$\lim_{T,n\to\infty} \mathbf{E}_f \left[a_{Tni}^{(j)}(z) - a_i(z)\right]^2 = 0 \qquad j = 1,2.$$

P r o o f. The statement of the Theorem follows from Lemmas 2 and 3, because $\mathbf{E}_\psi \left(Y^{[i]}(t)\right)^2 < \infty$ implies $\mathbf{E}_\psi \left|Y^{[i]}(t)\right| < \infty$ and

$$\mathbf{E}_f \left[a_{Tni}^{(j)} - a_i^{(j)}\right]^2 = \mathbf{D}_f a_{Tni}^{(j)} + \left[\mathbf{E}_f a_{Tni}^{(j)} - a_i^{(j)}\right]^2.$$

As a consequence of Lemmas 2 and 3 and Theorem 1, for a density estimate $\varphi_{Tn}(z)$ it can be formulated

Theorem 2. *Let the Assumptions* 2, 4 *be hold, and Assumptions* 5, 6 *be fulfilled for the process* $Z'(t)$ *and the sequence* Z_j'', $j = 1,2,\ldots$ *Then at each point* $z \in \mathcal{N}(\varphi)$:

1) $\lim\limits_{T,n\to\infty} \mathbf{E}_\varphi \varphi_{Tn}(z) = \varphi(z)$;

2) $\mathbf{D}_\varphi \varphi_{Tn}(z) \le$

$$\le \frac{4L\,\varphi(z)}{Th^{p'}(T)nh^{p''}(n)} \int_0^T \sqrt{\varphi_1(\tau)}\,d\tau \cdot \left[1 + 2\sum_{\mu=2}^{n} \sqrt{\varphi_2(\mu-1)}\right], \quad T, n \to \infty;$$

3) $\lim\limits_{T,n\to\infty} \mathbf{E}_\varphi \left[\varphi_{Tn}(z) - \varphi(z)\right]^2 = 0.$

P r o o f is self-evident, because under $Y^{[i]}(t) \equiv 1$, $Y_j^{[i]} \equiv 1$ we have

$$a_i(z) = a_i^2(z) = \varphi(z), \qquad a_{Tni}^{(k)}(z) = \varphi_{Tn}(z), \qquad k = 1,2.$$

Now let us switch our attention on studying properties of the estimates (4) and (5).

Lemma 4. *For any real* A, B, a, b $(B, b \neq 0)$ *and* $\alpha > 1$ *an inequality holds:*

$$\left|\frac{A}{B} - \frac{a}{b}\right|^{\alpha} \leq 2^{\alpha-1}\left[\left|\frac{A}{B}\right|^{\alpha}\left|\frac{B-b}{b}\right|^{\alpha} + \left|\frac{A-a}{b}\right|^{\alpha}\right]. \tag{10}$$

P r o o f. By elementary calculations with the help of the identity

$$\frac{1}{b} = \frac{1}{B} - \frac{b-B}{bB} \tag{11}$$

one can obtain

$$\left|\frac{A}{B} - \frac{a}{b}\right| \leq \left|\frac{A}{B} - \frac{A}{b}\right| + \left|\frac{A-a}{b}\right| \leq \left|\frac{A}{B}\right|\left|\frac{B-b}{b}\right| + \left|\frac{A-a}{b}\right|.$$

Then, using the triangle inequality [11] or c_r-inequality [12] under $r = \alpha$, we arrive to the statement of the Lemma 4.

Theorem 3. *Let the Theorem* 1 *be hold, and besides that* $K(z) \geq 0$, $\varphi(z) > 0$. *Then*

(i) $\lim\limits_{T,n\to\infty} \mathbf{E}R^{(1)}_{Tni}(z) = R_i(z)$ *if* $\sup_{y\in[0,\infty)} |Y^{[i]}(t)| = C_5 < \infty$;

(ii) $\lim\limits_{T,n\to\infty} \mathbf{E}R^{(2)}_{Tni}(z) = R_i(z)$ *if* $\sup_{j\in 1,2,\ldots} |Y^{[i]}_j| = C_6 < \infty$.

P r o o f. Letting $b = \varphi_{Tn}(z)$, $B = \mathbf{E}_{\varphi}\varphi_{Tn}(z)$ and using the identity (11) analogously to that in [12], we have

$$\mathbf{E}_f R^{(1)}_{Tni}(z) = \frac{\mathbf{E}_f a^{(1)}_{Tni}(z)}{\mathbf{E}_{\varphi}\varphi_{Tn}(z)} + \frac{F^{(1)}_{Tni}(z) + Q^{(1)}_{Tni}(z)}{\mathbf{E}_{\varphi}\varphi_{Tn}(z)^2}, \tag{12}$$

where

$$F^{(1)}_{Tni}(z) = \mathbf{E}_f\left\{\left[a^{(1)}_{Tni}(z) - \mathbf{E}_f a^{(1)}_{Tni}(z)\right]\left[\varphi_{Tn}(z) - \mathbf{E}_{\varphi}\varphi_{Tn}(z)\right]\right\},$$

$$Q^{(1)}_{Tni}(z) = \mathbf{E}_f\left\{a^{(1)}_{Tni}(z)\left[\varphi_{Tn}(z) - \mathbf{E}_{\varphi}\varphi_{Tn}(z)\right]\left[\varphi_{Tn}(z)\right]^{-1}\right\}.$$

Applying the Cauchy–Schwarz inequality and the Lemma 3 we obtain:

$$|F^{(1)}_{Tni}(z)| \leq \left[\mathbf{E}_f[a^{(1)}_{Tni}(z) - \mathbf{E}_f a^{(1)}_{Tni}(z)]^2\right]^{1/2}\left[\mathbf{E}_{\varphi}\left[\varphi_{Tn}(z) - \mathbf{E}_{\varphi}\varphi_{Tn}(z)\right]^2\right]^{1/2} =$$

$$= \left[\mathbf{D}_f a^{(1)}_{Tni}(z)\mathbf{D}_{\varphi}\varphi_{Tn}(z)\right]^{1/2} = O\left(\left(Th^{p'}(T)nh^{p''}(n)\right)^{-1}\right). \tag{13}$$

As it easily can be seen, under conditions of Theorem 3

$$\left|\frac{a_{Tni}^{(1)}(z)}{\varphi_{Tn}(z)}\right| \le \left[\sum_{j=1}^{n}\int_0^T |Y^{[i]}(t)|K\left(\frac{z-Z(t,j)}{h_{Tn}}\right)dt\right] \times$$

$$\times\left[\sum_{j=1}^{n}\int_0^T K\left(\frac{z-Z(t,j)}{h_{Tn}}\right)dt\right]^{-1} \le C_7. \tag{14}$$

Hence

$$|Q_{Tni}^{(1)}(z)| \le \mathbf{E}_f\left\{\left|\frac{a_{Tni}^{(1)}(z)}{\varphi_{Tn}(z)}\right|[\varphi_{Tn}(z)-\mathbf{E}_\varphi\varphi_{Tn}(z)]^2\right\} \le$$

$$\le C_1\mathbf{D}_\varphi\varphi_{Tn}(z) = O((Th^{p'}(T)nh^{p''}(n))^{-1}). \tag{15}$$

From (13), (15) and (12) it follows that

$$\mathbf{E}_f R_{Tni}^{(1)}(z) = \mathbf{E}_f a_{Tni}^{(1)}(z)[\mathbf{E}_\varphi\varphi_{Tn}(z)]^{-1} + O((Th^{p'}(T)nh^{p''}(n))^{-1}). \tag{16}$$

It proves the asymptotic unbiasedness of $R_{Tni}^{(1)}(z)$; the proof of the statemer (ii) can be made similarly.

Theorem 4. *Under conditions of Theorem* 3 *and* 2

$$\mathbf{D}_f R_{Tni}^{(j)}(z) = O((Th^{p'}(T)nh^{p''}(n))^{-1}), \qquad j=1,2.$$

P r o o f. Taking into account (16),

$$\mathbf{D}_f\left[\frac{a_{Tni}^{(j)}(z)}{\varphi_{Tn}(z)}\right] = \mathbf{E}_f\left\{\frac{a_{Tni}^{(j)}(z)}{\varphi_{Tn}(z)} - \mathbf{E}_f\left[\frac{a_{Tni}^{(j)}(z)}{\varphi_{Tn}(z)}\right]\right\}^2 =$$

$$= \mathbf{E}_f\left[\frac{a_{Tni}^{(j)}(z)}{\varphi_{Tn}(z)} - \frac{\mathbf{E}_f a_{Tni}^{(j)}(z)}{\mathbf{E}_\varphi\varphi_{Tn}(z)} + O\left(\frac{1}{Th^{p'}(T)nh^{p''}(n)}\right)\right]^2 \le$$

$$\le 2\left\{\mathbf{E}_f\left[\frac{a_{Tni}^{(j)}(z)}{\varphi_{Tn}(z)} - \frac{\mathbf{E}_f a_{Tni}^{(j)}(z)}{\mathbf{E}_\varphi\varphi_{Tn}(z)}\right]^2 + O\left(\frac{1}{T^2h^{2p'}(T)n^2h^{2p''}(n)}\right)\right\}.$$

Let us put in the inequality (10): $\alpha = 2$, $A = a_{Tni}^{(j)}(z)$, $B = \varphi_{Tn}(z)$, $a = \mathbf{E}_f a_{Tni}^{(j)}(z)$, $b = \mathbf{E}_\varphi\varphi_{Tn}(z)$. Then, with the help of (14), we have

$$\mathbf{E}_f\left[\frac{a_{Tni}^{(j)}(z)}{\varphi_{Tn}(z)} - \frac{\mathbf{E}_f a_{Tni}^{(j)}(z)}{\mathbf{E}_\varphi\varphi_{Tn}(z)}\right]^2 \le$$

$$\leq \frac{2}{\mathbf{E}_{\varphi}\varphi_{Tn}(z)}\left[C_j\mathbf{D}_{\varphi}\varphi_{Tn}(z)+\mathbf{D}_f a_{Tni}^{(j)}(z)\right], \qquad j=1,2.$$

The theorem is proved.

As a consequence of Theorems 3 and 4 it can be formulated

Theorem 5. *Let the conditions of Theorems* 3 *and* 4 *be hold. Then*

$$\lim_{T,n\to\infty}\mathbf{E}_f[R_{Tni}^{(j)}(z)-R_i(z)]^2=0, \qquad j=1,2.$$

4. Control Problems

Let the s' continuous and s'' discrete outputs y of a stochastic object should by the time $q > T$ be leveled to given values $\overset{\circ}{Y}{}'(q)$ and $\overset{\circ}{Y}{}''_{n+1}$ with the help of m controlled inputs u (m' of which are continuous and $m''=m-m'$ are discrete) and the rest $r=p-m$ inputs x (r' - continuous, $r''=r-r'$ - discrete) be only observed. For instance, let outputs correspond to some qualities of produced goods and inputs – to certain parameters of technology of production. There is a problem of finding of $\left(\overset{\circ}{U}{}'(q),\overset{\circ}{U}{}''_{n+1}\right)$.

Let the equations describing this object (as in case (1)) be

$$y^T=F^T(x,u), \tag{17}$$

and the controls u depend on inputs and outputs as

$$u^T=\Phi^T(x,y), \tag{18}$$

where $F=(F_1,\ldots,F_s)$, $\Phi=(\Phi_1,\ldots,\Phi_m)$; F_i and Φ_j, $i=\overline{1,s}$, $j=\overline{1,m}$ are one-to-one and continuous functions on their supports.

Taking into account the formula (18), the above mentioned properties of the functions F_i, and Φ_j, one can reformulate the problem for the object (17) into the equivalent problem for the object (18). Then the values of controlled inputs $\overset{\circ}{u}$ resulting in wanted outputs $\overset{\circ}{y}$ (as it follows from (18)) are defined as

$$\overset{\circ}{u}{}^T=\Phi^T(x,\overset{\circ}{y}). \tag{19}$$

As an approximation of $\left(\overset{\circ}{U}{}'(q),\overset{\circ}{U}{}''_{n+1}\right)$ let us take a system of m regression equations minimizing of $\left(\overset{\circ}{U}{}'(q),\overset{\circ}{U}{}''_{n+1}\right)$ and

$$\widetilde{R}_i\left(X(q,n+1),\quad \overset{\circ}{Y}(q,n+1)\right)=\widetilde{R}_i(*):$$

$$\widetilde{R}_i(*) = \left\{ \widetilde{R}_i^{(1)}(*), \quad i \in \mathcal{R}', \quad \widetilde{R}_i^{(2)}(*), \quad i \in \mathcal{R}'' \right\},$$

where

$$\mathcal{R}' = \{1, \ldots, r'\}, \qquad \mathcal{R}'' = \{r'+1, \ldots, r\},$$

$$\widetilde{R}_i^{(1)}(*) = \mathbf{E}\{U^{[i]}(t)/*\}, \tag{20}$$

$$\widetilde{R}_i^{(2)}(*) = \mathbf{E}\{U^{[i]}/*\}. \tag{21}$$

Analogously to (4) and (5), one can, from realizations of $U'(t)$, $X'(t)$, $Y'(t)$, $t \in [0,T]$, and from sequence (U_i'', X_i'', Y_i''), $i = \overline{1,n}$, using values $X'(q)$, X_{n+1}'', $\overset{\circ}{Y}{}'(q)$, $\overset{\circ}{Y}{}''_{n+1}$ construct estimates for (20) and (21):

$$[\widetilde{R}_{Tni}^{(1)}(*) = \int_0^T U^{[i]}(t) \sum_{j=1}^n K(\xi)K(\eta)\,dt \left[\int_0^T \sum_{j=1}^n K(\xi)K(\eta)\,dt \right]^{-1}, \tag{22}$$

$$\widetilde{R}_{Tni}^{(2)}(*) = \left[\sum_{j=1}^n U_j^{[i]}(t) \int_0^T K(\xi)K(\eta)\,dt \right] \left[\sum_{j=1}^n \int_0^T K(\xi)K(\eta)\,dt \right]^{-1}, \tag{23}$$

where

$$\xi = [X(q,n+1) - X(t,j)]h_{Tn}^{-1}, \qquad \eta = [\overset{\circ}{Y}(q,n+1) - Y(q,j)]h_{Tn}^{-1}.$$

The properties of estimates (22)–(23) are similar to these of (4)–(5).

It can be mentioned in conclusion that the algorithms (4), (5), (22), (23) will demand for their realization the use of digital and analog computers.

5. A Conclusion

In the paper the stochastic objects are considered, part of inputs and outputs of which are continuous stationary stochastic processes and the other part of inputs and outputs are weakly dependent discrete random sequences. For solving the problems of identification and control of such objects a class of algorithms is suggested and studied. The algorithms (4), (5), (22), (23), which are nonparametric, converge in a mean-square-sense. Their implementation may be fulfilled with combined digital and analog computers.

One can find the main parts of asymptotic MSE of the proposed estimates by using the methods presented in [14–21].

REFERENCES

1. *Koshkin G.M. and Tarasenko F.P.* Nonparametric algorithms for identifying and control of continuous-discrete stochastic objects // 8th IFAC-IFORS Symposium on Identification and System Parameter Estimation. China, Beijing: Pergamon Press, 1988. Vol. 2. P. 822-827.
2. *Pugachev V.S.* Random Functions Theory and Its Application to the Automatic Control Problems. Moscow: Fizmatgiz, 1962 (in Russian).
3. *Ibragimov I.A. and Linnik Yu.V.* Independent and Stationary Sequences of Random Variables. Moscow: Nauka, 1965 (in Russian).
4. *Koshkin G.M. and Tarasenko F.P.* Recursive estimation of probability density function and regression curve for dependent sample // Math. Statist. and its Applications. Tomsk: Tomsk University Publishing House, 1976. No. 4. P. 122-138 (in Russian).
5. *Koshkin G.M.* Asymptotic properties of nonparametric algorithms of identifying and control for continuous time random processes // Proceedings of 5th All-Union Conference on Nonparametric and Robust Methods in Statistics and Cybernetics. 1985. The paper is deposited at VINITI, No. 5259–85, 1985 (in Russian).
6. *Greblicki W., and Krzyzak A.* Asymptotic properties of kernel estimates of a regression function // J. Statist. Plan. and. Infer. 1980. No. 4. P. 81-90.
7. *Dobrovidov A.V. and Koshkin G.M.* Nonparametric signal estimation. Moscow: Nauka, Fizmatlit, 1997 (in Russian).
8. *Koshkin G.M.* Asymptotic properties of functions of statistics and their application to nonparametric estimation // Autom. Remote Contr. 1990. Vol. 51, No. 3. P. 345-357.
9. *Sweshnikov A.A.* Applied Methods of Random Functions Theory. Moscow: Nauka, 1968 (in Russian).
10. *Rosenblatt M.* Density estimates and Markov sequences // Nonparametric Techniques in Statistical Inference. Cambridge: Univ. Press, 1970. P. 199-210.
11. *Neveu J.* Mathematical Foundations of the Probability Calculus. Moscow: Mir, 1969 (in Russian).
12. *Loeve M.* Probability theory. Princeton, New Jersey, Toronto, New York, London: D. Van Nostrand Company Inc., 1960.
13. *Nadaraya E.A.* Nonparametric estimates of regression curve // Proceeding of the Computer Centre of the Academy of Sciences Gruz. SSR, Tbilisi, 1965. P. 56-68 (in Russian).
14. *Vasil'iev V.A. and Koshkin G.M.* Estimation of functions of a distribution density from dependent observations // Probl. Inf. Transm. 1997. Vol. 33, No. 4. P. 324-338.
15. *Koshkin G.M.* Deviation moments of the substitution estimator and its piecewise smooth approximations // Sib. Math. J. 1999. Vol. 40, No. 3. P. 515-527.

16. *Kitayeva A.V., Koshkin G.M. and Ryumkin V.I.* Nonparametric estimation of conditional functionals from dependent observations // Prague Stochastics'98: Proceedings of the Joint Session of Prague Symposium on Asymptotic Statistics and Prague Conference on Information Theory, Statistical Decision Functions and Random Processes (August 23 to 28, 1998, Prague). Vol. 1 / Ed. Marie Huskova, Petr Lachout, Jan Amos Visek. Prague: Union of Czech Mathematicians and Physicists, 1998. P. 295-300.
17. *Vasil'iev V.A. and Koshkin G.M.* Nonparametric identification of the autoregression // Teor. Veroyatn. Primen. 1998. Vol. 43, No. 3. P. 577-588 (in Russian).
18. *Dobrovidov A.V. and Koshkin G.M.* Nonparametric estimation of logarithmic derivative of a density from uniform strong mixing sequences // Avtom. Telemekh. 2001. Vol. 9. P. 63-88 (in Russian).
19. *Koshkin G.M. and Vasil'iev V.A.* Nonparametric estimation of derivatives of a multivariate density from dependent observations // Mathematical Methods of Statistics. New York: Allerton Press, Inc., 1998. Vol. 7, No. 4. P. 361-400.
20. *Vasil'iev V.A. and Koshkin G.M.* Nonparametric estimation of the ratios derivatives of a multivariate density from dependent observations // Sib. Math. J. 2000. Vol. 41, No. 2. P. 229-245.
21. *Kitaeva A.V., Koshkin G.M., Piven I.G. and Ryumkin V.I.* On nonparametric kernel identification of nonlinear autoregression process // The 5th Korea-Russian International Symposium on Science and Technology. (June 26 – July 3, 2001, Tomsk, Russia). Proceedings KORUS 2001. Tomsk: Tomsk Polytechnic University. No. 2. P. 208-211.

А.П. Красноштанов

Непараметрическая идентификация многосвязных систем при неполной информации

В настоящей работе предметом рассмотрения являются объекты, относящиеся к классу сложных взаимосвязанных статических систем.

1. Введение

В последние годы наметилась тенденция построения систем управлени различными сложными системами в реальных условиях, когда уровень априорной информации об объекте неполный, т.е. когда структура системы, величины помех и их статистические характеристики не известны [1]. В таких случаях для успешного управления необходимо построение моделей, описывающих процессы, происходящие в исследуемом объекте, и включение их в систему управления. Такие системы управления

получили название систем управления с идентификатором. При постановке задачи построения модели должны учитываться те технологические параметры или, иначе, те входные, выходные или приведенные к ним переменные, которые доступны измерению с помощью существующих средств контроля.

В настоящей работе предметом рассмотрения являются объекты, относящиеся к классу сложных взаимосвязанных статических систем. Примером таких объектов могут служить предприятия металлургической промышленности, стройиндустрии, энергетики и др. [2], а также производственно-экономические системы, имеющие дискретно-непрерывный характер технологического процесса и установившийся режим эксплуатации. Основные трудности, с которыми приходится сталкиваться при построении моделей подобного рода объектов традиционными методами идентификации, состоят не только в сложности определения структуры модели из-за ее большой размерности, но и во взаимосвязи переменных, а также из-за присутствия неконтролируемых факторов на самом объекте и в каналах связи. В зависимости от конкретных условий можно говорить о моделях с различной степенью детализации – на уровне технологических аппаратов и их групп, цехов, предприятий и их комплексов, которые представляют из себя отдельные модули. Объединяя эти модули в группы, можно строить модели и на их основе контуры управления более высокого иерархического уровня, вплоть до модели и контура управления производственного комплекса в целом.

Объекты исследования в данной работе представляют собой процессы, имеющие непрерывный характер и сложную, жесткую структуру взаимных технологических связей. Существующая структурная схема технологических связей между переменными самого объекта или его функциональных частей приводит к понятию многосвязности переменных, когда часть выходных переменных одних объектов является входными для других в технологической цепочке. Это обстоятельство, в свою очередь, приводит к тому, что исследуемый класс объектов описывается не в традиционной форме "вход – оператор связи – выход", а неявно, в виде некоторой системы уравнений, определяющих соответствующие неявные функции. Информационная обстановка характеризуется тем, что часть этих уравнений может быть задана точно. Часть может быть известна с точностью до набора параметров [3]. А некоторые уравнения из-за недостатка информации могут быть представлены лишь качественно (к примеру с точностью до переменных) соотношениями неизвестной структуры и выборкой наблюдений вектора состояний "вход–выход". Иными словами, имеет место ситуация, когда априорная

информация об исследуемом объекте имеет некий смешанный характер и может одновременно принадлежать к нескольким общепринятым уровням. Именно такой ситуации обычно соответствуют реальные условия функционирования большинства промышленных объектов. Таким образом, сама постановка задачи идентификации или управления осуществляется в форме, когда ее нельзя отнести ни к детерминированной, ни к байесовской, ни к параметрической или непараметрической.

2. Классификация многосвязных объектов, постановка задач

Рассмотрим задачу построения математических моделей производственных комплексов с непрерывным характером технологического процесса. Сложность этих объектов проявляется не только в большой размерности вектора переменных, но и в присутствии обратных и перекрестных потоков в технологической цепочке, что предопределяет зависимость некоторых выходных переменных не только от входа объекта, но и от некоторых других выходных переменных. Многосвязным объектом будем называть объект, который описывается некоторой системой неявных функций от входных и выходных переменных.

Проведем соответствующую классификацию исследуемых многосвязных объектов в зависимости от информационных условий. Примем к рассмотрению следующий набор уровней априорной информации [2]:

1. Частичная информация о модели процесса дана полностью. Иными словами, точно известна только часть модели (неполная модель). Недостаток информации восполняется только тем, что имеется выборка статистически независимых наблюдений вектора состояний объекта конечного объема, снятая со случайными ошибками с неизвестным законом распределения, но с нулевым первым и ограниченным вторым моментами;

2. Структура неполной модели известна с точностью до набора параметров. Имеется выборка независимых наблюдений вектора состояний, аналогичная п. 1;

3. Неизвестна параметризованная структура неполной модели, но известны некоторые качественные свойства объекта (например, объект статический). Имеется выборка независимых наблюдений вектора состояний, аналогичная п. 1;

4. Уровень различных комбинаций из первых трех уровней, приводящих к возникновению как незамкнутых, так и замкнутых относительно выхода многосвязных систем.

Введем обозначения: $X = (x_1, \ldots, x_k) \in \Omega(X) \subset R^k$ – вектор входных переменных объекта, $Y = (y_1, \ldots, y_L) \in \Omega(Y) \subset R^L$ – вектор выходных переменных, $\{X[t], Y[t]\}$, $t = \overline{1,N}$, – выборка статистически независимых наблюдений вектора состояний объекта $\{X, Y\}$ в дискретные моменты времени $t = \overline{1,N}$.

Рассмотрим более общий комбинированный уровень априорной информации в нашей классификации. Внутренние связи между X и Y на объекте в соответствии с многосвязной природой могут быть представлены следующей системой соотношений:

$$F_j\left(X^{(j)}, Y^{(j)}, \alpha_{1j}, \ldots, \alpha_{pj}\right) = 0,\ j = \overline{1,m}; \qquad (1)$$
$$m \le Ly_j - \varphi_j\left(X^{(j)}, Y^{(j)}\right) = 0,\ j = \overline{m+1,l}.$$

Здесь $F_j(\ldots)$, $j = \overline{1,m}$, – функции, известные с точностью до набора параметров $\alpha^{(j)} = (\alpha_{1j}, \ldots, \alpha_{pj})$, $\varphi_j(\ldots)$, $j = \overline{m+1,l}$, – функции неизвестного вида; $X^{(j)}$, $Y^{(j)}$ – векторы, составленные из компонент векторов X, Y, входящих в j-е уравнение.

Сформулируем основную задачу. Имея выборку наблюдений $\{X[t], Y[t]\}$, $t = \overline{1,N}$, требуется найти такое значение выходной величины $Y \in \Omega(Y)$ на объектах типа (1), которое соответствует заданному входному воздействию $X = \widetilde{X} \in \Omega(X)$.

3. Идентификация многосвязных объектов

Процесс моделирования данных многосвязных объектов обычно состоит из двух этапов. Первый завершается построением некоторой алгебраической системы уравнений:

$$F_N(X, Y, \alpha_N) = F_{Nj}\left(X^{(j)}, Y^{(j)}, \alpha_N^{(j)}\right),\quad j = \overline{1,m}, \qquad (2)$$
$$m \le Ly_j - \varphi_{Nj}\left(X^{(j)}, Y^{(j)}\right),\ j = \overline{m+1,L}.$$

В приведенных описаниях приняты обозначения: $F_j(\ldots)$ – вектор-функция; $\alpha_N^{(j)}$ – оценки векторов параметров $\alpha^{(j)}$ в параметризованных структурах при $j = \overline{1,m}$ (если таковые имеются) и непараметрические оценки "качественных" зависимостей при $j = \overline{m+1,L}$;

$$\varphi_{NJ}(X^j, Y^j) = \frac{\sum\limits_{t=1}^{N} y_j[t]\Phi(X^j - X^j[t], C, N)\Phi(Y^j - Y^j[t], C, N)}{\sum\limits_{t=1}^{N} \Phi(X^j - X^j[t], C, N)}. \qquad (3)$$

Ядра $\Phi(Z)$ в (3) – симметричные неотрицательные, имеющие свойства плотности на отрезке $\{Z : |Z| \le C\}$, дельтообразные при $N \to \infty$ функции; $C = c_N$ – параметр (коэффициент размытости), удовлетворяющий условиям [1]:

$$c_N > 0,\ N = 1, 2, \ldots;\ \lim_{N\to\infty} c_N = 0;\ \lim_{N\to\infty} (N c_N) = \infty.$$

При фиксированном объеме статистических данных $\{X[t], Y[t]\}$, $t = \overline{1, N}$, качество аппроксимации стохастических зависимостей с помощью непараметрической оценки кривой регрессии зависит от выбранных ядерных функций и значений коэффициентов размытости, определяющих полуширину интервала финитности ядра. При формировании статистик (3), например, применяются ядерные функции Епанечникова для скалярного аргумента $u = c_n^{-1}(x - x[t])$. В случае векторного аргумента обычно ядро является мультипликативной функцией своих скалярных составляющих. Определение конкретных значений параметров размытости c_N осуществляется из условия минимума критерия:

$$R(c_N) = N^{-1} \sum_{j=1}^{N} \left(y_i[j] - \varphi_{(Ni)}(X[j], Y[j], c_N)\right)^2,$$

в котором вычисление каждой j-й разности на основе статистики (3) проводится в скользящем режиме по выборке $\{X[t], Y[t]\}$, $t = \overline{1, N}$, исключая j-ю ситуацию.

4. Оценка параметров в параметризованных моделях

Предположим, что информация будет поступать последовательно, то есть в каждый момент времени t к нам поступает только одна пара значений входа и выхода $\{X[t], Y[t]\}$, тогда в качестве вычислительной процедуры оценки параметров будем использовать модификацию вероятностных итеративных алгоритмов – процедуру Кифера–Вольфовица:

$$\alpha_{j1}^{(i)} = \alpha_{j0}^{(i)} - \gamma_{j1}^{(i)} F^i\left(X^{(i)}[1] Y^{(i)}[1], \alpha_0^{(i)}\right),\ j = \overline{1, p_i},\ i = \overline{1, m},$$

где коэффициенты $\gamma_{j1}^{(i)}$, определяющие длину шага, на данном шаге итерации и на всех последующих шагах удовлетворяют известным условиям Роббинса–Монро. Далее берем вторую пару значений входа и выхода и так далее. На s-м шаге процедура для оценки параметров будет выглядеть так [3]:

$$\alpha_{js}^{(i)} = \alpha_{js-1}^{(i)} - \gamma_{js}^{(i)} F^i\left(X^{(i)}[s] Y^{(i)}[s], \alpha_{s-1}^{(i)}\right),\ j = \overline{1, p_i},\ i = \overline{1, m}.$$

На втором этапе идентификации осуществляется решение систем (3) при заданном входном воздействии $X = \widetilde{X}$. Алгоритмы прогноза для замкнутых и незамкнутых вариантов модельных систем несколько различаются. Выделим эти случаи.

5. Прогноз выхода замкнутых многосвязных систем

В качестве оценки решения замкнутых алгебраических систем нелинейных уравнений типа (3) принимается статистика

$$y_{jN} = \frac{\sum\limits_{t=1}^{N} y_j[t] \prod\limits_{j=1}^{L} \Phi(0 - \varepsilon_j[t], c_{2N})}{\sum\limits_{t=1}^{N} \prod\limits_{j=1}^{L} \Phi(0 - \varepsilon_j[t], c_{2N})}, \quad j = \overline{1,L}, \tag{4}$$

где $\varepsilon_j[t]$, $j = \overline{1,L}$, $t = \overline{1,N}$, – некоторая рабочая выборка невязок системы (3), специальным образом сгенерированная на основе исходной выборки "вход–выход", а именно:

$$\varepsilon_j[t] = F_{jN}(\widetilde{X}, Y[t]), \quad j = \overline{1,L},$$

где $F_{jN}(\cdot)$, $j = \overline{1,L}$, – компоненты соответствующей вектор-функции. Ядра $\Phi(\ldots)$ в (4) могут быть аналогичны тем, которые используются в оценке (3).

6. Прогноз выхода незамкнутых многосвязных систем

Если первый этап собственно построения модельной алгебраической системы, включающий в себя или оценку параметров, или оценку неизвестных функциональных зависимостей, не отличается от подобных действий для замкнутых многосвязных систем, то осуществление прогноза выхода имеет свои особенности. В условиях наблюдаемости объекта ситуация незамкнутости его модели в нашей классификации соответствует следующим информационным комбинациям:

- наличию детерминированной части модели и отсутствию знаний о другой ее части, восполненному только выборкой наблюдений;

- наличию параметризованной части модели и незнанию структуры другой ее части, восполненному только выборкой наблюдений;

– наличию непараметризованной части модели и незнанию даже качественных связей в другой ее части, восполненному только выборкой наблюдений.

Незамкнутость любой алгебраической системы относительно своих неизвестных приводит к тому, что система имеет бесконечное множество решений. Это происходит из-за наличия так называемых свободных переменных. Однако в связи с тем, что присутствуют наблюдения вектора состояний объекта, множество решений может быть "локализовано". Кроме того, если мы будем считать модели систем (1) некоторыми ограничениями типа равенств, то в этом случае можно оценить регрессию выхода при условиях, что вход $X = \widetilde{X}$ и выполняются все ограничения (уравнения модели). Последнее означает, что сгенерированная выборка невязок

$$\varepsilon_j[t] = F_N(\widetilde{X}, Y[t]), \quad j = \overline{1,m}, \quad t = \overline{1,N},$$

где $F_N(\ldots)$ обозначает вектор-функцию модельной системы объекта (1), должна захватывать точку ноль, что и соответствует выполнению ограничений равенства при $X = \widetilde{X}$. Таким образом, непараметрические оценки прогноза выхода для незамкнутых систем, являющихся соответствующими моделями, можно записать в виде статистики

$$y_{jN} = \frac{\sum\limits_{t=1}^{N} y_j[t] \prod\limits_{j=1}^{k} \Phi(\widetilde{x}_j - x_j[t], c_{1N}) \prod\limits_{j=1}^{L} \Phi(0 - \varepsilon_j[t], c_{2N})}{\sum\limits_{t=1}^{N} \prod\limits_{j=1}^{k} \Phi(\widetilde{x}_j - x_j[t], c_{1N}) \prod\limits_{j=1}^{L} \Phi(0 - \varepsilon_j[t], c_{2N})}, \; j = \overline{1,L}. \qquad (5)$$

Обратим внимание на то, что за незнание части модельной системы нам приходится платить увеличением размерности оценки (5) по сравнению с (4). Алгоритмы (4) и (5) допускают различные модификации, связанные с различными способами генерации рабочих выборок невязок.

СПИСОК ЛИТЕРАТУРЫ

1. *Медведев А.В.* Непараметрические системы адаптации. Новосибирск: Наука, 1983. 174 с.
2. *Красноштанов А.П.* Комбинированные многосвязные системы. Новосибирск: Наука, 2001. 176 с.
3. *Цыпкин Я.З.* Адаптация и обучение в автоматических системах. М.: Наука, 1968. 399 с.

Б.Ю. Лемешко, С.Н. Постовалов

Рекомендации ГОСТ Р 50.1-033-2001 и ГОСТ Р 50.1-037-2002 как итог исследований вопросов применения критериев согласия [1]

1. Введение

С 1.07.2002 Госстандарт России взамен устаревших нормативных документов вводит в действие новые рекомендации по прикладной статистике "Правила проверки согласия опытного распределения с теоретическим". ГОСТ Р 50.1-033-2001 регламентирует применение критериев типа χ^2 [1], ГОСТ Р 50.1-037-2002 – применение непараметрических критериев [2]. Рекомендации предназначены для использования в качестве руководства при статистической обработке результатов наблюдений в различных приложениях: анализе данных физических и технологических экспериментов, обработке результатов измерений, задачах исследования надежности и контроля качества.

Необходимость разработки рекомендаций вызвана двумя основными факторами. Во-первых, в нормативных документах практически не оговорены правила применения критериев согласия при проверке сложных статистических гипотез. А именно такие ситуации оказываются наиболее реальными в приложениях. Поэтому практика использования критериев согласия в задачах контроля качества, исследования надежности и в других приложениях зачастую приводит к их некорректному применению и, как следствие, неверным выводам. Во-вторых, в наиболее доступной литературе учебного характера, к сожалению, часто содержатся принципиальные ошибки и неверные рекомендации, приводящие к некорректному или неоптимальному применению критериев и увеличивающие вероятность неверных выводов. Не решены данные вопросы и в популярных программных системах статистического анализа. В основу рекомендаций были положены результаты, полученные нами при исследовании свойств критериев типа χ^2 и непараметрических критериев типа Колмогорова, типа ω^2 и Ω^2 Мизеса.

Основными методами проводимых исследований, наряду с математическим аппаратом, явились развиваемая методика компьютерного моделирования и исследования статистических закономерностей. Благодаря этой методике удалось получить достаточно точные решения для задач, которые не могли быть получены аналитическими средствами.

[1] Выполнена при поддержке Российского фонда фундаментальных исследований (проект є 00-01-00913)

2. Вопросы применения критериев типа χ^2

Проверка статистических гипотез о согласии эмпирических данных с теоретическим законом распределения с применением критериев согласия типа χ^2 обусловлена рядом условий, которые обеспечивают корректное решение задачи. К сожалению, не в каждом источнике, который используется в качестве руководства исследователем, находят отражение эти условия. Вследствие этого, несмотря на кажущуюся простоту, практика использования критериев согласия типа χ^2 изобилует примерами его некорректного или неэффективного применения, особенно при проверке сложных гипотез.

Анализ примеров "неудачного" применения критериев типа χ^2 позволяет выделить две группы причин, которые на практике могут приводить к неверным статистическим выводам. Во-первых, это часто совершаемые принципиальные ошибки, при которых использование в качестве предельного χ^2_{k-m-1}-распределения оказывается неправомерным. Такое происходит, когда, проверяя сложные гипотезы, применяют оценки, в основе которых лежат исходные негруппированные данные. Во-вторых, действия, использующие возможности критерия не наилучшим образом (например нерациональный выбор числа интервалов и способа разбиения выборки на интервалы). В первом случае возрастает вероятность ошибки первого рода α (отклонить верную проверяемую гипотезу), во втором – вероятность ошибки второго рода β (принять проверяемую гипотезу при справедливости альтернативы).

При использовании критериев согласия типа χ^2 неоднозначность при построении и вычислении статистик бывает связана с выбором числа интервалов и тем, каким образом область определения случайной величины разбивается на интервалы. Естественно, что такой произвол отражается на статистических свойствах применяемых критериев, в частности, на их мощности при различении близких конкурирующих гипотез. Очевидно, что выбор числа интервалов и способа разбиения на интервалы следует осуществлять с позиций обеспечения максимальной мощности критерия.

С использованием критериев согласия могут проверяться простые гипотезы вида H_0:$F(x) = F_0(x,\theta)$, где $F_0(x,\theta)$ – функция распределения вероятностей, с которой проверяется согласие наблюдаемой выборки независимых одинаково распределенных величин $X_1, X_2, ..., X_n$, а θ – известное значение параметра (скалярного или векторного), и сложные гипотезы H_0: $F(x) \in \{F_0(x,\theta), \theta \in \Theta\}$, где Θ – пространство параметров. В процессе проверки сложной гипотезы оценка параметра $\hat{\theta}$ вычисля-

ется по этой же самой выборке. Если оценка $\hat{\theta}$ вычислена по другой выборке, то гипотеза простая.

При использовании критериев согласия типа χ^2 область определения случайной величины разбивается на n интервалов граничными точками:

$$x_1 < x_2 < \ldots < x_{k-1} < x_k.$$

Статистика X_n^2 Пирсона вычисляется в соответствии с соотношением

$$X_n^2 = n \sum_{i=1}^{k} \frac{(n_i/n - P_i(\theta))^2}{P_i(\theta)}, \qquad (1)$$

где n_i – количество наблюдений, попавших в i-й интервал, $P_i(\theta) = \int\limits_{x_{i-1}}^{x_i} f_0(x,\theta)dx$ – вероятность попадания наблюдения в i-й интервал, $n = \sum\limits_{i=1}^{k} n_i$, $\sum\limits_{i=1}^{k} P_i(\theta) = 1$. При справедливой простой гипотезе H_0 предельное распределение статистики $G(X_n^2|H_0)$ есть χ^2-распределение с числом степеней свободы $k-1$. Если по выборке оценивалось m параметров закона в результате минимизации статистики X_n^2, статистика подчиняется χ^2-распределению с $k-m-1$ степенями свободы. При справедливой альтернативной гипотезе H_1 предельное распределение $G(X_n^2|H_1)$ представляет собой нецентральное χ^2-распределение с тем же числом степеней свободы и параметром нецентральности

$$s(\theta) = \sum_{i=1}^{k} \frac{c_i^2(\theta)}{P_i}, \qquad (2)$$

где $c_i(\theta) = \sqrt{n} \int\limits_{x_{i-1}(\theta)}^{x_i(\theta)} (f_1(x,\theta) - f_0(x,\theta))dx$ и $f_1(x,\theta)$ соответствуют альтернативе.

В случае проверки сложных гипотез и оценивания по выборке параметров наблюдаемого закона использование в качестве предельных χ^2_{k-m-1}-распределений справедливо лишь при определении оценок минимизацией статистики X_n^2 или при вычислении по сгруппированным данным ряда оценок, в том числе оценок максимального правдоподобия (ОМП). Распределения статистик χ^2 Пирсона и отношения правдоподобия при оценивании параметров наблюдаемого закона по точечным наблюдениям и различных способах построения интервалов нами были исследованы методами компьютерного моделирования. Было показано,

что распределения $G(X_n^2|H_0)$ статистики оказываются зависящими от того, как строятся интервалы, и хорошо аппроксимируются семейством гамма-распределений.

Предложенная в работах С.М. Никулина [3, 4] статистика типа χ^2 отличается от X_n^2 только при сложных гипотезах. Предельное распределение этой статистики есть распределение χ^2_{k-1} (количество степеней свободы не зависит от числа оцениваемых параметров). Неизвестные параметры распределения $F(x,\theta)$ в этом случае должны оцениваться по негруппированным данным методом максимального правдоподобия. При этом вектор вероятностей попадания в интервал $P = (P_1, P_2, ..., P_k)^T$ предполагается заданным, и граничные точки интервалов определяются соотношениями

$$x_i(\theta) = F^{-1}(P_1 + P_2 + ... + P_i), \quad i = \overline{1, k-1}.$$

Данная статистика имеет вид

$$Y_n^2(\theta) = X_n^2 + n^{-1} a^T(\theta)\Lambda(\theta)a(\theta), \tag{3}$$

где X_n^2 вычисляются в соответствии с (1). Элементы и размерность матрицы

$$\Lambda(\theta) = \left[J(\theta_l, \theta_j) - \sum_{i=1}^{k} \frac{w_{\theta_l i} w_{\theta_j i}}{P_i} \right]^{-1}_{m \times m}$$

определяются оцениваемыми компонентами вектора параметров θ; $J(\theta_l, \theta$ – элементы информационной матрицы

$$J(\theta) = \left[\int_{-\infty}^{\infty} \left(\frac{\partial \ln f_0(x,\theta)}{\partial \theta_l} \frac{\partial \ln f_0(x,\theta)}{\partial \theta_j} \right) f_0(x,\theta) dx \right]_{m \times m};$$

$a(\theta_l) = w_{\theta_l 1} n_1/P_1 + ... + w_{\theta_l k} n_k/P_k$ - элементы вектора $a(\theta)$; величины $w_{\theta_l i}$ определяются соотношением

$$w_{\theta_l i} = -f_0[x_i(\theta), \theta] \frac{\partial x_i(\theta)}{\partial \theta_l} + f_0[x_{i-1}(\theta), \theta] \frac{\partial x_{i-1}(\theta)}{\partial \theta_l}.$$

При верной конкурирующей гипотезе статистика Y_n^2 имеет в качестве предельного $G(Y_n^2|H_1)$ нецентральное χ^2_{k-1}-распределение с параметром нецентральности

$$s(\theta) = \sum_{i=1}^{k} \frac{c_i^2(\theta)}{P_i} + d^T(\theta)\Lambda(\theta)d(\theta), \tag{4}$$

где элементы вектора $d(\theta)$ определяются как

$$d(\theta_l) = \frac{w_{\theta_l 1} c_1(\theta)}{P_1} + \ldots + \frac{w_{\theta_l k} c_k(\theta)}{P_k}.$$

Зависимость мощности от способа группирования. Способ группирования особенно сильное влияние оказывает на предельное распределение $G(X_n^2|H_1)$. Было показано, что критерии согласия χ^2 Пирсона и отношения правдоподобия при проверке как простых, так и сложных гипотез имеют максимальную мощность против близких альтернатив, если использовать такое разбиение области определения случайной величины на интервалы, при котором потери в информации Фишера о параметрах закона, соответствующего гипотезе H_0, минимальны (асимптотически оптимальное группирование). Чем меньше потери в информации Фишера, связанные с группированием данных, тем больше параметр нецентральности, определяемый соотношением (2). В [1] для конкретных законов распределения представлен достаточно широкий состав (более 50 для примерно 20 законов распределений) построенных таблиц асимптотически оптимального группирования, минимизирующего потери в информации Фишера. Использование асимптотически оптимального группирования при заданном числе интервалов обеспечивает максимальную мощность при близких гипотезах. Причем выигрыш в мощности даже по отношению к разбиению на интервалы равной вероятности очень существенный.

Исследование распределений статистики Y_n^2 Никулина, которая отличается от X_n^2 только при сложных гипотезах, показало, что как $G(Y_n^2|H_0)$, так и $G(Y_n^2|H_1)$ несущественно зависят от способа группирования. Более того, наши исследования методами статистического моделирования показали, что с позиций наибольшей мощности разбиение на интервалы равной вероятности (равновероятное группирование) оказывается наиболее предпочтительным. Подчеркнем, что критерий типа χ^2 Никулина мощнее, чем критерии χ^2 Пирсона и отношения правдоподобия.

Зависимость мощности от числа интервалов k. За всю историю применения критериев типа χ^2 была предложена не одна формула для выбора числа интервалов, но ни одна из представленных в различных рекомендациях не выводилась с позиций максимальной мощности применяемого критерия, а в основном исходя из близости плотности к ее непараметрической оценке, гистограмме. Зная предельные распределения $G(S|H_0)$ и $G(S|H_1)$ статистики S, для любого заданного уровня значимости α можно оценить мощность соответствующего критерия,

рассматривая ее как функцию от числа интервалов k при заданном объеме выборки n. Исследование мощности критериев Пирсона и Никулина как функций от n и k проводилось аналитически и методами статистического моделирования. Результаты аналитических вычислений были полностью подтверждены оценками, полученными на основании моделирования.

Мощность критериев типа χ^2 определяется выражением [5]

$$1-\beta = P(s|r,\alpha) = \exp^{-s/2} \sum_{j=0}^{\infty} \frac{s^j}{j!2^{2j-1+r/2}\Gamma(j+r/2)} \times$$

$$\times \int_{\sqrt{x(\alpha,r)}}^{\infty} y^{2j-1+r} \exp^{-y^2/2} dy, \qquad (5)$$

где s – параметр нецентральности, определяемый соотношениями (2) и (4), $x(\alpha,r)$ представляет собой $(1-\alpha)$-процентную точку χ^2-распределения с r степенями свободы (α – заданная вероятность ошибки первого рода, β – вероятность ошибки второго рода). Приводимые ниже функции мощности строились при уровне значимости $\alpha = 0{,}1$.

Анализ функций мощности для различных альтернатив при проверке простых и сложных гипотез показал, что с увеличением числа интервалов мощность критериев типа χ^2 падает. Это соответствует результатам работ [6,7]. Максимальная мощность критериев при заданном объеме выборки n чаще всего достигается или при минимальном числе интервалов, или при некотором оптимальном значении k. Оптимальное число интервалов k зависит от объема выборки n и от конкретной пары конкурирующих гипотез H_0 и H_1. Чаще всего оптимальное k оказывается существенно меньше значений, рекомендуемых различными регламентирующими документами и задаваемых множеством эмпирических формул.

Таким образом, увеличение мощности критериев χ^2 Пирсона и отношения правдоподобия возможно за счет двух факторов: за счет выбора асимптотически оптимального группирования в качестве способа разбиения области определения случайной величины и подбора оптимального числа интервалов k при заданном объеме выборки n. Увеличение мощности критерия типа χ^2 Никулина возможно только за счет выбора оптимального числа интервалов.

Рассматривая пару альтернатив, всегда можно выбрать оптимальное число интервалов и подобрать оптимальное разбиение на интерва-

лы, в результате чего будем иметь критерий максимальной мощности, наилучшим образом различающий данные конкурирующие гипотезы.

3. Применение непараметрических критериев согласия при проверке сложных гипотез

К наиболее используемым критериям согласия относятся непараметрические критерии типа Колмогорова, типа ω^2 и Ω^2 Мизеса. В критерии Колмогорова в качестве расстояния между эмпирическим и теоретическим законом используется величина

$$D_n = \sup_{|x|<\infty} |F_n(x) - F(x,\theta)|,$$

где $F_n(x)$ – эмпирическая функция распределения, $F(x,\theta)$ – теоретическая функция распределения, n – объчм выборки. При проверке гипотез обычно используется статистика вида [5]

$$S_K = \frac{6nD_n + 1}{6\sqrt{n}}, \text{ где } D_n = \max(D_n^+, D_n^-),$$

$$D_n^+ = \max_{1\le i\le n}\left\{\frac{i}{n} - F(x_i,\theta)\right\},\ D_n^- = \max_{1\le i\le n}\left\{F(x_i,\theta) - \frac{i-1}{n}\right\},$$

n – объем выборки, $x_1, x_2, ..., x_n$ – упорядоченные по возрастанию выборочные значения, $F(x,\theta)$ – функция закона распределения, согласие с которым проверяется. Распределение величины S_K при простой гипотезе в пределе подчиняется закону Колмогорова $K(S)$ [5].

В критериях типа ω^2 расстояние между гипотетическим и истинным распределениями рассматривается в квадратичной метрике

$$\int_{-\infty}^{\infty} \{E[F_n(x)] - F(x)\}^2 \psi(F(x)) dF(x),$$

где $E[\star]$ – оператор математического ожидания.

При выборе $\psi(t) = 1$ в критериях типа ω^2 Мизеса пользуются статистикой (статистика Крамера–Мизеса–Смирнова) вида

$$S_\omega = \frac{1}{12n} + \sum_{i=1}^{n}\left\{F(x_i,\theta) - \frac{2i-1}{2n}\right\}^2,$$

при простой гипотезе подчиняющейся распределению $a1(S)$ [5].

При выборе $\psi(t) = 1/t(1-t)$ в критериях типа Ω^2 Мизеса применяемая статистика (статистика Андерсона–Дарлинга) имеет вид

$$S_\Omega = -n - 2\sum_{i=1}^{n}\left\{\frac{2i-1}{2n}\ln F(x_i,\theta) + \left(1-\frac{2i-1}{2n}\right)\ln(1-F(x_i,\theta))\right\}.$$

В пределе эта статистика подчиняется распределению $a2(S)$[5].

В случае простых гипотез предельные распределения статистик непараметрических критериев типа Колмогорова, ω^2 и Ω^2 Мизеса известны давно и не зависят от вида наблюдаемого закона распределения и значений его параметров. Говорят, что эти критерии являются "свободными от распределения". Это достоинство предопределяет широкое использование данных критериев в приложениях.

При проверке сложных гипотез, когда по той же самой выборке оцениваются параметры наблюдаемого закона $F(x,\theta)$, непараметрические критерии согласия теряют свойство "свободы от распределения". Однако мощность непараметрических критериев при проверке сложных гипотез при тех же объемах выборок n всегда существенно выше, чем при проверке простых. И если при проверке простых гипотез непараметрические критерии типа Колмогорова, ω^2 и Ω^2 Мизеса уступают по мощности критериям типа χ^2, при условии, что в последних используется асимптотически оптимальное группирование, то при проверке сложных гипотез непараметрические критерии оказываются более мощными. Для того чтобы воспользоваться их преимуществами, надо только знать распределение $G(S|H_0)$ для проверяемой сложной гипотезы.

Различия в предельных распределениях тех же самых статистик при проверке простых и сложных гипотез очень существенны. Поэтому предостережения против неаккуратного применения критериев согласия при проверке сложных гипотез неоднократно поднимались на страницах печати. При проверке сложных гипотез на условный закон распределения статистики $G(S|H_0)$ влияет целый ряд факторов, определяющих "сложность" гипотезы: вид наблюдаемого закона $F(x,\theta)$, соответствующего истинной гипотезе H_0; тип оцениваемого параметра и количество оцениваемых параметров; в некоторых ситуациях конкретное значение параметра (например в случае гамма-распределения); используемый метод оценивания параметров и точность вычисления оценок.

Исходной точкой для исследований предельных распределений статистик непараметрических критериев согласия при сложных гипотезах послужила работа [8]. В литературных источниках изложен ряд подходов к использованию непараметрических критериев согласия в случае

проверки сложных гипотез. При достаточно большом объеме выборки ее можно разбить на две части, и по одной из них оценивать параметры, а по другой проверять согласие [9]. В некоторых частных случаях предельные распределения статистик исследовались аналитическими методами [10], процентные точки распределений строились методами статистического моделирования [11–14]. Для приближенного вычисления вероятностей "согласия" вида $P\{S > S^*\}$ (достигаемого уровня значимости) строились формулы, дающие достаточно хорошие приближения при малых значениях соответствующих вероятностей [15–17]. В наших работах исследование распределений статистик непараметрических критериев согласия и построение моделей этих распределений осуществлялись с использованием методики компьютерного анализа статистических закономерностей.

Наши исследования показали, что распределения статистик критериев согласия существенно зависят от метода оценивания параметров. Вообще говоря, каждому типу оценок при конкретной сложной проверяемой гипотезе соответствует свой предельное распределение $G(S|H_0)$ статистики. В случае метода максимального правдоподобия распределения статистик $G(S|H_0)$ очень сильно зависят от закона, соответствующего гипотезе H_0. В то же время, разброс распределений $G(S|H_0)$ при использовании MD-оценок, минимизирующих статистику критерия, в существенно меньшей степени зависит от вида закона $F(x,\theta)$, соответствующего гипотезе H_0. Это позволяет говорить об определенной "свободе от распределения" для рассматриваемых критериев. Если опираться только на этот факт, то, казалось бы, что только такие методы оценивания и следует применять при проверке сложных гипотез. Однако исследование мощности рассматриваемых критериев при различных методах оценивания показало, что наибольшую мощность данные критерии при близких альтернативах имеют в случае использования ОМП. На рис. 1 иллюстрируется изменение распределений $G(S_K|H_0)$ статистики типа Колмогорова S_K в зависимости от числа оцениваемых параметров закона распределения Su-Джонсона с плотностью

$$f(x) = \frac{\theta_1}{\sqrt{2\pi}\sqrt{(x-\theta_3)^2+\theta_2^2}} \times$$

$$\times \exp\left\{-\frac{1}{2}\left[\theta_0 + \theta_1 \ln\left\{\frac{x-\theta_3}{\theta_2} + \sqrt{\left(\frac{x-\theta_3}{\theta_2}\right)^2 + 1}\right\}\right]^2\right\}.$$

На рисунке кривая *1* – функция распределения Колмогорова, ко-

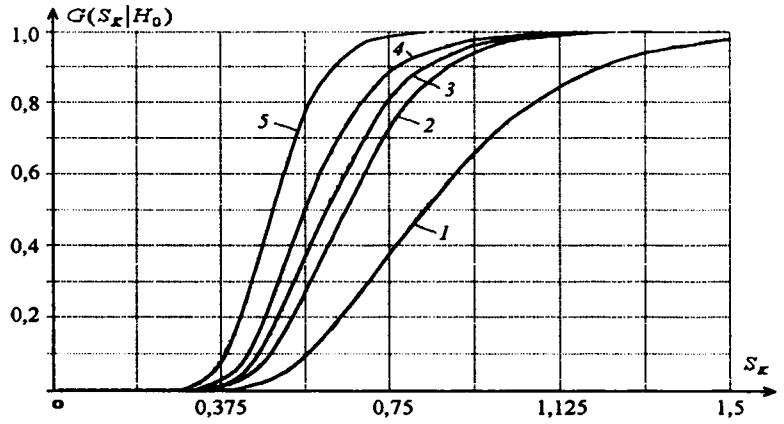


Рис. 1. Изменение распределений статистики типа Колмогорова в зависимости от числа оцениваемых параметров наблюдаемого распределения Su-Джонсона

торому подчиняется статистика при проверке простой гипотезы, *2* – распределение статистики при проверке сложной гипотезы и оценивании по данной выборке методом максимального правдоподобия только параметра θ_0, *3* – при оценивании только параметров θ_0 и θ_1, *4* – при оценивании параметров θ_0, θ_1 и θ_2, *5* – при оценивании всех четырех параметров θ_0, θ_1, θ_2 и θ_3.

Распределения статистик непараметрических критериев согласия существенно зависят от вида и числа оцениваемых параметров. И даже при оценивании единственного параметра предельное распределение статистики резко отличается от предельного распределения той же самой статистики в случае проверки простой гипотезы. Различие возрастает с увеличением числа оцениваемых параметров. Пренебрежение этим фактом в практике применения критериев согласия приводит к большим ошибкам в вычислении вероятности вида $P\{S > S^*\}$ и неоправданному принятию проверяемой гипотезы.

При малых объемах выборки n распределения $G(S_n|H_0)$ зависят от n. Однако существенная зависимость распределения статистик от n наблюдается только при небольших объемах выборки. Как показали наши исследования, при $n \geq 15 - 20$ распределения $G(S_n|H_0)$ достаточно близки к предельным $G(S|H_0)$, и зависимостью от n можно пренебречь.

Построенные на настоящий момент другими исследователями таблицы процентных точек и предельные распределения статистик непараметрических критериев ограничены относительно узким кругом сложных гипотез. Это объясняется тем, что построение предельного распределения аналитическими методами выливается в чрезвычайно непростую задачу. В то же время методика компьютерного анализа статистических закономерностей, хорошо зарекомендовавшая себя при моделировании распределений статистик критериев, позволяет при необходимости без существенных трудностей расширить этот круг.

В рекомендациях [2] нами были построены модели предельных распределений статистик рассматриваемых критериев при проверке сложных гипотез относительно 17 различных законов наблюдаемых случайных величин при использовании ОМП и MD-оценок. Это в общей сложности 350 моделей предельных распределений для такого же количества вариантов сложных гипотез.

В соответствии с этой методикой следует по закону $F(x,\hat{\theta})$ смоделировать N выборок того же объема n, что и выборка, для которой необходимо проверить гипотезу H_0: $F(x) \in \{F(x,\theta), \theta \in \Theta\}$. Далее для каждой из N выборок вычислить оценки тех же параметров закона, а затем значение статистики S соответствующего критерия. В результате будет получена выборка значений статистики S_1, S_2,...,S_N с законом распределения $G(S_n|H_0)$ для проверяемой гипотезы H_0. По этой выборке при значительных N можно построить достаточно гладкую эмпирическую функцию распределения $G_N(S_n|H_0)$, которой можно непосредственно воспользоваться для вывода, следует ли принимать гипотезу H_0. При необходимости можно по $G_N(S_n|H_0)$ построить приближенную аналитическую модель, аппроксимирующую $G_N(S_n|H_0)$, и тогда уже, опираясь на эту модель, принимать решение относительно проверяемой гипотезы.

Как показали исследования, хорошей аналитической моделью для $G_N(S_n|H_0)$ часто оказывается один из следующих четырех законов: гамма-распределение, распределение Su-Джонсона, распределение Sl-Джонсона или логарифмически нормальное. В крайнем случае, всегда можно, опираясь на ограниченное множество законов распределения, построить модель в виде смеси законов.

Построенные и представленные в [2] аппроксимации предельных распределений статистик непараметрических критериев согласия расширяют область корректного применения этих критериев и рекомендуются широкому кругу исследователей. Расширение множества моделей распределений статистик, то есть расширение множества корректно про-

веряемых сложных гипотез, не является самоцелью. Это имеет смысл только для тех законов распределения случайных величин, которые наиболее часто используются в приложениях. Именно для этих законов следует продолжить исследования. В целом же мы имеем дело с бесконечными множествами существующих случайных величин, их законов, возможных формулировок сложных гипотез. Следовательно, речь должна идти о постепенном *создании информационных технологий* исследования статистических закономерностей, технологий обеспечивающих статистический анализ и обработку экспериментальных наблюдений в различных областях знаний и человеческой жизнедеятельности.

4. Заключение

Основу ГОСТ Р 50.1-033-2001 составляют полученные таблицы асимптотически оптимального группирования, обеспечивающие максимальную мощность критериев типа χ^2 при простых и сложных гипотезах, и рекомендации по их применению.

Содержание ГОСТ Р 50.1-037-2002 определяют построенные модели распределений статистик непараметрических критериев согласия и соответствующие таблицы процентных точек, обеспечивающие корректные статистические выводы при проверке сложных гипотез.

СПИСОК ЛИТЕРАТУРЫ

1. *ГОСТ Р 50.1-033-2001* Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть I. Критерии типа χ^2.
2. *ГОСТ Р 50.1-037-2002* Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть II. Непараметрические критерии.
3. *Никулин М.С.* Критерий хи-квадрат для непрерывных распределений с параметрами сдвига и масштаба // Теория вероятностей и ее применение. 1973. Т. XVIII. № 3. С. 583-591.
4. *Никулин М.С.* О критерии хи-квадрат для непрерывных распределений // Теория вероятностей и ее применение. 1973. Т. XVIII. № 3. С. 675-676.
5. *Большев Л.Н., Смирнов Н.В.* Таблицы математической статистики. М.: Наука, 1983. 416 с.
6. *Чибисов Д.М., Гванцеладзе Л.Г.* О критериях согласия, основанных на группированных данных // III советско-японский симпозиум по теории вероятностей. Ташкент: Фан, 1975. С. 183-185.
7. *Боровков А.А.* О мощности критерия χ^2 при увеличении числа групп // Теория вероятностей и ее применение. 1977. Т. XXII. № 2. С. 375-378.

8. *Kac M., Kiefer J., Wolfowitz J.* On tests of normality and other tests of goodness of fit based on distance methods // Ann. Math. Stat. 1955. Vol. 26. P. 189-211.
9. *Durbin J.* Kolmogoriv–Smirnov test when parameters are estimated // Lect. Notes Math. 1976. Vol. 566. P. 33-44.
10. *Мартынов Г.В.* Критерии омега-квадрат. М.: Наука, 1978. 80 с.
11. *Pearson E.S., Hartley H.O.* Biometrica tables for statistics. Vol. 2. Cambridge: University Press, 1972. 634 p.
12. *Stephens M.A.* Use of Kolmogorov–Smirnov, Cramer – von Mises and related statistics – vithout extensive table // J. R. Stat. Soc. 1970. Vol. 32. P. 115-122.
13. *Stephens M.A.* EDF statistics for goodness of fit and some comparisons // J. Am. Statist. Assoc. 1974. Vol. 69. P. 730-737.
14. *Chandra M., Singpurwalla N.D., Stephens M.A.* Statistics for test of fit for the extrem-value and weibull distribution // J. Am. Statist. Assoc. 1981. Vol. 76. P. 375.
15. *Тюрин Ю.Н.* О предельном распределении статистик Колмогорова–Смирнова для сложной гипотезы // Изв. АН СССР. Сер. Математич. 1984. Т. 48, № 6. С. 1314-1343.
16. *Тюрин Ю.Н., Саввушкина Н.Е.* Критерии согласия для распределения Вейбулла–Гнеденко // Изв. АН СССР. Сер. Техн. кибернетика. 1984. № 3. С. 109-112.
17. *Саввушкина Н.Е.* Критерий Колмогорова–Смирнова для логистического и гамма-распределения // Сб. тр. ВНИИ систем. исслед. 1990. № 8. С. 50-56.

Г. А. Медведев

Непараметрические методы оценивания нейтральных к риску вероятностей по данным рыночных наблюдений

Настоящая статья представляет собой краткий обзор результатов по непараметрическим методам оценивания нейтральных к риску распределений вероятностей с помощью наблюдения цен финансовых производных (в основном опционов европейского типа) в дату их истечения. Статья содержит в основном не описание методов, а информацию о задачах, которые решаются с их помощью, с соответствующими ссылками.

1. Введение

Инвесторы, действующие на финансовом рынке, крайне заинтересованы в наличии математических средств, позволяющих им в условиях неопределенности оценивать стоимости тех или иных финансовых контрактов. Однако эта проблема является, как правило, достаточно сложной в условиях стохастического поведения рыночных показателей. В то же самое время природа рынка определяет его стохастичность, которая описывается не какими-либо строгими законами, а только рациональным поведением его участников, имеющих разнообразные предпочтения. Поэтому стохастические закономерности рынка могут быть определены исключительно с помощью статистических выводов на основе наблюдений за рыночными процессами. Одной из таких проблем является идентификация вероятностных распределений, управляющих поведением цен рыночных активов. Широко распространенным классом рисковых активов являются финансовые производные, а среди них наибольшую известность имеют опционы, дающие право покупать (или продавать) финансовые инструменты по фиксированной цене независимо от состояния рынка. Уже в течение долгого времени определение стоимости опциона производится довольно систематически. Сначала делается предположение относительно процесса, порождающего данные. То есть принимается конкретный вид случайного процесса, управляющего ценой основного актива. Чаще всего в качестве процесса цены актива принимается геометрическое броуновское движение. Затем делается преобразование этого процесса таким образом, чтобы он описывался нейтральным к риску распределением вероятностей (НРРВ), с помощью рассуждений относительно динамического воспроизведения или рассматрения влияния цены риска явно. Наконец, оцениваются параметры модели и определяется стоимость опциона.

Отступление от этого систематического пути происходит только в немногих случаях, когда удается воспользоваться имеющимися аналитическими результатами, полученными при рассмотрении типичных рыночных ситуаций. Наиболее характерным из таких исключений является использование модели Блэка и Шоулса [1] (BS-модели). В рамках этой модели получается формула для получения стоимости опциона, в которой цена опциона определяется через известные рыночные переменные, кроме одной – волатильности. Волатильность – это единственный параметр в BS-формуле, который непосредственно не наблюдается. Поэтому практики начали искать соответствующее значение волатильности, приспосабливая BS-модель к наблюдаемым ценам опционов. В течение долгого времени после начала торговли опционами в начале

1970-х число их типов было невелико, поэтому не возникало проблем оценивания каких-либо других параметров, отличающихся от волатильности. Это коренным образом изменилось в последние годы прошлого столетия, когда финансовые производные стали основываться на разнообразных рыночных индексах финансового рынка США, а процедура их погашения стала усложняться введением многих параметров и различных гарантий. Было обнаружено [2], что модели опционов, продаваемых на рынках Японии, Германии и Великобритании, аналогичны.

Одни исследователи стали рассматривать новые модели, которые могли бы объяснить эти цены опционов, а другие пытались изучать разнообразные случайные процессы цен активов, основываясь на наблюдаемых ценах опционов. Это стало целесообразным, так как возникли сотни одновременно обращающихся опционов на рыночные индексы, отличающихся ценой исполнения и/или сроками истечения. При условии такого обилия данных естественно было поинтересоваться информацией, содержащейся в ценах опционов относительно нейтрального к риску случайного процесса, лежащего в основе актива.

Для дискретного класса случайных процессов на этот вопрос был дан ответ в серии статей [3–5] о классах моделей, которые в [3] названы порождаемыми биноминальными деревьями (implied binomial trees). Эти порождаемые наблюдаемыми ценами опционов деревья являются обобщением первоначально введенных в [6] биноминальных деревьев.

Здесь приводится краткий обзор литературы этого направления. В разделе 2 рассматривается понятие нейтральных к риску вероятностей (risk-neutral probability) и показывается, как определить эти вероятности по набору цен опционов с заданным сроком до истечения. В разделе 3 обсуждаются непараметрические модели определения стоимости опционов.

2. Нейтральные к риску вероятности

Определим зависимую от состояния цену π_i как цену актива, по которому выплачивается единица стоимости, если в заданную дату T происходит определенное событие i, и ничего не выплачивается в других случаях. Сумма таких зависимых от состояния цен по всем состояниям в дату T должна равняться цене безрисковой облигации, по которой достоверно выплачивается единица стоимости в дату погашения T. Когда арбитражные возможности на финансовом рынке отсутствуют, зависимые от состояния цены являются только положительными.

Часто удобно нормировать зависимые от состояния цены относительно цены облигации. Тогда совокупность этих зависимых от состо-

яния цен, нормированных ценами облигаций, обладает всеми свойствами распределения вероятностей: они неотрицательны и нормируются на единицу. Финансистами принято называть их "нейтральными к риску вероятностями". В результате появляется НРРВ $\{P_i\}$, которое является удобным средством для расчета цен финансовых инструментов:

$$P_i = r^{-T}\pi_i,$$

где r равно 1 плюс годовая доходность облигации. Теперь можно определять стоимость любого актива C с платежом X_i в дату T, вычисляя математическое ожидание по нейтральным к риску вероятностям:

$$C = E_P[X]/r^T.$$

На полном рынке или имеется столько торгуемых ценных бумаг (активов), сколько имеется состояний среды, или реализуется достаточное число промежуточных торговых сессий, чтобы имелось столько динамически ребалансируемых портфелей, сколько реализуется состояний среды. На таких полных рынках, как показано в [7], можно полностью идентифицировать нейтральное к риску распределение вероятностей через набор цен опционов европейского типа. Это распределение оказывается единственным, и в [8] дана точная формула для него:

$$\frac{\partial^2 C}{\partial K^2} = \pi_{S=K}.$$

Вторая производная цены европейского колла C относительно его цены исполнения K является зависимой от состояния ценой π будущей цены актива, погашаемого точно по цене исполнения опциона. Чтобы вывести нейтральные к риску вероятности через цены опционов, мы можем следовать любым из трех подходов. Во-первых, найдем кандидата на нейтральное к риску распределение и адаптируем окончательные цены опционов к наблюдаемым ценам. Во-вторых, приспособим функцию цен опционов по ценам исполнения с помощью наблюдаемых цен опционов. Затем используем модель [8], чтобы получить нейтральное к риску распределение. В-третьих, с помощью наблюдаемых волатильностей аппроксимируем порождаемые волатильности как функцию цен исполнения. Затем вычислим функцию цен опционов по ценам исполнения и, наконец, получим нейтральное к риску распределение.

Эта процедура применялась в многочисленных вариациях, и каждая из них имеет некоторые общие свойства, которые влияют на ее эффективность. Получение нейтрального к риску распределения сталкивается с определенными проблемами, так как распределения вероятностей

должны обладать определенными свойствами: они уменьшаются до нуля для достаточно малых и достаточно больших цен акций, поэтому центральная часть распределения имеет экстремумы, а сумма всех вероятностей распределения должна равняться единице.

Аппроксимация функции цен опционов по ценам исполнения затрудняется тем, что эта функция принимает значение цены акции минус настоящая стоимость потока платежей дивидендов по времени действия опциона при нулевой цене исполнения и приближается к нулю для очень высоких цен исполнения. Кроме того, цены опционов должны удовлетворять условиям отсутствия арбитража.

Лучшим методом аппроксимации функции порождаемых волатильностей является, по-видимому, численный. В BS-модели вместо реальной U-образной формы волатильность предполагается постоянной. В этом случае нейтральное к риску распределение обладает хорошим поведением, но это только до тех пор, пока функция не приобретает реальную форму.

Однако точное соответствие нейтрального к риску распределения множеству цен опционов сильно зависит от полноты рынка. Если рынки неполны, то имеются состояния среды, для которых нельзя воспроизвести опцион. Это может возникать или из-за торговых ограничений, или из-за дополнительных неторговых факторов, таких, как стохастическая волатильность или случайные скачки цен акций. Тогда наблюдаемые цены опционов могут порождаться целым набором нейтральных к риску распределений, а не одним единственным распределением. В свою очередь, это означает, что можно только устанавливать границы на цены опционов, а не определять их единственным образом через цены исполнения. В [9] исследованы такие границы при наличии скачков, а в [10, 11] найдены границы, которые основаны на предельных доходностях инвестиций. В [10] ограничено отношение прибыли к потерям, а в [11] ограничен коэффициент Шарпа (ожидаемая отдача на единицу волатильности).

В течение 1980-х гг. наблюдалось затишье в исследованиях, а затем в [12] была предложена первая версия процедуры, в результате которой получается гистограмма вероятностей непосредственно через цены опционов в предположении, что вероятности равномерно распределены между любыми двумя соседними ценами исполнения опциона. К сожалению, эта схема оказалась неустойчивой в вычислительном плане и допускает получение отрицательных вероятностей, если выборка цен опционов не имеет достаточный объем. В [13] детально исследована устойчивость этого метода.

Имеются обзорные статьи [14–16] в этом направлении; в них сосредоточено внимание на конкретных методах идентификации НРРВ [14], уделено внимание как методам, так и их применениям [15], установлены принципы классификации различных методов: разложения, параметрические и непараметрические [15, 16]. Здесь уделяется внимание непараметрическому подходу.

3. Непараметрические методы

Непараметрические методы используются для достижения большей гибкости в приспособлении НРРВ к ценам опционов, а не к какой-то параметрической форме распределения. Поэтому они допускают использование более широкого класса функций. Мы разделим эти методы на три группы.

1) Ядерные методы, которые концептуально относятся к регрессионным в том смысле, что они стремятся приспосабливать функцию к данным наблюдений. Главное отличие состоит в том, что ядерная регрессия не определяет параметрическую форму функции.

2) Методы максимума энтропии, позволяющие находить непараметрическое распределение вероятностей, которое является наиболее близким (с точки зрения объема имеющейся информации) к априорному распределению при удовлетворении некоторых ограничений, таких, как точное воспроизведение цены наблюдаемых опционов.

3) Методы подбора кривой, слабо связанная группа методов, которые стремятся приспособить порождаемые волатильности или нейтральные к риску вероятности как можно ближе к некоторой подходящей функции. Обычно используемой мерой качества приспособления является сумма квадратов отклонений между наблюдаемыми ценами опционов и ценами опционов, получающимися методом подбора кривой.

Ядерные методы основаны на концепции, что каждая точка данных (скажем, порождаемая волатильность при некоторой цене исполнения) должна быть проверена как центр области, где могла бы находиться истинная функция. Чем дальше значение переменной от точки данных наблюдения, тем менее вероятно, что истинная функция проходит через эту точку. Ядро $k(x)$, в качестве которого часто принимается плотность нормального распределения, измеряет соответствующий вклад в правдоподобие при отклонении переменной от точки данных.

Таким образом, если мы наблюдаем порождаемые волатильности как функции цен исполнения $\sigma_i(K_i)$, ядерная регрессия в одномерном

случае имеет вид

$$\sigma(K) = \frac{\sum_{i=1}^{n} k\left(\frac{K-K_i}{h}\right)\sigma_i}{\sum_{i=1}^{n} k\left(\frac{K-K_i}{h}\right)},$$

где h – коэффициент размытости, который управляет гладкостью ядерной регрессии. Ядерная регрессия предполагает наличие большого объема выборочных данных, что ограничивает ее использование.

В [17] используются пятимерные ядерные оценки как функции цены акции, цены исполнения, срока до истечения, процентной ставки и дивидендной доходности. Там же используется упрощенный трехмерный случай, когда переменными являются форвардная цена, цена исполнения и время до истечения. Эта методология используется в [18] для того, чтобы исследовать нейтральное к риску распределение процентных ставок. Однако там имеется предупреждение, что наличие инерции в процентных ставках нарушает предположение о независимости и тождественной распределенности наблюдений, и, таким образом, асимптотические результаты [17] не должны выполняться. И действительно, в [18] найдено существенное выборочное смещение. В касающемся этой темы исследовании [19] предлагается двумерная ядерная формула оценки как функции времени до истечения и наличия денежных средств. Кроме того, там предлагается моделировать случайный процесс порождаемой волатильности в течение делового дня.

Методы максимума энтропии связаны с максимизацией количества пропущенной информации при наличии ограничений; другими словами, они выражают стремление нанести наименьший ущерб относительно неизвестной или пропущенной информации. Чтобы реализовать эту идею, а также, в некотором смысле, аналогичную идею байесовской постановки, выбирается априорное распределение Q_i и находится апостериорное нейтральное к риску распределение P_i с помощью максимизации взаимной энтропии

$$-\sum_i P_i \ln\left(P_i/Q_i\right)$$

при таких ограничениях, как положительность вероятностей, равенство единице суммы вероятностей и точное совпадение цен опционов и основного актива с наблюдаемыми. Если априорное распределение оказывается равномерным, то взаимная энтропия сводится к обычной энтропии:

$$-\sum_i P_i \ln P_i.$$

В [20] использовались равномерное и логнормальное распределения как априорные, и получены некоторые результаты путем моделирования. В [21] найдено нейтральное к риску распределение при использовании распределения (в прошлом) цены актива как априорного. НРРВ там ограничено только требованием точного совпадения цены основного актива. Кроме того, в [21] определены цены опционов и найдено подтверждение асимметрии волатильности, которая оказалась довольно чувствительной к длительности выборочного интервала наблюдений, использованного для получения априорного распределения.

Методы подбора кривой составляют довольно широкий класс методов, в которых или порождаемая волатильность как функция цен исполнения, или само нейтральное к риску распределение аппроксимируется некоторой достаточно общей функцией. Эти функции обычно являются гладкими и образуют класс полиномов от цены исполнения.

Для увеличения гибкости функции такие многочлены могут представлять сплайны, покусочно составляющие функцию. Это достигается гладким сшиванием полиномиальных сегментов в так называемых узлах путем согласования значений и производных сегментов в узлах. Главное преимущество сплайнов состоит в том, что узлы служат для того, чтобы локализовать согласование многочленов и, таким образом, избегать колебаний и наличия экстремумов многочленов в областях с небольшим количеством данных.

И, наконец, гладкость может быть сделана явной целью в непараметрических методах подбора кривой для отыскания функции с помощью метода наименьших квадратов или минимизации интеграла от возведенных в квадрат вторых производных, что учитывает совокупную кривизну функции.

Мы рассмотрим, в первую очередь, методы, удовлетворяющие требованию реальной U-образной формы волатильности, а затем методы аппроксимации НРРВ. В [22] сначала использован метод подбора кривой к U-образной форме, аппроксимирующий функцию квадратичным полиномом. Затем порождаемые волатильности там переводятся в цены опционов, и получается НРРВ через вторые производные цен опционов по ценам исполнения согласно [8]. После чего логнормальные хвосты припасовываются к найденному НРРВ.

Разновидность этого метода, в котором квадратичный полином приспосабливается к порождаемой волатильности через дельта опциона – колл, а не через цену исполнения, предложена в [23]. Дельта – это производная цены опциона, вычисляемой по BS-формуле, по цене акции. В этом случае преимущество состоит в том, что не требуется припасо-

вывать логнормальные хвосты распределению, к нейтральному к риску. Подход [22] распространен в [24] на сплайны седьмого порядка. В [25] используются кубические сплайны для аппроксимации U-образной формы волатильности, а в [26] используются кубические B-сплайны, которые составляют особое семейство многочленов, используемых для сегментов аппроксимируемой функции. Полиномиальная аппроксимация логарифма U-образной функции используется в [27], чтобы автоматически исключить отрицательные значения порождаемой волатильности. Этот метод распространен на двумерный случай в [28].

Вычисления в методе подбора кривой при аппроксимации U-образной функции не требуют итераций и могут быть выполнены за один проход. Однако они не гарантируют, что получающиеся вероятности будут всегда положительны, что должно контролироваться отдельно. Вместе с тем, с практической точки зрения, аппроксимация U-образной формы обычно проще, чем непосредственная аппроксимация распределения вероятности.

Вторая группа методов подбора кривой стремится аппроксимировать нейтральное к риску распределение непосредственно. В [29] минимизируется расстояние между дискретизированными вероятностями и логнормальным априорным распределением при условии точного определения наблюдаемых цен опционов и основного актива. В [30] показано, что различные способы измерения этого расстояния, такие, как вычисление квадратов или абсолютных значений отклонений, или взаимной энтропии, незначительно влияют на результирующее НРРВ, пока наблюдается достаточно много цен опционов. Авторы [30] также вводят новый метод, который не нуждается в априорном распределении и максимизирует гладкость нейтрального к риску распределения при условии точного определения наблюдаемых цен опционов. В [31] используется та же самая идея относительно максимизации гладкости распределения и аппроксимируется плотность, задаваемая с помощью кубических сплайнов. Близкий метод, в котором конструируется логарифм нейтрального к риску распределения из кусочно линейных сегментов предлагается в [32]. Затем в байесовой постановке там используется показательное априорное распределение и находится НРРВ, которое согласовывается с наблюдаемыми ценами опционов.

Хотя методы определения нейтральных к риску распределений многочисленны, результаты имеют тенденцию быть довольно похожими, если имеется достаточно много наблюдаемых цен опционов. Причина кроется в том, что если наблюдается только от десяти до пятнадцати опционных продаж, безарбитражные границы настолько сжимают-

ся, что цена очередного опциона часто определяется в пределах спреда предложения–спроса. Это встречается реже на хвостах распределения, но довольно точно характеризует случаи ликвидных опционов. В [30] делается вывод, что можно использовать любой разумный метод и результирующие нейтральные к риску распределения будут похожими. Это наблюдение подтверждается в [25], где используются три различных метода: кубические сплайны, метод минимального расстояния из [29] и смесь логнормальных распределений, чтобы аппроксимировать U-образную форму волатильности опционов на обменные курсы. Подобные результаты также получились в [34] при использовании трех различных методов: смеси логнормальных распределений, разложения по полиномам Эрмита и максимизации энтропии для аппроксимации цен опционов на фьючерсы облигаций.

4. Заключение

Начиная с кризиса 1987 года исследователи поняли, что BS-модель не может объяснить U-образную форму волатильности на рынках опционов на рыночные индексы. Для опционов на индивидуальные акции и обменные курсы также характерна U-образная форма волатильности. Это привело к возобновлению интереса к параметрическим расширениям BS-модели, которые сейчас обычно включают стохастическую волатильность и наличие случайных скачкообразных изменений. Анализ таких моделей является довольно сложным.

Другой подход, рассмотренный здесь, состоит в том, чтобы использовать наблюдаемые цены опционов и с помощью соответствующего статистического анализа изучить более глубоко случайный процесс цены актива. При наличии достаточно большой выборки цен опционов с одинаковым сроком истечения можно найти нейтральные к риску распределения вероятностей, которые позволяют рассчитать эти цены. Кроме того, можно также изучать с помощью наблюдаемых цен опционов структуру случайного процесса цены основного актива, что может привести к аппроксимации распределения вероятностей и этого процесса.

СПИСОК ЛИТЕРАТУРЫ

1. *Black F., Scholes M.* Pricing of options and corporate liabilities // J. Polit. Econ. 1973. Vol. 81. P. 637-659.
2. *Tompkins R.* Implied volatility surfaces: uncovering the regularities for options on financial futures. Working paper, Vienna University of Technology and Warwick University, 1998.
3. *Rubinstein M.* Implied binomial trees // J. Finance. 1994. Vol. 49, No. 3. P. 771-818.

4. *Derman E., Kani I., Chriss N.* Implied trinomial trees of the volatility smile // J. Derivatives. 1996. Vol. 3, No. 4. P. 7-22.
5. *Dupire B.* Pricing with a smile // Risk. 1994. Vol. 7, No. 1. P. 18-20.
6. *Cox J., Ross S., Rubinstein M.* Option pricing: a simplified approach // J. Financial Economics. 1979. Vol. 7, No. 3. P. 229-263.
7. *Ross S.* Options and efficiency // Quarterly J. Economics. 1976. Vol. 90. P. 75-89.
8. *Breeden D., Litzenberger R.* Prices of state-contingent claims implicit in options prices // J. Business. 1978. Vol. 51. P. 621-651.
9. *Masson J., Perrakis S.* A jumping smile. Working paper, University of Ottawa, 1997.
10. *Bernardo A., Ledoit O.* Gain, loss and asset pricing // J. Polit. Econ. 1999. Vol. 107. P. 79-83.
11. *Cochrane J., Saa-Requejo J.* Beyond arbitrage: good-deal asset price bounds in incomplete markets // J. Polit. Econ. 1999. Vol. 107. P. 117-160.
12. *Longstaff F.* Martingale restriction tests of option pricing models. Working paper, University of California, Los Angeles, 1990.
13. *Mayhew S.* On estimating the risk-neutral probability distribution implied by option prices. Working paper, Purdue University, 1995.
14. *Cont R.* Beyond implied volatility // Econophysics / Eds. J. Kertesz and I. Kondor. Dordrecht: Kluwer, 1997.
15. *Bahra B.* Implied risk-neutral probability density functions from option prices: theory and application. Working paper, Bank of England, 1997.
16. *Jackwerth J. C.* Option-implied risk-neutral distributions and implied binomial trees: literature review // J. Derivatives. 1999. Vol. 7. P. 66-82.
17. *Ait-Sahalia Y., Lo A.* Nonparametric estimation of state-price densities implicit in financial asset prices // J. Finance. 1998. Vol. 53, No. 2. P. 499-547.
18. *Pritsker M.* Nonparametric density estimation and tests of continuous time interest rate models. Working paper, Federal Reserve Board, Washington, DC, 1997.
19. *Rookley C.* Fully exploiting the information content of intra day option quotes: applications in option pricing and risk management. Working paper, University of Arizona, 1997.
20. *Buchen P., Kelly M.* The maximum entropy distribution of an asset inferred from option prices // J. Financial and Quantitative Analysis. 1996. Vol. 31, No. 1. P. 143-159.
21. *Stutzer M.* A simple nonparametric approach to derivative security valuation // J. Finance. 1996. Vol. 51. P. 1633-1652.
22. *Shimko D.* Bounds on probability // Risk. 1993. Vol. 6. P. 33-37.
23. *Malz A.* Estimating the probability distribution of the future exchange rate from option prices // J. Derivatives. 1997. Vol. 5, No. 2. P. 18-36.
24. *Brown G., Toft K.B.* Constructing binomial trees from multiple implied probability distributions // J. Derivatives. 1999. Vol. 7, No. 2.
25. *Campa J., Chang K., Reider R.* Implied exchange rate distributions: evidence from OTC option markets // J. International Money and Finance. 1998. Vol. 17, No. 1. P. 117-160.

25. *Aparicio S., Hodges S.* Implied risk-neutral distribution: a comparison of estimatic methods. Working paper, Warwick University, 1998.
26. *Rosenberg J., Engle R.* Option hedging using empirical pricing kernels. Working paper, New York University, 1997.
27. *Rosenberg J.* Pricing multivariate contingent claims using estimated risk-neutral density functions. Working paper, University of California, San Diego, 1996.
28. *Rubinstein M.* Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978 // J. Finance. 1985. Vol. 40, No. 2. P. 455-480.
29. *Jackwerth J., Rubinstein M.* Recovering probability distributions from option prices // J. of Finance. 1996. Vol. 51. P. 1611-1631.
30. *Mayhew S.* On estimating the risk-neutral probability distribution implied by option prices. Working paper, Purdue University, 1995.
31. *Hartvig N., Jensen J., Pedersen J.* Risk neutral densities of the 'Christmas Tree' type. Working paper, Centre for Mathematical Physics and Stochastics, Aarhus University, 1999.
32. *Coutant S., Jondeau E., Rockinger M.* Reading interest rate and bond futures options' smiles: how PIBOR and notional operators appreciated the 1997 French snap election. Working paper, HEC, 1998.

В. М. Недельк

К задаче оценивания качества решающей функции [1]

Работа посвящена проблеме статистической устойчивости решающих функций, или оценивания риска. Для случая дискретной "независимой" переменной получена зависимость между средними риском и эмпирическим риском решения для "наихудшего" распределения ("стратегии природы"). Полученный результат позволяет оценить степень завышенности оценок Вапника–Червоненкиса.

Предлагается эвристический подход к применению полученных результатов для оценивания качества решения в общем случае.

1. Введение

Пусть X – пространство значений "независимых переменных", а Y – целевое пространство прогнозируемых значений. Пусть C – множество вероятностных мер на $D = X \times Y$. Конкретную меру $c \in C$ будем обозначать $P_c[D]$. Здесь и далее квадратные скобки будут использоваться

[1] Работа выполнена при поддержке РФФИ, грант 01-01-00839

для указания множества, на σ-алгебре подмножеств которого задается мера, а круглые скобки — для указания меры множества (вероятности события). В общем случае необходимо дополнительно потребовать также существование $\forall x \in X$ условных мер $P_c[Y|x]$.

Под решающей функцией понимается отображение $f : X \to Y$. Для определения качества решающей функции необходимо задать функцию потерь $L : Y^2 \to [0, +\infty)$. Теперь под риском будем понимать средние потери

$$R(c, f) = \int L(y, f(x))\, dP_c[D] = \int R_x(c, f)\, dP_c[X],$$

где $R_x(c, f) = \int L(y, f(x))\, dP_c[Y|x]$.

Решающая функция строится на основе случайной независимой выборки $v_c = \{(x^i, y^i) \in D |\ i = \overline{1, N}\}$ из распределения $P_c[D]$.

В качестве выборочной оценки риска вводится понятие эмпирического риска

$$\widetilde{R}(v, f) = \frac{1}{N} \sum_{i=1}^{N} L(y^i, f(x^i)).$$

Для всех практически используемых алгоритмов построения решающих функций эмпирический риск является смещенной оценкой риска, причем всегда заниженной, поскольку данные алгоритмы минимизируют эмпирический риск.

Актуальной является задача оценки этого смещения.

Введем обозначения:

$$F(c, Q) = ER(c, f_{Q,v});$$
$$\widetilde{F}(c, Q) = E\widetilde{R}(v, f_{Q,v}).$$

Здесь $Q : \{v\} \to \{f\}$ – алгоритм построения решающей функции, а $f_{Q,v}$ – решающая функция, построенная по выборке v алгоритмом Q.

Математическое ожидание берется по всем выборкам объема N.

Введем функцию наибольшего смещения

$$S_Q(\widetilde{F}_0) = \widehat{F}_Q(\widetilde{F}_0) - \widetilde{F}_0, \tag{1}$$

где $\widehat{F}_Q(\widetilde{F}_0) = \sup\limits_{c:\, F(c,Q) = F_0} F(c, Q)$.

Основной результат данной работы заключается в нахождении зависимости $S_Q(\widetilde{F}_0)$ для случая дискретных X и Q, минимизирующих эмпирический риск в каждой точке x.

2. Дискретный случай

Пусть X дискретно, то есть $X = \{1, ..., n\}$.

Тогда $R(c, f) = \sum\limits_{x=1}^{n} p_x R_x(\xi_x, f(x))$, где $p_x = P_c(x)$, $R_x(\xi_x, v_x) = P(y \neq f(x)|x)$ – условный риск в точке x, ξ_x – краткое обозначение условной меры $P_c[Y|x]$, v_x – подмножество выборки для значения x.

Зададимся решающей функцией, минимизирующей эмпирический риск: $f_v^*(x) = \arg\min\limits_{y \in Y} \widetilde{L}(y, v_x)$, $\widetilde{L}(y, v_x) = \sum\limits_{v_x} L(y, y^i)$.

Определим, от каких величин зависят искомые средние.

$$F(c, Q) = ER(c, f_v^*) = \sum_{x=1}^{n} p_x ER_x(\xi_x, f_{v_x}^*(x)) = \sum_{x=1}^{n} F_x(p_x, \xi_x),$$

где введено обозначение $F_x(p_x, \xi_x) = p_x ER(\xi_x, f_{v_x}^*(x))$.

Аналогично

$$\widetilde{F}(c, Q) = \frac{1}{N} E\widetilde{L}(v, f_v^*) = \sum_{x=1}^{n} E\widetilde{L}_x(v_x, f_{v_x}^*(x)) = \sum_{x=1}^{n} \widetilde{F}_x(p_x, \xi_x).$$

Введем функцию

$$\widehat{F}_x(p_x, \tilde{F}_x^0) = \sup_{\xi_x : F_x(p_x, \xi_x) = F_x^0} F_x(p_x, \xi_x).$$

Теперь

$$\widehat{F}_Q(\tilde{F}_0) = \max \sum_{x=1}^{n} \widehat{F}_x(p_x, \tilde{F}_x^0), \tag{2}$$

где максимум берется по всем p_x и $\tilde{F}_x^0$, $x = \overline{1, n}$, при ограничениях $p_x \geq 0$, $\tilde{F}_x^0 \geq 0$, $\sum\limits_{x=1}^{n} p_x = 1$, $\sum\limits_{x=1}^{n} \tilde{F}_x^0 = \tilde{R}^0$.

Исходная экстремальная задача существенно упростилась, разбившись на два этапа: нахождение $\widehat{F}_x(p_x, \tilde{F}_x^0)$ и нахождение максимума функции на простой подобласти евклидова пространства. Функция $\widehat{F}_x(p_x, \tilde{F}_x^0)$, как показано ниже, может быть легко аппроксимирована путем численного моделирования.

Однако решать задачу (2) непосредственно численными методами проблематично, поскольку размерность пространства, по которому производится максимизация, равна $2n$, где n может быть достаточно велико (практический интерес представляют значения порядка десятков).

3. Решение экстремальной задачи

Перепишем задачу (2) в абстрактных обозначениях:

$$\sum_{x=1}^{n} \Phi(z^x) \to \max_{z^x}, \tag{3}$$

$z^x = (z_1^x, ..., z_m^x)$, $z_j^x \geq 0$, $\sum\limits_{x=1}^{n} z_j^x = 1$, $j = \overline{1,m}$.

Применительно к задаче (2) Φ соответствует $\widehat{F}_x$, $m = 2$, $z_1^x = p_x$, $z_2^x = \widetilde{F}_x^0 / \widetilde{F}_0$.

Дискретизуем пространство значений вектора z^x. То есть предположим, что $z^x \in \{t^1, ..., t^l\}$.

Тогда задачу (3) можно переписать в эквивалентном виде:

$$\sum_{i=1}^{l} \Phi(t^i)\kappa^i \to \max_{\kappa^i}, \tag{4}$$

где $\kappa^i \geq 0$, $\sum\limits_{i=1}^{l} t_j^i \kappa^i = 1$, $j = \overline{1,m}$, $\sum\limits_{i=1}^{l} \kappa^i = 1$, причем для κ^i допустимы лишь значения, кратные $1/n$, или 0.

Будем решать задачу без последнего ограничения дискретности κ^i. Получаем задачу линейного программирования, решение которой находится в некоторой вершине области значений аргументов. Это означает, что только $m + 1$ из κ^i могут быть ненулевыми.

Теперь легко показать, что вносимая снятием ограничения дискретности κ^i погрешность оказывается порядка m/n, и ей можно пренебречь.

Поскольку вывод о числе ненулевых κ^i не зависит от шага дискретизации пространства значений вектора z^x, результат можно распространить и на исходную задачу.

Утверждение 1. *Решение z^x задачи* (3) *включает (использует) не более $m + 1$ различных векторов.*

Таким образом, размерность пространства, по которому проводится поиск максимума, сокращается до $m(m + 1)$.

Применительно к задаче (2) это составляет 6, и задача легко решается численно.

4. Результаты

Предложенный метод позволяет найти зависимость $S_Q(\tilde{F}_0)$ при любых параметрах n и N.

Однако наиболее нагляден и удобен для сравнения с другими подходами асимптотический случай: $N/n = M = \text{const}$, $N \to \infty$, $n \to \infty$.

Рассматриваемое приближение вполне приемлемо уже при $n = 10$, при этом имеем лишь один входной параметр M.

Для иллюстрации рассмотрим задачу классификации (k классов). Функция потерь $L(y, y') = \begin{cases} 0, & y = y' \\ 1; & y \neq y'. \end{cases}$

Оценим $\widehat{F}_x(p_x, \widetilde{F}_x^0)$ методом машинного моделирования. Для этого будем случайным образом генерировать ξ_x, то есть условное распределение $P[Y|x]$.

Условное распределение ξ_x задается k параметрами (вероятности для каждого образа). В соответствии с равномерным распределением на множестве параметров генерируем ξ_x, для которого вычисляем значения $F_x(p_x, \xi_x)$ и $\widetilde{F}_x(p_x, \xi_x)$, и полученную пару изображаем графически в виде точки.

В рассматриваемом асимптотическом приближении два входных параметра N и p_x можно заменить одним: $\alpha = Np_x$.

Результаты моделирования при различных значениях параметра α приведены на рис. 1 (величины нормированы делением на $(k-1)/k$ – максимальную вероятность ошибки).

При $k = 2$ точки составляют сплошную линию. При $k > 2$ точки заполняют некоторую область, по которой также легко можно с достаточной точностью аппроксимировать $\widehat{F}_x(p_x, \widetilde{F}_x^0)$.

Аналогичным образом моделирование для оценки $\widehat{F}_x(p_x, \widetilde{F}_x^0)$ можно провести и в случае прогнозирования вещественной переменной [5].

Теперь мы можем определить точность оценок Вапника–Червоненкиса для рассмотренного случая дискретного X, поскольку мы нашли точную зависимость средних риска от эмпирического риска для "наихудшей" стратегии природы.

Для $\widehat{S}(\tilde{F}_0)$ в [1] приводится оценка $\widehat{S}'_V(\tilde{F}_0) = \tau$, а также улучшенная оценка $\widehat{S}_V(\tilde{F}_0) = \tau^2\left(1 + \sqrt{1 + \frac{2\tilde{F}_0}{\tau^2}}\right)$, где τ асимптотически стремится к $\sqrt{\ln 2/(2M)}$.

На рис. 2 для $M = 5$ приведены зависимость $\widehat{S}(\tilde{F}_0)$ и ее оценка $\widehat{S}_V(\tilde{F}_0)$. График демонстрирует значительную завышенность последней.

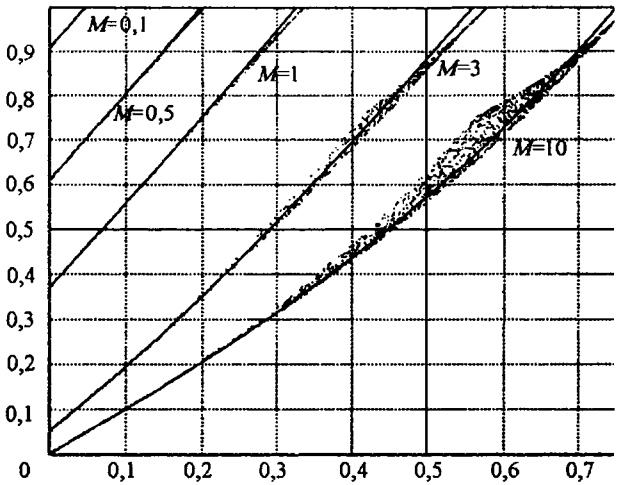

M=0,1
M=1
M=3
M=0,5
M=10
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7

Рис. 1

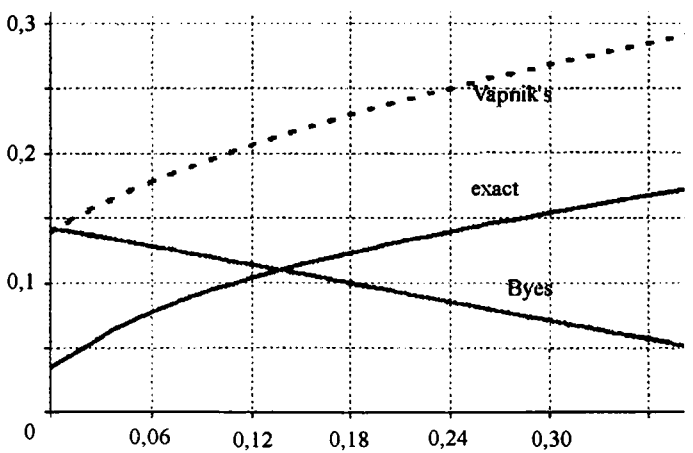

0,3
0,2
0,1
0
Vapnik's
exact
Byes
0,06
0,12
0,18
0,24
0,30

Рис. 2

Третья линия на графике (прямая) представляет вариант искомой зависимости, получаемый байесовским подходом, основанным на введении равномерного распределения на стратегиях природы [2–4, 8]. В работе [8] приводится результат $\widehat{S}_B(\widetilde{F}_0) = (1 - 2\widetilde{F}_0)/(M + 2)$.

Несколько неожиданным представляется тот факт, что $\widehat{S}(\widetilde{F}_0)$ на значительном интервале оказывается существенно меньше $\widehat{S}_B(\widetilde{F}_0)$, хотя байесовский подход подразумевает усреднение по всем стратегиям природы, а изложенный подход — выбор наихудшей.

К сожалению, обсуждение этого факта выходит за рамки данной работы, поскольку требует детального анализа байесовского подхода.

5. Практическое использование результатов

Полученная зависимость может быть практически использована для оценки ожидаемого риска. Для этого достаточно к величине полученного эмпирического риска $\widetilde{F}_0$ прибавить поправку $S_Q(\widetilde{F}_0)$.

Однако с прикладной точки зрения случай дискретного X и независимого принятия решения в каждой его точке имеет малый интерес.

Для использования полученного результата в реальных прикладных задачах предлагается следующий эвристический подход [6].

Метод оценивания вероятности ошибки основывается на гипотезе, что зависимость $\widehat{S}_Q(\widetilde{F}_0)$ в общем случае будет близкой к найденной оценке, параметры которой однозначно определяются M.

Параметр M можно оценить как $N/\log_2 \mathbf{C}$, где $\mathbf{C}$ – емкость класса решающих правил. Однако данная оценка будет в точности адекватна только для алгоритма выбора решающей функции, осуществляющего полный перебор. Для алгоритмов направленного поиска фактическая емкость может оказаться меньше.

Оценить фактическую емкость используемого алгоритма построения решающей функции можно путем моделирования по следующей схеме.

Задаемся равномерным распределением в пространстве D.

В соответствии с равномерным распределением моделируем выборки того же объема, что и выборка в реальной задаче.

Исследуемым алгоритмом построения решающей функции строим решающее правило и находим для него значение эмпирического риска.

Многократным моделированием оцениваем $\widetilde{F}_0$ – среднее значение эмпирического риска.

Поскольку вероятность ошибки для любого правила при равномерном распределении известна (например, при распознавании двух образов она равна 0,5), теперь знаем одну из точек (крайнюю правую)

кривой $\widehat{S}_Q(\widetilde{F}_0)$, а значит, можем определить M и восстановить всю зависимость.

Оценки ожидаемого риска часто бывает достаточно для практических нужд. При необходимости для риска могут быть дополнительно построены оценки доверительного интервала [7].

СПИСОК ЛИТЕРАТУРЫ

1. *Вапник В. Н., Червоненкис А. Я.* Теория распознавания образов. М.: Наука, 1974. 415 с.
2. *Hyghes G. F.* On the mean accuracy of statistical pattern recognizers // IEEE Trans. Inform. Theory. 1968. Vol. IT-14, No. 1. P. 55-63.
3. *Лбов Г. С., Старцева Н. Г.* Сложность распределений в задачах классификации // Докл. РАН. 1994. Т. 338, № 5. С. 592-594.
4. *Лбов Г. С., Старцева Н. Г.* Логические решающие функции и вопросы статистической устойчивости решений. Новосибирск: Институт математики, 1999. 211 с.
5. *Nedelko V. M.* An asymptotic estimate of the quality of a decision function based on empirical risk for the case of a discrete variable // Pattern Recognition and Image Analysis. 2001. Vol. 11, No. 1. P. 69-72.
6. *Недельko В. М.* Критерий оценки качества решающей функции по эмпирическому риску в задаче классификации // Искусственный интеллект. Изд-во НАН Украины, 2000. № 2. С. 172-178.
7. *Недельko В. М.* Оценивание доверительного интервала для вероятности ошибки решающей функции распознавания по эмпирическому риску // Доклады Всероссийской конференции "Математические методы распознавания образов". Москва: Изд-во ВЦ РАН, 1999.
8. *Бериков В. Б.* Об устойчивости алгоритмов распознавания в дискретной постановке // Искусственный интеллект. Изд-во НАН Украины, 2000. № 2. С. 5-8.
9. *Лбов Г. С., Недельko В. М.* Байесовский подход к решению задачи прогнозирования на основе информации экспертов и таблицы данных // Докл. РАН. 1997. Т. 357, № 1. С. 29-32.

И.А. Пестунов

О некоторых аспектах реализации классификатора Розенблатта–Парзена в условиях нестационарности и малых выборок

В работе рассматриваются особенности использования непараметрического классификатора Розенблатта–Парзена при наличии малых обучающих выборок и в условиях изменения вероятностных характеристик классов с течением времени.

Адаптация непараметрического классификатора Розенблатта–Парзена к условиям нестационарности

Традиционный подход к построению непараметрических классификаторов, основанных на оценках плотности Розенблатта–Парзена, заключается в подстановке в байесовское решающее правило (БРП) вместо неизвестных истинных вероятностных характеристик классов соответствующих им оценок, полученных по обучаемым выборкам [1]. При этом предполагается, что вероятностные характеристики классов с течением времени не изменяются. Однако в действительности все исследуемые объекты и процессы в большей или меньшей мере дрейфуют во времени и часто непредсказуемым образом. Причины этого дрейфа могут быть самые разные. Например, в задачах распознавания типов растительного покрова, по данным дистанционного зондирования, источником дрейфа может явиться изменение оптических свойств фотоэлементов (листьев, стеблей, цветков и т.п.), изменение их биомассы. В задачах оценки (прогноза) качества продукции по косвенным признакам это может происходить из-за старения технического оборудования, изменения подлежащих обработке сырьевых компонентов и т.п. В результате при решении практических задач нередко возникает необходимость адаптации непараметрического классификатора Розенблатта–Парзена к изменению вероятностных характеристик классов.

Рассмотрим задачу распознавания образов с позиции байесовской теории принятия решений, предполагая при этом, что условные плотности распределения $\{f_i(\cdot)\}$ классов $\{\Omega_i\}$ являются известными функциями времени, т.е. $f_1(\cdot), \ldots, f_M(\cdot) \in \mathcal{F}$, где $\mathcal{F} = f(t, x)$ — совокупность непрерывных плотностей распределения, зависящих от времени $t \in T$; T — некоторый числовой интервал $[a, b]$. Тогда БРП в текущий момент времени $t \in [a, b]$ примет вид

$$\delta_0 = \delta_0(t, x) = \arg\min_{j \in S} g_j(t, x),$$

где дискриминантная функция

$$g_j(t,x) = \sum_{i=1}^{M} \lambda_{ij} q_i f_i(t,x), \quad x \in \Re^k.$$

Здесь λ_{ij} — величина потерь при отнесении в Ω_j объектов из Ω_i; q_i — априорная вероятность класса Ω_i; $j \in S = \{1,2,\ldots,M\}$.

Предположим теперь, что плотности $\{f_i(t,x)\}$ неизвестны, а имеется лишь обучающая выборка $V = \bigcup\limits_{i=1}^{M} V^{(i)}$ объема $N = \sum\limits_{i=1}^{M} N_i$, где $V^{(i)} = \{x_j^{(i)} : x_j^{(i)} \in \Re^k$ — наблюдение из Ω_i, зарегистрированное в момент времени $t_j^{(i)}$, $j = 1,2,\ldots,N_i$ $(t_1^{(i)} \le \cdots \le t_{N_i}^{(i)})\}$. Требуется построить решающее правило (РП) для классификации наблюдений, поступающих в момент времени $t \ge t_{\max} = \max\limits_{i \in S} t_{N_i}^{(i)}$.

Ясно, что при наличии такой исходной информации плотности $\{f_i(t,x)\}$ не могут быть оценены, поскольку они являются неизвестными функциями времени t, а регистрация выборочной информации производится через некоторые интервалы Δt. Такая постановка вопроса принципиально отличается от известных тем, что в данном случае нельзя говорить о состоятельности каких-либо оценок плотностей $\{f_i(t,x)\}$, поскольку в разные моменты времени истинная функция принимает разные формы. С практической точки зрения, эта постановка может быть оправдана только при условии, если скорость изменения вероятностных характеристик классов значительно меньше темпа поступления обучающей информации.

Для конструирования непараметрического алгоритма Розенблатта–Парзена, предназначенного для работы в нестационарных условиях, предлагается использовать известный прием [2], суть которого состоит в том, что при оценивании плотностей $\{f_i(t,x)\}$ вновь поступившая информация получает больший вес по сравнению со "старой". Используя этот прием, непараметрическое РП можно записать в виде

$$\hat{\delta}_0 = \hat{\delta}_0(t_{\max}, x; V) = \arg\min_{j \in S} \left\{ \sum_{i=1}^{M} \lambda_{ij} q_i \hat{f}_{N_i}^{(i)}(t_{\max}, x) \right\}. \qquad (1)$$

Здесь

$$\hat{f}_{N_i}^{(i)}(t_{\max}, x) = \frac{1}{N_i} \sum_{j=1}^{N_i} \left(\prod_{l=1}^{k} h_l(N_i) \right)^{-1} \Phi\left[\frac{x_1 - x_{j1}^{(i)}}{h_1(N_i)}, \ldots, \frac{x_k - x_{jk}^{(i)}}{h_k(N_i)} \right] \rho(j, N_i),$$

где $\Phi(\cdot)$ — некоторая заданная колоколообразная функция, называемая ядром, $\{h_1(N_i),\ldots,h_k(N_i)\}$ — набор положительных параметров сглаживания (размытости), а $\rho(j,N_i)$ — некоторая определяемая функция памяти, дающая больший вес более поздним по времени регистрации элементам обучающей выборки и меньший вес "старым" (более ранним) элементам.

Приведем результаты двух экспериментов с псевдослучайными числами, подтверждающих эффективность РП (1).

В первом эксперименте рассматривался случай двух равновероятных классов, размерность пространства признаков $k=2$. Первый класс описывался инвариантным во времени двухмерным нормальным законом распределения $N(\mu^{(1)},I)$ с математическим ожиданием $\mu^{(1)}=(5,5)^T$ и единичной ковариационной матрицей I.

Второй класс в каждый дискретный момент времени $s=\overline{0,99}$ описывался также нормальным законом распределения $N(\mu^{(2)}(s),I)$ с математическим ожиданием $\mu^{(2)}(s)$ и единичной ковариационной матрицей I, при этом рассматривались два варианта.

В первом варианте

$$\mu^{(2)}(s)=(5+\sqrt{10{,}89}-sl,5)^T,$$

а во втором

$$\mu^{(2)}(s)=(5+\sqrt{4{,}33}+sl,5)^T,$$

что при $s=0$ соответствует $\sim$5 и 15% байесовской вероятности ошибки, $l=(\sqrt{10{,}89}-\sqrt{4{,}33})/100$.

Для каждого из этих вариантов моделировалось по одной обучающей выборке объема $N=N_1+N_2=200$ (в каждый дискретный момент времени $s=0,1,\ldots,99$ регистрировалось по одному наблюдению из 1-го и 2-го классов) и по одной контрольной выборке объема $N_k=200$ ($N_{k_1}=N_{k_2}=100$), составленной из наблюдений, зарегистрированных в момент времени $s=99$. Далее строилось РП (1). При этом функция ядра определялась выражением

$$\Phi\left(\frac{x-x_j^{(i)}}{h}\right)=(2\pi)^{-k/2}\prod_{l=1}^{k}\exp\left(-\frac{(x_l-x_{jl}^{(i)})^2}{2h^2}\right),$$

функция памяти $\rho(s,N_i)=(s/N_i)^\alpha$, где числовой параметр $\alpha\geq 0$. Заметим, что если $\alpha=0$, то получаем обычный классификатор Розенблатта–Парзена. Нахождение оптимального значения параметра сглаживания h осуществлялось обычным образом с использованием оценки "скользящего экзамена". Затем для некоторого фиксированного набора значений

α методом "контрольной выборки" оценивалось значение ВОК. Результаты этого эксперимента приведены в табл. 1.

Таблица 1

Вариант	Оценки	α						
		0	1	2	3	4	6	8
I	ВОК	0,165	0,135	0,12	0,105	0,105	0,12	0,12
	h	1,72	1,76	1,96	1,96	1,96	1,96	1,92
II	ВОК	0,05	0,045	0,045	0,045	0,045	0,045	0,045
	h	1,52	2,1	2,48	2,48	2,52	2,56	2,55

Во втором эксперименте рассматривался также случай двух равновероятных классов, размерность пространства признаков $k = 2$. Первый класс описывался инвариантным во времени законом распределения. Для объектов этого класса значения второго признака были распределены равномерно на отрезке $[-4, 6]$, а значения первого признака определялись в соответствии с выражением

$$x = -\sqrt{25 - (y-1)^2} + \xi,$$

где x, y — значения 1-го и 2-го признаков соответственно, а ξ — случайная величина, имеющая равномерное распределение на отрезке $[-1, 2]$. Второй класс в каждый дискретный момент времени $s = 0, 1, \ldots, 99$ описывался нормальным законом распределения $N(\mu^{(2)}(s), \Sigma)$ с математическим ожиданием $\mu^{(2)}(s)$ и диагональной ковариационной матрицей

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}.$$

В зависимости от значения σ^2 рассматривались три варианта: 1-й — $\sigma^2 = 0{,}5$; 2-й — $\sigma^2 = 0{,}7$ и 3-й — $\sigma^2 = 1$. Вектор $\mu^{(2)}(s)$ определялся выражением $\mu^{(2)}(s) = (-sd/99, 0)^T$, где d — величина максимального смещения вектора математического ожидания второго класса, которое произошло с момента времени $s = 0$ до момента $s = 99$. Далее для каждого из вариантов модели выполнялись действия, аналогичные действиям, выполненным в ходе 1-го эксперимента. Полученные в результате значения оценок ВОК и соответствующие им значения параметров сглаживания приведены в табл. 2.

Приведенные в табл. 1 и 2 результаты показывают, что достоверность распознавания в условиях нестационарности при использовании РП (1) во всех рассмотренных вариантах превосходит достоверность распознавания с помощью обычного классификатора Розенблатта–Парзена. Таким образом, в условиях дрейфа вероятностных характеристик

Таблица 2

σ	d		α			
			0	2	4	8
0,5	3	ВОК	0,13	0,102	0,085	0,10
		h	1,88	1,4	1,2	1,24
0,7	4	ВОК	0,275	0,213	0,165	0,17
		h	2,6	1,5	0,92	0,68
1	4	ВОК	0,32	0,22	0,195	0,195
		h	1,8	2,0	0,84	0,79

классов, используя РП (1), можно значительно уменьшить вероятность ошибочной классификации.

Настройка непараметрического классификатора Розенблатта–Парзена в условиях малых выборок

При решении прикладных задач распознавания образов обычно используется (0–1)-матрица потерь. В этом случае λ_{ij} $(i,j=\overline{1,M})$ определяется выражением $\lambda_{ij}=1-\delta_{ij}$, где δ_{ij} – символ Кронекера, а подстановочное РП записывается в виде

$$\hat{\delta}_0=\hat{\delta}_0(x;V)=\arg\max_{i\in S} q_i\hat{f}_{N_i}^{(i)}(x).$$

Для случая двух классов это правило можно переписать следующим образом:

$$\frac{\hat{f}_{N_1}^{(1)}(x)}{\hat{f}_{N_2}^{(2)}(x)}\gtrless\frac{q_2}{q_1}\to x\in\begin{cases}\Omega_1,\\ \Omega_2.\end{cases}$$

Величину $t=q_2/q_1$ называют пороговым значением отношения правдоподобия, или решающим порогом. В случае равновероятных классов этот порог равен 1.

Если в этом правиле в качестве $\{\hat{f}_{N_i}^{(i)}(x)\}$ используются непараметрические оценки Розенблатта–Парзена, то такое РП является состоятельным (в смысле байесовской ВОК) и при наличии представительных обучающих выборок обеспечивает высокое качество распознавания. Однако в условиях малых обучающих выборок из-за смещенности оценок Розенблатта–Парзена качество классификации в большинстве случаев оказывается неудовлетворительным.

С целью компенсации смещения оценок плотности предлагается отказаться от выбора решающего порога традиционным способом, а производить оптимизацию подстановочного РП не только по параметру сглаживания h, но и по значению параметра t.

Процедура оптимизации осуществляется в два этапа:

1) при фиксированном значении t выбирается оптимальное значение параметра h исходя из минимума оценки ВОК, полученной методом "скользящего экзамена";

2) из всех полученных наборов t и h выбирается тот, на котором достигается минимальное значение оценки "скользящего экзамена".

Построенные таким способом алгоритмы классификации исследовались методом статистического моделирования. Приведем результаты исследования для двух моделей: параметрической и непараметрической.

Параметрическая модель. Рассматривался случай двух классов ($q_1 = q_2 = 0{,}5$) в восьмимерном признаковом пространстве. Первый класс описывался нормальной плотностью с математическим ожиданием $\mu^{(1)} = (0,0,0,0,0,0,0,0)$ и единичной ковариационной матрицей $\Sigma^{(1)} = I$, а второй — нормальной плотностью с математическим ожиданием $\mu^{(2)} = (0,2{,}563,0,0,0,0,0,0)$ и ковариационной матрицей $\Sigma^{(2)} = 9I$.

Для данной модели генерировались обучающая выборка объема $N = (N_1 + N_2) = 200$ ($N_i/k \approx 12$) и контрольная, включающая по 500 реализаций из каждого класса. Затем для каждой из полученных обучающих выборок строился классификатор Розенблатта–Парзена с регулируемым решающим порогом. В качестве функции ядра использовалось мультипликативное нормальное ядро. Результаты эксперимента приведены в табл. 3, в которой представлено значение решающего порога t, значение оптимального параметра сглаживания h_{opt}, количество неправильно распознанных объектов на этапе обучения R_H и процент ошибок на этапе классификации R.

Таблица 3

t	h_{opt}	R_H	R, %
1	0,7	32	**12,3**
2	1,2	13	11,1
6	2,9	12	4,0
8	2,4	11	3,2
10	2,2	10	**3,0**

Из этой таблицы следует, что для данной восьмимерной модели за счет выбора решающего порога t удалось в 4 раза уменьшить ошибку классификации.

Непараметрическая модель. Моделировался случай двух равновероятных классов в четырехмерном признаковом пространстве. Первый класс состоял из точек, описываемых нормальными плотностя-

ми с ковариационной матрицей $\Sigma^{(1)} = 4I$. При этом для первой половины точек математическое ожидание $\mu_1^{(1)} = (0,0,0,0)$, а для второй — $\mu_2^{(1)} = (12,0,0,0)$. Второй класс состоял из точек, описываемых нормальными плотностями с единичной ковариационной матрицей $\Sigma^{(2)} = I$. Первая половина точек класса имела математическое ожидание $\mu_1^{(2)} = (6,0,0,0)$, а вторая — $\mu_2^{(2)} = (18,0,0,0)$.

Таблица 4

N	t	h_{opt}	R_H	R, %
20	**1**	08	1	**9,1**
	0,7	3,0	1	7,1
	0,45	2,6	1	**5,0**
	0,3	2,0	1	5,5
100	**1**	1,1	3	**4,9**
	0,7	1,4	2	4,9
	0,45	1,6	2	3,6
	0,4	2,6	1	**2,5**
400	**1**	0,4	10	**4,7**
	0,3	2,4	5	**2,4**
600	**1**	0,7	16	**3,3**
	0,3	1,6	10	**2,3**

Для данной модели генерировались обучающие выборки объема $N = (N_1 + N_2) = 20, 100, 400, 600$. Для каждой обучающей выборки генерировалась контрольная выборка, включающая по 500 реализаций из каждого класса. Далее выполнялись действия, аналогичные действиям для параметрической модели. Результаты эксперимента представлены в табл. 4.

Анализ табл. 3 и 4 показывает, что при наличии малых обучающих выборок ($N_i/k < 50$) за счет настройки решающего порога удается существенно улучшить качество классификации.

СПИСОК ЛИТЕРАТУРЫ

1. *Харин Ю.С.* Робастность в статистическом распознавании образов. Минск: Университетское, 1992. 232 с.
2. *Раудис Ш., Пикялис В., Юшкявичюс К.* Экспериментальное сравнение тринадцати алгоритмов классификации // Статистические проблемы управления. Вильнюс, 1975. Вып. 11. С. 53-80.
3. *Медведев А.В.* Адаптация в условиях непараметрической неопределенности // Адаптивные системы и их приложения. Новосибирск: Наука, 1978. С. 4-34.

П.Ф. Тарасенко

Проверка гипотез для индикаторного анализа линейных моделей

Предлагаются тесты проверки простых гипотез о параметрах линейной модели для случая, когда о погрешностях наблюдений известно только то, что они имеют, например, общий (но неизвестный) интерквартильный размах. В рамках этой задачи одновременно проверяется гипотеза о значениях квартилей. Рассмотрены обобщения этой задачи, различные подходы к построению тестов, получены предельные распределения их статистик.

1. Постановка задачи

Рассмотрим линейную модель наблюдений $Y=(Y_1,..,Y_N)^\top$ в виде

$$Y = X^\top\theta + \varepsilon, \tag{1}$$

где $\theta = (\theta_1,..,\theta_T)^\top \in \mathbf{R}^T$ – неизвестные параметры, X – неслучайная матрица плана, со столбцами $X_1,..,X_N$, и $\varepsilon = (\varepsilon_1,..,\varepsilon_N)^\top$ – независимые в совокупности случайные погрешности. Относительно распределения погрешностей $F(u) = P(\varepsilon_i < u)$ будем предполагать, что оно порождает априорно известное распределение $\overline{p} = \{p_1,..,p_K\}$ на множествах некоторого конечного разбиения $C(\mu) = \{C_1(\mu),..,C_K(\mu)\}$ пространства $\mathbf{R}^1$, известного с точностью до параметров $\mu = (\mu_1,..,\mu_M)^\top$. Иными словами, для всех $k \in \{1,\ldots,K\}$ известны значения $p_k = P(\varepsilon_i \in C_k(\mu))$.

Поставим задачу проверки простой гипотезы $H_0 : (\theta,\mu) = (\theta_0,\mu_0)$ против сложной альтернативы $H_1 : (\theta,\mu) \neq (\theta_0,\mu_0)$. Решение этой задачи является шагом на пути к построению оценок параметров θ и μ. Оценками можно считать такие θ_0 и μ_0, которые обеспечивают наибольший достигнутый уровень значимости при проверке гипотезы $(\theta,\mu) = (\theta_0,\mu_0)$.

Например, в знаковой (медианной) модели погрешностей разбиение $C = \{(-\infty,0),0,(0,\infty)\}$ не зависит от параметров и $\overline{p} = \{\frac{1}{2},0,\frac{1}{2}\}$. Здесь проверке подлежат гипотезы только о параметрах θ модели (1). Второй пример описывает квартильную априорную информацию о погрешностях, если $C(\mu) = \{(-\infty,-\mu),(-\mu,0),0,(0,\mu),(\mu,\infty)\}$ зависит от скалярных параметров μ и $\overline{p} = \{\frac{1}{4},\frac{1}{4},0,\frac{1}{4},\frac{1}{4}\}$. В этом случае можно выдвигать гипотезы как о параметрах θ модели (1), так и о масштабном параметре μ разбиения $C(\mu)$. Другие примеры будут рассмотрены ниже.

Уточним модель, которая описывает априорную информацию о погрешностях. Чтобы охватить сразу широкий спектр ситуаций, зададим семейство $\mathcal{C} = \{C(\mu') : \mu' \in \mathcal{M}\}$ разбиений пространства $\mathbf{R}^1$, параметризованное с помощью $\mu' = (\mu'_1, \ldots, \mu'_M)^\top \in \mathcal{M} \subseteq \mathbf{R}^M$. Каждое разбиение в $\mathcal{C}$ состоит из K непустых множеств, т.е.

$$C(\mu') = \{C_k(\mu') : k = \overline{1,K}\}.$$

Каждое из множеств $C_k(\mu')$ является интервалом в $\mathbf{R}^1$ или объединением нескольких несмежных интервалов, т.е.

$$C_k(\mu') = \bigcup_{l=1}^{K_k} C_{k,l}(\mu') = \bigcup_{l=1}^{K_k} \langle C_{k,l,0} + d_{k,l,0}^\top \mu',\ C_{k,l,1} + d_{k,l,1}^\top \mu' \rangle. \quad (2)$$

Здесь угловые скобки означают, что граница интервала может быть открытой или закрытой, при этом случай $C_{k,l,0} = C_{k,l,1}$, когда закрытый интервал является точкой, не исключается. Постоянные векторы $d_{k,l,j}$ состоят из нулей, кроме, возможно, одного элемента, принимающего значения $+1$ или -1. Иными словами, параметры μ' отвечают за сдвиги границ интервалов, из которых состоят множества разбиения. При этом сдвиги двух разных границ могут быть согласованными, если за них отвечает один и тот же параметр. Естественно, что набор точек $C_{k,l,j}$, набор векторов $d_{k,l,j}$ и область $\mathcal{M}$ допустимых значений параметров должны быть согласованы так, чтобы система множеств $\{C_{k,l}(\mu') : k = \overline{1,K}, l = \overline{1,K_k}\}$ оставалась разбиением при всех $\mu' \in \mathcal{M}$.

Априорная информация о погрешностях определяется тем, что для некоторого неизвестного набора параметров $\mu \in \mathcal{M}$ и известного набора значений вероятностей $\bar{p} = \{p_k : k = \overline{1,K}, p_k \geq 0, \sum_{k=1}^{K} p_k = 1\}$ имеет место

$$P(\varepsilon_i \in C_k(\mu)) = p_k, \quad i = \overline{1,N}, \quad k = \overline{1,K}. \quad (3)$$

Поскольку общий случай заслуживает отдельного рассмотрения, в данной работе будем считать, что все отличные от нуля априорные вероятности p_k одинаковы. Если среди K элементов разбиения (2) только K_1 имеют ненулевые вероятности (3), то либо $p_k = 0$, либо $p_k = 1/K_1$. Обозначим $p = 1/K_1$, однако далее будем продолжать использовать p_k.

2. Индикаторные признаки и метки множеств разбиения

Рассмотрим задачу проверки простой гипотезы H_0 о параметрах модели (1)–(3) против сложной альтернативы H_1, где

$$H_0 : (\theta, \mu) = 0, \quad H_1 : (\theta, \mu) \neq 0. \tag{4}$$

К (4) сводится и случай проверки более общей гипотезы $(\theta, \mu) = (\theta_0, \mu_0)$, если в (2) заменить Y на $Y - X^\top \theta_0$ и в (1) $C_{k,l,j}$ на $C_{k,l,j} + d^\top_{k,l,j}\ \mu_0$.

В качестве признаков для построения статистических выводов на основе априорного знания (3) достаточно использовать индикаторы $s = (s_1, \ldots, s_N)^\top$ попадания невязок модели (1) в множества (2), где

$$s_i = s_i(0,0), \quad s_i(\theta, \mu) = s(Y_i - X_i^\top \theta,\ \mu), \quad i = \overline{1, N}, \tag{5}$$

а функция $s(u, \mu)$ принимает значение k, если $u \in C_k(\mu)$. Пространство индикаторных признаков, таким образом, содержит K^N элементов. Обозначим его через $\mathcal{S}$. Функция мощности любого индикаторного теста (основанного на признаках (5)) с критической областью $\mathcal{S}_1 \subset \mathcal{S}$ для проверки гипотез о параметрах θ и μ может быть записана в виде $P(\mathcal{S}_1|\theta, \mu) = \sum\limits_{s \in \mathcal{S}_1} P(s|\theta, \mu)$, где $P(s|\theta, \mu)$ – правдоподобие набора признаков s. В силу независимости шумов имеем

$$P(s|\theta, \mu) = \prod_{i=1}^{N} P(s_i|\theta, \mu). \tag{6}$$

Отметим, что при гипотезе $P(s|0,0) = \prod\limits_{i=1}^{N} p_{s_i}$. Поэтому когда все отличные от нуля априорные вероятности p_k одинаковы, то либо $P(s|0,0) = 0$, либо $P(s|0,0) = p^N$. Если же набор признаков s рассматривается как случайный, то при выполнении гипотезы имеем $P(s|0,0) = p^N$ почти наверняка. Обозначив границы альтернативного разбиения через $C_{k,l,j,i}(\theta, \mu) = C_{k,l,j} + d^\top_{k,l,j}\ \mu - X_i^\top \theta$, получим

$$P(s_i = k|\theta, \mu) = \sum_{l=1}^{K_k} [F(C_{k,l,1,i}(\theta, \mu)) - F(C_{k,l,0,i}(\theta, \mu))]. \tag{7}$$

В дальнейшем анализе важную роль играют локальные (в окрестности гипотезы) свойства распределения (6), поэтому дадим их описание.

При условии непрерывной дифференцируемости F в окрестностях точек $C_{k,l,j}$, из (7) получаем градиенты распределения каждого признака по параметрам θ и μ при их гипотетических значениях

$$\nabla_\theta P(s_i = k|0,0) = -X_i \sum_{l=1}^{K_k} [f(C_{k,l,1}) - f(C_{k,l,0})];$$

$$\nabla_\mu P(s_i = k|0,0) = \sum_{l=1}^{K_k} [f(C_{k,l,1}) d_{k,l,1} - f(C_{k,l,0}) d_{k,l,0}],$$

где f – производная функции распределения F в соответствующей точке. Если записать разложение в ряд Тейлора для (7), подставить результат в (6) и раскрыть полученное произведение, выделяя члены первого порядка, то при тех же условиях на F получим

$$\nabla_\theta P(s|0,0) = -P(s|0,0) A_1 \sum_{i=1}^{N} X_i R_1(s_i); \tag{8}$$

$$\nabla_\mu P(s|0,0) = P(s|0,0) A_2 \sum_{i=1}^{N} R_2(s_i),$$

где $R_1(k)$ – скаляр, а $R_2(k)$ – вектор размерности M:

$$R_1(k) = (A_1 p_k)^{-1} \sum_{l=1}^{K_k} [f(C_{k,l,1}) - f(C_{k,l,0})]; \tag{9}$$

$$R_2(k) = (A_2 p_k)^{-1} \sum_{l=1}^{K_k} [f(C_{k,l,1}) d_{k,l,1} - f(C_{k,l,0}) d_{k,l,0}].$$

Здесь предполагается, что если $p_k = 0$, то $R_j(k) = 0$. Использование в (9) нормирующих постоянных A_1 и A_2 во многих случаях позволяет добиться того, чтобы значения $R_j(k)$ зависели только от вероятностей p_k, и именно на этих значениях строить анализ модели (1), перейдя, таким образом, к новому пространству признаков. В связи с этим новые признаки $R_j(k)$ далее будем называть метками множеств разбиения и будем говорить о наборах меток. Так, метки R_1 образуют один набор $\{R_1(1), \ldots, R_1(K)\}$, а метки R_2 образуют M наборов $\{R_{2,m}(1), \ldots, R_{2,m}(K)\}$, $m = \overline{1, M}$.

Важным свойством наборов меток является их нулевое математическое ожидание при гипотезе. Однако иногда вместо (9) приходится

использовать модифицированные наборы меток. Модификация меток требуется в случаях, когда путем их масштабирования с помощью констант A_j не удается добиться, чтобы все значения $R_j(k)$ зависели только от априорно известного набора вероятностей $\overline{p}$. В некоторых случаях мотивом для модификации меток может служить не незнание f, а необходимость придать тесту определенные свойства, например свойство локальной несмещенности. Тогда приходится заменять в метках набор вероятностей $\overline{p}$, несмотря на то, что он известен. Чтобы отличать метки, полученные по формулам (9), от возможно модифицированных меток, будем везде далее в первом случае употреблять обозначения R_j, а во втором – B_j. Если модификация не потребовалась, то эти обозначения относятся к одному и тому же набору меток.

3. Синтез тестов

Выбор критической области $S_1 \in S$ имеет целью улучшить поведение функции мощности теста, по крайней мере локально, в окрестности гипотезы. Обычно это приводит к тесту вида $H(s) > \text{const}$, где для отбора точек можно использовать различные критерии $H(s)$. Среди них прежде всего назовем локальный критерий средней кривизны функции мощности [1]:

$$\operatorname{tr}\left(\nabla^2 P(s|0,0)\right) > \text{const}, \tag{10}$$

который основан на анализе локальных свойств матрицы вторых производных $\nabla^2 P(S_1|0,0)$ при условии локальной несмещенности теста, т.е. когда $\nabla P(S_1|0,0) = 0$. Для случая знакового подхода, когда разбиение $C = \{(-\infty,0),0,(0,+\infty)\}$ не является параметризованным и $\overline{p}=\{\frac{1}{2},0,\frac{1}{2}\}$, с помощью критерия (10) удается получить соответствующий тест, как это сделано в [1, § 2.2], за счет естественного предположения $f'(0)=0$:

$$\sum_{i=1}^{N}\sum_{j=1}^{N} X_i^{\top} X_j \operatorname{sign}(Y_i)\operatorname{sign}(Y_j) > \text{const}. \tag{11}$$

В общем случае, когда для описания априорной информации используются параметризованные разбиения (2), построение теста с помощью критерия (10) затрудняется по целому ряду причин. В частности, не всегда удается построить локально несмещенный тест. Чтобы преодолеть эту ситуацию, можно предложить другое правило выбора критической области, основанное не на вторых, а на первых производных функции правдоподобия (6):

$$\frac{\|\nabla_\theta P(s|0,0)\|^2}{A_1^2} + \frac{\|\nabla_\mu P(s|0,0)\|^2}{A_2^2} > \text{const}. \tag{12}$$

Далее критерий (12) с использованием (9) записывается в виде

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left(X_i^{\top}X_jR_1(s_i)R_1(s_j)+R_2^{\top}(s_i)R_2(s_j)\right) > \text{const}. \qquad (13)$$

В знаковом анализе, когда разбиение $C=\{(-\infty,0),0,(0,\infty)\}$ не зависит от параметров и $\overline{p}=\{\frac{1}{2},0,\frac{1}{2}\}$, получим, что $R_2=0$. Поэтому если в (8) положить $A_1=-2f(0)$, то $R_1(1)=-1$, $R_1(2)=0$ и $R_1(3)=1$, т.е. $R_1(s_i)=\operatorname{sign}(Y_i)$ п.н. Таким образом, (13) представляется в виде (11). Иными словами, критерии (10) и (12) в знаковом анализе приводят к одному и тому же тесту. В общем случае можно показать, что критерии (10) и (12) при некоторых дополнительных условиях асимптотически эквивалентны.

4. Распределение индикаторных статистик

Запишем тест (13) в векторном виде

$$H(s)=N\begin{pmatrix}\overline{XB}_1\\ \overline{B}_2\end{pmatrix}^{\top}\begin{pmatrix}\overline{XB}_1\\ \overline{B}_2\end{pmatrix} > t(\alpha), \qquad (14)$$

где $\overline{XB}_1=\frac{1}{N}\sum\limits_{i=1}^{N}X_iB_1(s_i)$, $\overline{B}_2=\frac{1}{N}\sum\limits_{i=1}^{N}B_2(s_i)$, и рассмотрим вопрос о вычислении процентных точек $t(\alpha)$ на основе точного и предельного распределения индикаторной статистики $H(s)$.

При гипотезе для $H(s)$ можно найти точное распределение, так как число элементов пространства признаков $\mathcal{S}$ является конечным и равно K^N. Для больших объемов выборки можно воспользоваться асимптотическим распределением. Чтобы его получить, рассмотрим случайный вектор $\xi_N=N^{1/2}(\overline{XB}_1,\overline{B}_2)^{\top}$ размерности $T+M$ и введем следующие обозначения: $d_1^2=E\{B_1^2\}=\sum\limits_{k=1}^{K}B_1^2(k)p_k$, $D_2=E\{B_2B_2^{\top}\}$, $C=E\{B_1B_2\}$, $\overline{XX}=\frac{1}{N}\sum\limits_{i=1}^{N}X_iX_i^{\top}=\frac{1}{N}XX^{\top}$, $\overline{X}=\frac{1}{N}\sum\limits_{i=1}^{N}X_i$. Для последовательности матриц плана X будем предполагать существование предельной матрицы вторых моментов $V_X=\lim\limits_{N\to\infty}\overline{XX}$ и предельного вектора средних $E_X=\lim\limits_{N\to\infty}\overline{X}$. Кроме того, записывая $A_N<\infty$ для некоторой зависящей от N квадратной матрицы A_N, будем иметь в виду, что все ее элементы ограничены равномерно по N. Иными словами, существует

постоянная K, такая, что $|[A_N]_{i,j}| < K\ \forall i, j, N$, где $[\cdot]_{i,j}$ – элемент матрицы. Будем также обозначать знаком $\sim$ асимптотическую эквивалентность случайных величин (в смысле равенства их предельных распределений).

Лемма 1. *Если* $0 < V_X < \infty$, $d_1^2 > 0$ *и* $D_2 > 0$, *то при гипотезе* (4) *предельное распределение случайного вектора* ξ_N *является нормальным с нулевым средним и матрицей ковариаций*

$$V = \begin{pmatrix} d_1^2 V_X & E_X C^\top \\ C E_X^\top & D_2 \end{pmatrix}.$$

Доказательство. В силу того, что среднее значение меток B_1 и B_2 при гипотезе равно нулю, вектор ξ_N при гипотезе тоже будет иметь нулевое математическое ожидание $E\{\xi_N\} = 0$. Если учесть независимость погрешностей в модели (1), то матрица ковариаций для ξ_N запишется в виде

$$V_N = E\left\{\xi_N \xi_N^\top\right\} = \begin{pmatrix} d_1^2 \overline{XX} & \overline{X} C^\top \\ C \overline{X}^\top & D_2 \end{pmatrix}.$$

По условиям леммы $0 < V_X < \infty$, откуда следует, что $\|E_X\| < \infty$ и далее, с учетом условий $d_1^2 > 0$ и $D_2 > 0$, получаем, что $0 < V < \infty$. Отсюда в силу центральной предельной теоремы следует утверждение леммы.

Из леммы 1 вытекает целый ряд следующих ниже утверждений, доказательства которых аналогичны известным классическим результатам [2], поэтому здесь не приводятся.

Следствие 1. *В условиях леммы 1 при гипотезе имеет место*

$$V_N^{-1/2} \xi_N \sim V^{-1/2} \xi_N \sim \eta,$$

где η *– стандартный нормальный случайный вектор размера* $T+M$.

Замечание 1. Матрицы $V_N^{1/2}$ и $V^{1/2}$ не обязательно являются симметричными, а могут быть взяты, например, в верхнем диагональном виде. В любом случае, везде далее мы предполагаем, что квадратный корень всякой симметричной положительно определенной матрицы A строится так, чтобы $(A^{1/2})^\top A^{1/2} = A$, но не обязательно $A^{1/2}(A^{1/2})^\top = A$.

Теорема 1. *В условиях леммы 1 для статистики теста* (14) *при гипотезе имеет место*

$$H(s) = \xi_N^\top \xi_N \sim \sum_{k=1}^{T+M} \lambda_{N,k} \eta_k^2 \sim \sum_{k=1}^{T+M} \lambda_k \eta_k^2, \qquad (15)$$

где $\lambda_{N,k}$ и λ_k – собственные числа матриц V_N и V соответственно, а η_k – стандартные независимые гауссовские случайные величины.

Замечание 2. На практике для определения порогов $t(\alpha)$ теста (14) с помощью (15) можно воспользоваться собственными числами $\lambda_{N,k}$ матрицы V_N, если она не вырождена.

Теорема 2. *В условиях леммы 1 при гипотезе имеет место*

$$\xi_N^{\top} V_N^{-1} \xi_N \sim \xi_N^{\top} V^{-1} \xi_N \sim \chi^2_{T+M},$$

где χ^2_{T+M} – случайная величина, имеющая распределение хи-квадрат с $T+M$ степенями свободы.

Замечание 3. Теорема 2 дает основание для использования вместо теста (14) его нормированного варианта, более удобного с точки зрения вычисления процентных точек в асимптотическом случае:

$$H_1(s) = N \begin{bmatrix} \overline{XB}_1 \\ \overline{B}_2 \end{bmatrix}^{\top} \begin{bmatrix} d_1^2 \overline{XX} & \overline{X} C^{\top} \\ C\overline{X}^{\top} & D_2 \end{bmatrix}^{-1} \begin{bmatrix} \overline{XB}_1 \\ \overline{B}_2 \end{bmatrix} > t_1(\alpha), \qquad (16)$$

где $t_1(\alpha)$ – процентная точка распределения хи-квадрат с $T+M$ степенями свободы. Кроме того, для теста (16), как и для (14), можно построить точное распределение.

Замечание 4. Будем рассматривать еще один тест, который в отличие от (16) не требует дополнительных условий для своей локальной несмещенности (см. следующий раздел). Это тест с частично нормированной статистикой:

$$H_0(s) = N \left(d_1^{-2} \overline{XB}_1^{\top} \overline{XX}^{-1} \overline{XB}_1 + \overline{B}_2^{\top} D_2^{-1} \overline{B}_2 \right) > t_0(\alpha). \qquad (17)$$

При умеренных объемах выборки можно найти точное распределение статистики теста (17). Чтобы найти предельное распределение этой статистики, представим ее в виде $(A_N^{-1/2}\xi_N)^{\top}(A_N^{-1/2}\xi_N)$, где матрица A_N является блочно-диагональной: $A_N = \operatorname{diag}(d_1^2 \overline{XX}, D_2)$. Введем также предельную матрицу $A = \operatorname{diag}(d_1^2 V_X, D_2)$.

Теорема 3. *В условиях леммы 1 для статистики теста* (17) *при гипотезе имеет место*

$$H_0(s) = \xi_N^{\top} A_N^{-1} \xi_N \sim (1 - v_N\rho)\eta_1^2 + (1 + v_N\rho)\eta_2^2 + \sum_{i=3}^{T+M} \eta_i^2 \sim$$

$$\sim \xi_N^{\top} A^{-1} \xi_N \sim (1 - v\rho)\eta_1^2 + (1 + v\rho)\eta_2^2 + \sum_{i=3}^{T+M} \eta_i^2,$$

где $\eta_1, \ldots, \eta_{T+M}$ *– независимые случайные величины, имеющие стандартное нормальное распределение,* $\rho^2 = (C^\top D_2^{-1} C)/d_1^2$ *– квадрат множественного коэффициента корреляции наборов меток* B_1 *и* B_2, $v_N^2 = \overline{X}^\top \overline{XX}^{-1} \overline{X}$, $v^2 = E_X^\top V_X^{-1} E_X$.

Доказательство. По лемме 1 предельное распределение вектора ξ_N является непрерывным, а матрица A невырождена. Поэтому, если учесть, что равенство $A = \lim\limits_{N\to\infty} A_N$ имеет место для неслучайных матриц, то соотношение $\xi_N^\top A_N^{-1} \xi_N \sim \xi_N^\top A^{-1} \xi_N$ становится очевидным. Далее, по условиям леммы 1 матрица A_N невырождена для достаточно больших N, поэтому случайный вектор $A_N^{-1/2} \xi_N$ при гипотезе имеет нулевое среднее и матрицу ковариации

$$A_N^{-1/2} V_N (A_N^{-1/2})^\top = \begin{pmatrix} I_T & G_N^\top \\ G_N & I_M \end{pmatrix}, \tag{18}$$

где $G_N = D_2^{-1/2} C (\overline{XX}^{-1/2} \overline{X})^\top / d_1$, а через I_m обозначена единичная матрица размера $m \times m$. Чтобы найти собственные значения матрицы (18), применим тождества

$$\begin{vmatrix} 1 & b^\top \\ b & Q \end{vmatrix} = |Q|(1 - b^\top Q^{-1} b), \quad |Q + ab^\top| = |Q|(1 + a^\top Q^{-1} b), \tag{19}$$

которые выполняются для любых векторов a, b и невырожденных матриц Q. Запишем характеристический многочлен в виде

$$P(\lambda) = (1-\lambda)^{T+M} \begin{vmatrix} I_T & G_N^\top/(1-\lambda) \\ G_N/(1-\lambda) & I_M \end{vmatrix},$$

пользуясь тем, что возникающая при делении неопределенность в точке $\lambda = 1$ будет разрешена за счет непрерывности. Для j-го столбца матрицы G_N имеем $G_{N,j} = D_2^{-1/2} C [\overline{XX}^{-1/2} \overline{X}]_j / d_1$, где $[\cdot]_j$ – соответствующий элемент вектора. Поэтому применение T раз первого из тождеств (19) к полученному характеристическому многочлену позволяет получить

$$P(\lambda) = (1-\lambda)^{T+M} \left| I_M - (1-\lambda)^{-2} \sum_{j=1}^{T} G_{N,j} G_{N,j}^\top \right| =$$

$$= (1-\lambda)^{T+M} \left| I_M - \frac{D_2^{-1/2} C (D_2^{-1/2} C)^\top}{d_1^2 (1-\lambda)^2} v_N^2 \right|.$$

Отсюда, применяя второе из тождеств (19), получаем, что

$$P(\lambda) = (1-\lambda)^{T+M-2} [(1-\lambda)^2 - v_N^2 \rho^2].$$

Таким образом, все собственные значения матрицы (18) равны единице, кроме двух из них, равных $1-v_N\rho$ и $1+v_N\rho$, что и доказывает теорему.

5. Условия локальной несмещенности

Еще одно желательное свойство критической области $\mathcal{S}_1$ – локальная несмещенность – определяется как $\nabla P(\mathcal{S}_1|0,0) = 0$. Это требование, с учетом (8), выражается следующей парой равенств:

$$\begin{aligned} \nabla_\theta P(\mathcal{S}_1|0,0) &= -A_1 \sum_{s\in\mathcal{S}_1} P(s|0,0) \sum_{i=1}^{N} X_i R_1(s_i) = 0; \\ \nabla_\mu P(\mathcal{S}_1|0,0) &= A_2 \sum_{s\in\mathcal{S}_1} P(s|0,0) \sum_{i=1}^{N} R_2(s_i) = 0. \end{aligned} \tag{20}$$

При этом можно говорить отдельно о несмещенности как по параметрам θ модели (1), так и по параметрам μ разбиения (2), или даже по каждому из параметров θ_t или μ_m в отдельности. Обозначим множество $\{1,\ldots,K\}$ номеров элементов разбиения через $\mathcal{K}$ и сформулируем достаточные условия локальной несмещенности построенных тестов.

Теорема 4. *Если на множестве $\mathcal{K}$ существует биекция g, для которой имеет место*

$$\begin{aligned} &\forall\, k \in \mathcal{K}\setminus\mathcal{K}_g: \quad p_k = p_{g(k)}, \quad R_1(k) = -R_1(g(k)); \\ &\forall\, k_1,k_2 \in \mathcal{K}: \quad B_2^{\mathsf{T}}(k_1)B_2(k_2) = B_2^{\mathsf{T}}(g(k_1))B_2(g(k_2)), \end{aligned} \tag{21}$$

то тесты (14) *и* (17) *являются локально несмещенными по параметрам θ модели* (1).

Доказательство. При выполнении первых двух условий в (21) из (6) следует, что $P(s|0,0) = P(g(s)|0,0)$, а учитывая (8), получаем $\nabla_\theta P(s|0,0) = -\nabla_\theta P(g(s)|0,0)$. Последнее условие в (21) приводит к тому, что наборы признаков s и $g(s)$ будут иметь одинаковую степень предпочтения как по критерию $H(s)$, так и по критерию $H_0(s)$. Поэтому s и $g(s)$ оба будут входить в критическую область $\mathcal{S}_1$ тестов (14) и (17) либо оба не будут в нее входить. Следовательно, в выражении для $\nabla_\theta P(\mathcal{S}_1|0,0)$ из (20) все слагаемые разбиваются на пары противоположных по величине значений, что обеспечивает локальную несмещенность тестов (14) и (17) по θ.

Замечание 5. Условия (21) обеспечивают несмещенность сразу по всем параметрам θ модели (1). Условия, аналогичные (21), можно получить и для локальной несмещенности тестов (14) и (17) по каждому

из параметров μ_m, $m = \overline{1,M}$, в отдельности. Эти условия даются следующей теоремой, которая доказывается аналогично теореме 4.

Теорема 5. *Если на множестве $\mathcal{K}$ существует биекция g_m, для которой имеет место*

$$\begin{array}{l} \forall\ k \in \mathcal{K} \setminus \mathcal{K}_{g_m}: \quad p_k = p_{g_m(k)}, \quad R_{2,m}(k) = -R_{2,m}(g_m(k)), \\ \forall\ k_1, k_2 \in \mathcal{K}: \quad B_1(k_1)B_1(k_2) = B_1(g_m(k_1))B_1(g_m(k_2)), \end{array} \tag{22}$$

то тесты (14) *и* (17) *являются локально несмещенными по параметру μ_m разбиения* (2).

Замечание 6. Если одновременно выполняются условия (21) теоремы 4 и (22) теоремы 5 для всех $m = \overline{1,M}$, то тесты (14) и (17) будут локально несмещенными по всем параметрам.

Замечание 7. Для локальной несмещенности нормированного теста (16) по параметрам θ и μ, кроме условий (21) и (22), требуется еще одно условие: либо $\overline{X} = 0$, либо $C = 0$. Действительно, если оба эти равенства не выполняются, то для того, чтобы наборы меток s, $g(s)$ и $g_m(s)$ оценивались одинаково критерием (16), кроме (21) и (22), должны выполняться равенства

$$\forall k_1, k_2 \in \mathcal{K},\ m \in \overline{1,M}: \quad B_1(k_1)B_2(k_2) = B_1(g(k_1))B_2(g(k_2)) = \\ = B_1(g_m(k_1))B_2(g_m(k_2)),$$

однако с учетом (21) и (22) это равносильно тому, что $C = 0$. Отметим, что выполнение условия $\overline{X}{=}0$ является скорее исключением и может быть достигнуто только в случае, когда в модели (1) отсутствует свободный параметр и имеется редкая возможность в план эксперимента X специально включать "балансирующее" наблюдение $X_N = \sum\limits_{i=1}^{N-1} X_i$. Поэтому основным условием несмещенности теста (16) является некоррелированность меток $C = 0$. В любом случае выполнение хотя бы одного из условий $\overline{X} = 0$ или $C = 0$ приводит к тому, что статистики тестов (16) и (17) совпадают и в условиях леммы 1 имеют предельное распределение χ^2 с $T{+}M$ степенями свободы.

6. Применение принципа локальной оптимальности по Рою

Рассмотрим еще один подход к построению тестов для проверки гипотезы (4), который использует принцип оптимальности Роя [3]. Суть этого подхода состоит в том, чтобы найти тест, локально наиболее мощный против любой из скалярных линейных односторонних альтернатив

в пространстве параметров, который бы рассматривал все эти частные альтернативы как одинаково возможные (равноправные).

Для описания частных альтернатив введем линейное полупространство, параметризованное скалярным параметром t (луч, исходящий из точки гипотетических значений параметров $(\theta,\mu)=0$ в направлении ψ):

$$\Psi(\psi)=\left\{(\theta,\mu)^{\top}\in\mathbf{R}^{T+M}:(\theta,\mu)^{\top}=t\psi,\quad t>0\right\}. \tag{23}$$

Вектор направления $\psi\neq 0$ состоит из двух частей: $\psi=(\psi_1,\psi_2)^{\top}$, где $\psi_1=(\psi_{1,1},\ldots,\psi_{1,T})^{\top}$ и $\psi_2=(\psi_{2,1},\ldots,\psi_{2,M})^{\top}$, отвечающих за параметры $\theta=(\theta_1,\ldots,\theta_T)^{\top}$ модели (1) и параметры $\mu=(\mu_1,\ldots,\mu_M)^{\top}$ разбиения (2) соответственно.

Для каждого частного множества альтернатив $\Psi(\psi)$ может быть поставлена задача проверки гипотезы $H_0:t=0$ о скалярном параметре t против односторонней альтернативы $H_1:t>0$. Если обозначить $P_{\psi}(s|t)=P(s|t\psi_1,t\psi_2)$, то $\frac{d}{dt}P_{\psi}(s|0)=\psi^{\top}\nabla P(s|0,0)$. Как известно, для проверки гипотезы о скалярном параметре против односторонней альтернативы локально наиболее мощным будет являться тест с критической областью

$$\mathcal{S}_1(\psi)=\left\{s:\frac{d}{dt}P_{\psi}(s|0)=\psi^{\top}\nabla P(s|0,0)>\mathrm{const}\right\}. \tag{24}$$

Действительно, критическая область $\mathcal{S}_1(\psi)$ при фиксированном уровне значимости $\alpha=\sum\limits_{s\in\mathcal{S}_1(\psi)}P_{\psi}(s|0)$ должна выбираться так, чтобы функция мощности $\sum\limits_{s\in\mathcal{S}_1(\psi)}P_{\psi}(s|t)$ возрастала в окрестности точки $t=0$ с максимальной скоростью.

Добиваясь равноправия частных альтернатив, представим множество альтернатив задачи (4) в виде объединения линейных однопараметрических односторонних альтернатив вида (23): $\bigcup\limits_{\nu\in O}\Psi(W_N^{\top}\nu)$, где $O=\{\nu\in\mathbf{R}^{T+M}:\|\nu\|=1\}$ – сфера единичного радиуса, W_N – некоторая положительно определенная матрица. Это эквивалентно рассмотрению некоторого эллипсоида в пространстве параметров. Далее для каждого отдельного направления (т.е. при фиксированном $\psi=W_N^{\top}\nu$) рассмотрим оптимальный тест вида (24) против альтернативы $\Psi(W_N^{\top}\nu)$ и зададим для него порог $t/h_N(\nu)$, где h_N – некоторая положительная скалярная функция. Цель такого специального задания порога – добиться одинакового уровня значимости для всех частных тестов. Если

этого удастся достичь хотя бы в асимптотическом смысле, то тест

$$\max_{\nu \in O} \{h_N(\nu)\nu^\top W_N \nabla P(s|0,0)\} > t \tag{25}$$

можно рассматривать как реализацию права вето при голосовании равноправных участников за принятие гипотезы. При этом в голосовании принимают участие все тесты, локально наиболее мощные против скалярных линейных альтернатив.

Таким образом, тест (25) будем называть локально оптимальным (или асимптотически локально оптимальным) по Рою, если можно подобрать положительно определенную матрицу W_N и положительную скалярную функцию h_N, определенную на единичной сфере в $\mathbf{R}^{T+M}$ так, чтобы тесты $h_N(\nu)\nu^\top W_N \nabla P(s|0,0) > t$, будучи примененными против соответствующих частных одномерных альтернатив $\Psi(W_N^\top \nu)$ при фиксированном пороге t, для всех $\nu \in O$ имели одинаковые (или сколь угодно близкие при достаточно больших N) уровни значимости.

Замечание 8. Если рассматривается асимптотическая локальная оптимальность по Рою, то на последовательность матриц W_N следует наложить условие существования положительно определенной предельной матрицы $W = \lim\limits_{N\to\infty} W_N$, $0 < W < \infty$. Это необходимо для того, чтобы придать асимптотический смысл последовательности множеств $\{W_N^\top \nu : \nu \in O\}$ направлений в пространстве параметров.

Теорема 6. *В условиях леммы 1 нормированный тест* (16) *является асимптотически локально оптимальным по Рою.*

Доказательство. Учитывая (8), можно заметить, что $\nabla P(s|0,0) = \sqrt{N} P(s|0,0) A \xi_N$, где $A = \operatorname{diag}(-A_1 I_T, A_2 I_M)$.

Далее, задав $h_N(\nu) = (\sqrt{N} p^N)^{-1}$ и $W_N = V_N^{-1/2} A^{-1}$, с учетом того, что $P(s|0,0) = p^N$ п.н., получим оптимальный тест против частной альтернативы $\Psi((V_N^{-1/2} A^{-1})^\top \nu)$ в виде $\nu^\top V_N^{-1/2} \xi_N > t$. С учетом утверждения леммы 1 и в силу того, что $\nu^\top \nu = 1$, статистика $\nu^\top V_N^{-1/2} \xi_N$ будет иметь стандартное нормальное предельное распределение вне зависимости от ν. Таким образом, тесты, построенные против каждой из частных альтернатив с использованием заданных выше h_N и W_N, будут иметь асимптотически одинаковые уровни значимости. Теперь остается только найти конкретную форму теста (25) для нашего случая. Величина $\nu^\top V_N^{-1/2} \xi_N$ достигает своего максимума среди всех $\nu \in O$ при $\nu = V_N^{-1/2} \xi_N / \| V_N^{-1/2} \xi_N \|$, поэтому тест (25) можно записать в виде $\| V_N^{-1/2} \xi_N \| > t$, что эквивалентно $\xi_N^\top V_N{}^{-1} \xi_N > t^2$ и совпадает с (16).

7. Примеры

Приведем наиболее важные примеры, когда априорная информация о распределении погрешностей может быть описана моделью (2), (3). Пример 1 касается уже обсуждавшейся знаковой модели, в примерах 2–4 рассматриваются разбиения с одним параметром, в примерах 5–7 – разбиения с двумя параметрами. Во всех примерах предполагается, что при проверке гипотезы выполнены условия симметрии распределения погрешностей на границах гипотетического разбиения, т.е. $F(C_{k,l,j}) = 1-F(-C_{k,l,j})$ и $f(C_{k,l,j}) = f(-C_{k,l,j})$.

Пример 1. Для знаковой модели погрешностей разбиение (2) состоит из трех элементов $C = \{(-\infty, 0), 0, (0, +\infty)\}$, оно не является параметризованным, а распределение (3) на нем задано в виде $\overline{p}\{\frac{1}{2}, 0, \frac{1}{2}\}$. Все рассмотренные методы построения тестов приводят к (11), при этом $A_1 = -2f(0)$, метки $R_1 = \{-1, 0, 1\}$ и $R_2 = 0$ не требуют модификации и тест (11) является локально несмещенным по параметрам θ и асимптотически локально оптимальным по Рою.

Пример 2. Априорная информация об интерквантильном размахе уровня 1/3 может быть описана с помощью разбиения с одним положительным параметром $C(\mu') = \{(-\infty, -\mu'), [-\mu', \mu'], (\mu', +\infty)\}$ и вероятностями $\overline{p} = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$. Положив $A_1 = A_2 = -3f(\mu_0)$, из (9) получаем метки $R_1 = \{-1, 0, 1\}$, $R_2 = \{1, -2, 1\}$. При этом наборы меток R_1 и R_2 некоррелированы ($C = 0$) и имеют дисперсии $d_1^2 = 2/3$, $d_2^2 = 2$. Наконец, для биекции $g : \{1, 2, 3\} \to \{3, 2, 1\}$ на номерах элементов разбиения выполняются условия (21) локальной несмещенности по θ. Достаточные условия (22) локальной несмещенности по параметру μ' не выполняются, однако тест (17) является асимптотически локально оптимальным по Рою.

Пример 3. Информацию об интерквартильном размахе можно описать разбиением

$$C(\mu') = \{(-\infty, -\mu'), [-\mu', 0), 0, (0, \mu'], (\mu', +\infty)\}$$

с одним положительным параметром при $\overline{p} = \{\frac{1}{4}, \frac{1}{4}, 0, \frac{1}{4}, \frac{1}{4}\}$. Если положить $A_1 = A_2 = -4f(\mu_0)$, то получаем некоррелированные метки $R_1 = \{-1, -a, 0, a, 1\}$ и $R_2 = \{1, -1, 0, -1, 1\}$ с дисперсиями $d_1^2 = (1+a^2)/2$ и $d_2^2 = 1$. Условия (21) несмещенности по θ выполняются, если задать биекцию $g : \{1, 2, 3, 4, 5\} \to \{5, 4, 3, 2, 1\}$. Модифицируем метки, положив в них $a=1$. При этом все условия несмещенности по θ сохраняются. Кроме того, для биекции $g_1 : \{1, 2, 3, 4, 5\} \to \{2, 1, 3, 5, 4\}$ условия (22) несмещенности по μ' тоже будут выполнены. Таким образом, чтобы добиться

несмещенности по μ' и по θ, можно использовать модифицированные метки $B_1 = \{-1,-1,0,1,1\}$ и $B_2 = \{1,-1,0,-1,1\}$, однако при этом (14) будет тестом максимального правдоподобия, а тест (17) асимптотически локально оптимальным по Рою только в случае, если $f(0) = 2f(\mu_0)$.

Пример 4. Рассмотрим случай, когда один из элементов разбиения $C(\mu') = \{(-\infty,-\mu') \bigcup (\mu',\infty), [-\mu',0), 0, (0,\mu']\}$ состоит из двух несмежных интервалов при наборе вероятностей $\overline{p} = \{\frac{1}{3}, \frac{1}{3}, 0, \frac{1}{3}\}$. Полагая $A_1 = 3(f(\mu_0) - f(0))$ и $A_2 = -3f(\mu_0)$, получим следующие некоррелированные наборы меток: $R_1 = \{0,-1,0,1\}$ и $R_2 = \{2,-1,0,-1\}$ с дисперсиями $d_1^2 = 2/3$ и $d_2^2 = 2$. Условия (21) несмещенности по параметрам θ выполняются для биекции $g : \{1,2,3,4\} \to \{1,4,3,2\}$. Как и в примере 2, условия (22) локальной несмещенности по μ' не выполняются, однако тест (17) является асимптотически локально оптимальным.

Пример 5. Рассмотрим вариант примера 2, когда разбиение имеет те же априорные вероятности $\overline{p} = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$, но зависит от двух параметров

$$C(\mu') = \{(-\infty,-\mu_1'), [-\mu_1', \mu_2'], (\mu_2', +\infty)\}$$

при $\mathcal{M} = \{(\mu_1', \mu_2')^{\mathsf{T}} \in \mathbf{R}^2 : -\mu_1' < \mu_2'\}$. Полагая $A_1 = A_2 = -3f(\mu_{20})$, можно получить наборы меток

$$R_1 = \{-1,0,1\},\ R_{2,1} = \{1,-1,0\} \text{ и } R_{2,2} = \{0,-1,1\}.$$

Здесь, как и в примере 2, $d_1^2 = 2/3$, но на этот раз все три набора меток коррелированы: $C^{\mathsf{T}} = (-\frac{1}{3}, \frac{1}{3})$ и $D_2 = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

Рассматривая биекции $g : \{1,2,3\} \to \{3,2,1\}$, $g_1 : \{1,2,3\} \to \{2,1,3\}$ и $g_2 : \{1,2,3\} \to \{1,3,2\}$, обнаруживаем, что в отличие от примера 2 выполняются не только условия (21), но и (22). Учитывая коррелированность меток, заключаем, что тесты (14) и (17) являются локально несмещенными, а тест (16) асимптотически локально оптимальным по Рою. Интересно, что за счет добавления параметра, по сравнению с примером 2, удается добиться локальной несмещенности по параметрам разбиения.

Пример 6. Модифицируя пример 3, при тех же вероятностях $\overline{p} = \{\frac{1}{4}, \frac{1}{4}, 0, \frac{1}{4}, \frac{1}{4}\}$ можно использовать два положительных параметра: $C(\mu') = \{(-\infty,-\mu_1'), [-\mu_1', 0), 0, (0,\mu_2'], (\mu_2', +\infty)\}$. Будем считать, что $f(0) = (a+1)f(\mu_{02})$. Полагая $A_1 = A_2 = -4f(\mu_{02})$, получим следующие наборы меток:

$$R_1 = \{-1,-a,0,a,1\},\ R_{2,1} = \{1,-1,0,0\} \text{ и } R_{2,2} = \{0,0,-1,1\}.$$

Положив $a = 1$ в R_1, получаем $B_1 = \{-1,-1,0,1,1\}$, при этом $d_1^2 = 1$, $D_2 = \frac{1}{2}I_2$, $C = 0$. В результате модификации получаем такой же эф-

факт, как и в примере 3: условия локальной несмещенности по θ и μ выполняются, однако при этом (14) будет тестом максимального правдоподобия, а тест (17) асимптотически локально оптимальным только в случае, если $f(0) = 2f(\mu_0)$.

Пример 7. Рассмотрим вариант примера 4 с теми же априорными вероятностями $\overline{p} = \{\frac{1}{3}, \frac{1}{3}, 0, \frac{1}{3}\}$, но не с одним, а с двумя положительными параметрами $C(\mu') = \{(-\infty, -\mu_1') \bigcup (\mu_2', \infty), [-\mu_1', 0), 0, (0, \mu_2']\}$. Положив $A_1 = 3(f(0) - f(\mu_{02}))$ и $A_2 = 3f(\mu_{02})$, получим коррелированные наборы меток $R_1 = \{0, -1, 0, 1\}$, $R_{2,1} = \{-1, 1, 0, 0\}$ и $R_{2,2} = \{-1, 0, 0, 1\}$, для которых $d_1^2 = 2/3$, $C^\top = \frac{1}{3}(-1, 1)$ и $D_2 = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

Условия (21) несмещенности по θ выполнены в этом примере при $g : \{1, 2, 3, 4\} \to \{1, 4, 3, 2\}$. Условия (22) несмещенности по параметрам μ' не выполняются, однако нормированный тест (17) является асимптотически локально оптимальным по Рою.

Замечание 9. В заключение отметим, что когда (1) является моделью парной линейной регрессии, в условиях примера 3 восстановление параметров θ и μ имеет целью найти полосу на корреляционном поле, в которой лежит около половины из всех выборочных точек, и еще по одной четверти точек находятся выше и ниже этой полосы. Для наблюдений, полученных в результате физических экспериментов, это условие означает, что используется априорное знание об одинаковой точности измерений в разных опытах. При этом понятие точности выражено в терминах интерквартильного размаха, что близко к тому, как трактуется точность измерений различных приборов. Аналогичную интерпретацию можно дать не только примеру 3, но и каждому из приведенных выше примеров.

СПИСОК ЛИТЕРАТУРЫ

1. *Болдин М. В., Симонова Г. И., Тюрин Ю. Н.* Знаковый статистический анализ линейных моделей. М.: Наука, Физматлит, 1997.
2. *Боровков А. А.* Математическая статистика. Новосибирск: Наука, 1997.
3. *Roy S. N.* Some aspects of multivariate analysis. New York: Wiley, 1957.

II. Робастные методы

В.В. Домбровский, Е.А. Ляшенко

Робастное управление дискретными системами со случайными параметрами и мультипликативными шумами

Рассматривается задача управления по квадратичному критерию дискретными стохастическими системами со случайными параметрами и аддитивными и мультипликативными шумами, зависящими от состояний и управлений. Получены уравнения для оптимального линейного динамического регулятора по выходу. Синтезированный регулятор обладает свойством робастности к виду распределения вектора случайных параметров.

1. Введение

Широкий класс реальных систем, параметры которых точно неизвестны и меняются случайным образом, описывается линейными моделями со случайными параметрами. Проблема синтеза регуляторов для таких систем рассматривалась во многих работах (подробный обзор которых дан в [1]), что обусловлено ее теоретической и практической важностью. В частности, в [1–4] рассматриваются системы со случайными скачкообразно меняющимися параметрами, принимающими конечное множество значений в соответствии с эволюцией дискретной марковской цепи с известными переходными вероятностями (системы случайной структуры или гибридные системы), возмущенные аддитивными и мультипликативными шумами.

В данной работе также рассматриваются системы со случайными параметрами, возмущенные аддитивными и мультипликативными шумами, однако в отличие от [1–4] предполагается, что параметры представляют собой последовательность независимых случайных величин, для которых известны только первый и второй моменты распределения. К такого класса системам можно отнести, например, системы с дважды стохастическими аддитивными и мультипликативными шумами (когда статистические характеристики шумов являются случайными величинами). Моделями с дважды стохастическими возмущениями описываются многие реальные объекты, в частности в финансовой математике [5, 6]. В классе линейных получены уравнения оптимального

динамического регулятора по выходу. Синтезированный регулятор обладает свойством робастности к виду распределения вектора случайных параметров.

Отметим, что задачи синтеза регуляторов для подобных систем со случайными параметрами, но без мультипликативных шумов, рассматривались в [7, 8].

2. Синтез динамического регулятора

Пусть объект управления описывается уравнениями

$$x(k+1) = \Big(A_0(k) + \sum_{j=1}^{p} A_j[\theta(k),k]v_j(k)\Big)x(k) +$$

$$+\Big(B_0(k) + \sum_{j=1}^{p} B_j[\theta(k),k]v_j(k)\Big)u(k) + D_1[\theta(k),k]w_1(k); \qquad (1)$$

$$y(k) = \Big(C_0(k) + \sum_{j=1}^{p} C_j[\theta(k),k]v_j(k)\Big)x(k) + D_2(k)w_2(k), \qquad (2)$$

где $x(k)$ – n-мерный вектор состояния, $y(k)$ – m-мерный вектор наблюдений ($m \leqslant n$), $u(k)$ – l-мерный вектор управления, $v(k)$, $w_1(k)$ и $w_2(k)$ – векторы белых шумов размерностей p, p_1 и p_2 с нулевыми средними и единичными матрицами ковариаций, причем

$$\mathbf{M}\{w_1(k)w_2^T(k)\} = 0,\ \mathbf{M}\{v(k)w_1^T(k)\} = 0,\ \mathbf{M}\{v(k)w_2^T(k)\} = 0;$$

$\theta(k)$ – последовательность независимых s-мерных случайных векторов с известными первым и вторым моментами:

$$\mathbf{M}\{\theta(k)\} = \bar{\theta}(k),\ \mathbf{M}\{\theta(k)\theta^T(k)\} = \Theta(k),\ \mathbf{M}\{\theta(k)v^T(k)\} = 0,$$

$$\mathbf{M}\{\theta(k)w_1^T(k)\} = 0,\ \mathbf{M}\{\theta(k)w_2^T(k)\} = 0.$$

$A_0(k)$, $B_0(k)$, $C_0(k)$, $D_1[\theta(k),k]$, $D_2(k)$, $A_j[\theta(k),k]$, $B_j[\theta(k),k]$ и $C_j[\theta(k),k]$ ($j = \overline{1,p}$) – матрицы соответствующих размерностей; $D_1[\theta(k),k]$, $A_j[\theta(k),k$ $B_j[\theta(k),k]$ и $C_j[\theta(k),k]$ ($j = \overline{1,p}$) зависят от $\theta(k)$ линейно, причем для любого фиксированного j ($j = \overline{1,p}$) $B_j[\theta(k),k]$ и $C_j[\theta(k),k]$ не могут быть одновременно ненулевыми; $v_j(k)$ – j-я компонента вектора $v(k)$; символ T означает транспонирование. Начальный вектор $x(0)$ предполагается случайным с известными характеристиками $\mathbf{M}\{x(0)\} = x^0$ и $\mathbf{M}\{x(0)x^T(0)\} = P_x^0 \geqslant 0$, предполагается также, что $\mathbf{M}\{x(0)v^T(k)\}=0$, $\mathbf{M}\{x(0)w_1^T(k)\}=0$, $\mathbf{M}\{x(0)w_2^T(k)\}=0$, $\mathbf{M}\{x(0)\theta^T(k)\}=0$, $k = 0, 1, 2, \ldots, N$.

Структуру динамического регулятора определим соотношениями

$$u(k) = K(k)\hat{x}(k); \tag{3}$$

$$\hat{x}(k+1) = A_0(k)\hat{x}(k) + B_0(k)u(k) + G(k)\left[y(k) - C_0(k)\hat{x}(k)\right], \tag{4}$$

где $\hat{x}(k)$ – несмещенная оценка вектора состояния системы (1), (2); $K(k)$, $G(k)$ – матрицы соответствующих размерностей.

Необходимо определить параметры регулятора – последовательности матриц $K(k)$, $G(k)$ $(k = 0, 1, \ldots, N)$, минимизирующие критерий

$$J = \mathbf{M}\left\{\sum_{k=0}^{N-1}\left[x^T(k)R_1(k)x(k) + u^T(k)R_2(k)u(k)\right] + \right.$$

$$\left. + x^T(N)R_1(N)x(N) + e^T(N)\Pi e(N)\right\}, \tag{5}$$

где $e(k) = x(k) - \hat{x}(k)$ – ошибка оценки вектора состояния, $R_1(k) \geqslant 0$, $R_2(k) > 0$ и Π – симметричные весовые матрицы, M – оператор математического ожидания. В дальнейшем будем использовать следующие обозначения: для любой матрицы $\Psi[\theta(k), k]$, зависящей от $\theta(k)$, $\overline{\Psi}(k) = \mathbf{M}\{\Psi[\theta(k), k]\}$ и $\widetilde{\Psi}(k) = \Psi[\theta(k), k] - \overline{\Psi}(k)$.

Теорема. *Оптимальные последовательности матриц $K(k)$, $G(k)$, минимизирующие функционал (5) на траекториях замкнутой системы (1)–(4), определяются системой матричных уравнений*

$$G(k) = \left[A_0(k)P_2(k)C_0^T(k) + \sum_{j=1}^{p}\overline{A}_j(k)P_1(k)\overline{C}_j^T(k) + \right.$$

$$\left. + \sum_{j=1}^{p}\mathbf{M}\left\{\widetilde{A}_j(k)P_1(k)\widetilde{C}_j^T(k)\right\}\right]\left[C_0(k)P_2(k)C_0^T(k) + \sum_{j=1}^{p}\overline{C}_j(k)P_1(k)\times\right.$$

$$\left.\times\,\overline{C}_j^T(k) + \sum_{j=1}^{p}\mathbf{M}\left\{\widetilde{C}_j(k)P_1(k)\widetilde{C}_j^T(k)\right\} + D_2(k)D_2^T(k)\right]^{-1}; \tag{6}$$

$$K(k) = -\left[B_0^T(k)Q_1(k+1)B_0(k) + \sum_{j=1}^{p}\overline{B}_j^T(k)[Q_1(k+1) + Q_2(k+1)]\overline{B}_j(k) + \right.$$

$$\left. + \sum_{j=1}^{p}\mathbf{M}\left\{\widetilde{B}_j^T(k)[Q_1(k+1) + Q_2(k+1)]\widetilde{B}_j(k)\right\} + R_2(k)\right]^{-1}\times$$

$$\times\left[B_0^T(k)Q_1(k+1)A_0(k) + \sum_{j=1}^{p}\overline{B}_j^T(k)[Q_1(k+1) + Q_2(k+1)]\overline{A}_j(k) + \right.$$

$$+\sum_{j=1}^{p}\mathbf{M}\left\{\widetilde{B}_j^T(k)[Q_1(k+1)+Q_2(k+1)]\widetilde{A}_j(k)\right\}\Big]; \tag{7}$$

$$P_1(k+1)=(A_0(k)+B_0(k)K(k))(P_1(k)-P_2(k))(A_0(k)+B_0(k)K(k))^T+$$

$$+A_0(k)P_2(k)A_0^T(k)+\sum_{j=1}^{p}\left\{\overline{A}_j(k)P_2(k)\overline{A}_j^T(k)+(\overline{A}_j(k)+\overline{B}_j(k)K(k))\times\right.$$

$$\left.\times(P_1(k)-P_2(k))(\overline{A}_j(k)+\overline{B}_j(k)K(k))^T\right\}+\sum_{j=1}^{p}\mathbf{M}\left\{(\widetilde{A}_j(k)+\right.$$

$$+\widetilde{B}_j(k)K(k))(P_1(k)-P_2(k))(\widetilde{A}_j(k)+\widetilde{B}_j(k)K(k))^T+$$

$$\left.+\widetilde{A}_j(k)P_2(k)\widetilde{A}_j^T(k)\right\}+\overline{D}_1(k)\overline{D}_1^T(k)+\mathbf{M}\left\{\widetilde{D}_1(k)\widetilde{D}_1^T(k)\right\}; \tag{8}$$

$$P_2(k+1)=A_0(k)P_2(k)A_0^T(k)-G(k)C_0(k)P_2(k)C_0^T(k)G^T(k)+$$

$$+\sum_{j=1}^{p}\left\{\overline{A}_j(k)P_2(k)\overline{A}_j^T(k)-G(k)\overline{C}_j(k)P_1(k)\overline{C}_j^T(k)G^T(k)+\right.$$

$$\left.+(\overline{A}_j(k)+\overline{B}_j(k)K(k))(P_1(k)-P_2(k))(\overline{A}_j(k)+\overline{B}_j(k)K(k))^T\right\}+$$

$$+\sum_{j=1}^{p}\mathbf{M}\left\{\widetilde{A}_j(k)P_2(k)\widetilde{A}_j^T(k)-G(k)\widetilde{C}_j(k)P_1(k)\widetilde{C}_j^T(k)G^T(k)+\right.$$

$$\left.+(\widetilde{A}_j(k)+\widetilde{B}_j(k)K(k))(P_1(k)-P_2(k))(\widetilde{A}_j(k)+\widetilde{B}_j(k)K(k))^T\right\}+$$

$$+\overline{D}_1(k)\overline{D}_1^T(k)+\mathbf{M}\left\{\widetilde{D}_1(k)\widetilde{D}_1^T(k)\right\}-G(k)D_2(k)D_2^T(k)G^T(k); \tag{9}$$

$$Q_1(k)=A_0^T(k)Q_1(k+1)A_0(k)-K^T(K)B_0^T(k)Q_1(k+1)B_0(k)K(k)+$$

$$+\sum_{j=1}^{p}\left\{\overline{A}_j^T(k)Q_1(k+1)\overline{A}_j(k)-K^T(k)\overline{B}_j^T(k)[Q_1(k+1)+Q_2(k+1)]\times\right.$$

$$\left.\times\overline{B}_j(k)K(k)+(\overline{A}_j(k)-G(k)\overline{C}_j(k))^TQ_2(k+1)(\overline{A}_j(k)-G(k)\overline{C}_j(k))\right\}+$$

$$+\sum_{j=1}^{p}\mathbf{M}\left\{\widetilde{A}_j^T(k)Q_1(k+1)\widetilde{A}_j(k)-K^T(k)\widetilde{B}_j^T(k)[Q_1(k+1)+\right.$$

$$\left.+Q_2(k+1)]\widetilde{B}_j(k)K(k)+(\widetilde{A}_j(k)-G(k)\widetilde{C}_j(k))^TQ_2(k+1)(\widetilde{A}_j(k)-\right.$$

$$-G(k)\widetilde{C}_j(k))\Big\} + R_1(k) - K^T(k)R_2(k)K(k); \qquad (10)$$

$$Q_2(k) = K^T(k)B_0^T(k)Q_1(k+1)B_0(k)K(k) + (A_0(k) - G(k)C_0(k))^T \times$$

$$\times Q_2(k+1)(A_0(k) - G(k)C_0(k)) + \sum_{j=1}^{p} K^T(k)\overline{B}_j^T(k)[Q_1(k+1) +$$

$$+Q_2(k+1)]\overline{B}_j(k)K(k) + \sum_{j=1}^{p} \mathrm{M}\Big\{K^T(k)\widetilde{B}_j^T(k)\left[Q_1(k+1) + Q_2(k+1)\right] \times$$

$$\times \widetilde{B}_j(k)K(k)\Big\} + K^T(k)R_2(k)K(k) \qquad (11)$$

с граничными условиями $P_1(0) = P_x^0$, $P_2(0) = P_x^0 - x^0\left(x^0\right)^T$, $Q_1(N) = R_1(N)$, $Q_2(N) = \Pi$.

Минимальное значение функционала (5)

$$J^* = \mathrm{tr}\Big\{\sum_{k=0}^{N-1}\left[\left(R_1(k) + K^T(k)R_2(k)K(k)\right)P_1(k) -\right.$$

$$\left.-K^T(k)R_2(k)K(k)P_2(k)\right] + R_1(N)P_1(N) + \Pi P_2(N)\Big\}.$$

Теорема дает необходимые условия оптимальности в виде дискретной двухточечной краевой задачи. Для стационарных систем в установившемся состоянии (при $k \to \infty$) параметры регулятора определяются стационарными решениями разностных уравнений (6)–(11).

СПИСОК ЛИТЕРАТУРЫ

1. *Малышев В. В., Пакшин П. В.* Прикладная теория стохастической устойчивости и оптимального стационарного управления (обзор). Ч. II // Изв. АН СССР. Техническая кибернетика. 1990. № 2. С. 97-119.
2. *Пакшин П. В.* Оптимальное линейное управление дискретными объектами при случайном скачкообразном изменении их параметров // Пробл. управления и теории информации. 1982. Т. 11. № 3. С. 53-61.
3. *Пакшин П. В.* Устойчивость дискретных систем со случайной структурой при постоянно действующих возмущениях // АиТ. 1983. № 6. С. 75-85.
4. *Hopkins W. R.* Optimal stabilization of families of linear differential equations with jump coefficients and multiplicative noise // SIAM J. Control and Optimiz. 1987. Vol 26. No. 6. P. 93-120.
5. *Merton R. C.* Continuous-time finance. Cambr. Ma. Blackwell, 1990.
6. *Мельников А. В., Волков С. Н., Нечаев М. Л.* Математика финансовых обязательств. М.: ГУ ВШЭ, 2001.

7. *Домбровский В. В., Чикунова Е. В.* Синтез динамических регуляторов пониженного порядка для систем со случайными параметрами // Изв. РАН. Теория и системы управления. 1998. №. 2. С. 97-101.
8. *Chen S., Yong J.* Stochastic linear guadratic optimal control problems // Applied Mathematics and Optimisation. 2001. Vol. 43. P. 21-45.

Б.Ю. Лемешко, С.С. Помадин

Корреляционный анализ многомерных случайных величин при нарушении предположений о нормальности [1]

1. Введение

В различных приложениях статистического анализа многомерных величин одну из ключевых позиций занимают задачи корреляционного анализа. В процессе их решения выявляются наличие и характер взаимосвязи величин, взаимозависимости величин при устранении влияния совокупности других или зависимости одной случайной величины от группы величин. Вычисляются оценки коэффициентов и матриц парной, частной и множественной корреляций, проверяются различные статистические гипотезы относительно параметров многомерного распределения и коэффициентов корреляции. На основании результатов корреляционного анализа может делаться вывод о наличии и характере функциональной зависимости или предпочтительности для описания исследуемого объекта регрессионной модели того или иного вида.

В основе классического аппарата корреляционного анализа лежит предположение о принадлежности наблюдаемого случайного вектора к многомерному нормальному закону. Базируясь на этом, получили предельные распределения статистик, используемых в корреляционном анализе. На практике предпосылки классического корреляционного анализа выполняются далеко не всегда. Поэтому возникает вопрос о справедливости выводов, сделанных при применении классического аппарата, при нарушении основного предположения.

Основной целью данной работы явилось стремление разобраться, что будет происходить с распределениями различных статистик корреляционного анализа, если наблюдаемый закон будет отличаться от

[1] Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 00-01-00913)

многомерного нормального. Очевидно, что ответить на поставленный вопрос, используя чисто аналитические методы, чрезвычайно сложно из-за нетривиальности возникающих задач. Поэтому в основу проводимого исследования положена развиваемая методика компьютерного анализа статистических закономерностей.

Для ответа на поставленный вопрос было предусмотрено решение следующих подзадач:

– моделирование отличных от нормального многомерных законов распределения;

– исследование эмпирических распределений статистик корреляционного анализа при наблюдении многомерного нормального закона;

– исследование распределений статистик корреляционного анализа при законах распределения наблюдаемых многомерных величин, отличающихся от нормального.

Второй пункт данного перечня задач был предусмотрен для подтверждения работоспособности методики при совпадении результатов моделирования распределений статистик с классическими.

2. Моделирование многомерных случайных величин

Моделирование многомерного нормального закона. Многомерное нормальное распределение случайного вектора $\overline{X} = \|X_1, X_2, \ldots, X_m\|^\top$ размерности m полностью определяется вектором математических ожиданий $\overline{M} = \|M_1, M_2, \ldots, M_m\|^\top$ и ковариационной матрицей $\Sigma = \|\sigma_{ij}\| = \|E[(X_i - M_i)(X_j - M_j)]\|$.

Функция плотности многомерного нормального закона имеет вид

$$f(\overline{X}) = \frac{1}{\sqrt{(2\pi)^m |\Sigma|}} e^{-\frac{1}{2}(\overline{X}-\overline{M})^\top \Sigma^{-1} (\overline{X}-\overline{M})}.$$

Пусть мы имеем совокупность случайных величин $\{Z_i\}, i = \overline{1,m}$, где Z_i подчиняется стандартному нормальному закону с параметрами (0,1). Тогда вектор $\overline{X}$, распределенный по многомерному нормальному закону с параметрами $\overline{M}$ и Σ, получается через линейное преобразование вида [1]

$$\overline{X} = A\overline{Z} + \overline{M}. \tag{1}$$

Обычно предполагают, что A является нижней треугольной матрицей

$$A = \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{bmatrix},$$

коэффициенты a_{ij} которой определяются рекуррентной процедурой

$$a_{ij} = \frac{\sigma_{ij} - \sum\limits_{k=1}^{j-1} a_{ik}a_{jk}}{\sqrt{\sigma_{jj} - \sum\limits_{k=1}^{j-1} a_{jk}^2}}, \quad 1 \le j \le i \le m. \tag{2}$$

Моделирование многомерного закона, отличного от нормального. Задача моделирования многомерных величин, распределенных по закону, отличному от нормального, с заданными математическим ожиданием и ковариационной матрицей в самой работе решается в соответствии с описанным выше алгоритмом. Однако совокупность $\{Z_i\}$, $i = \overline{1,m}$, формируется уже не по стандартному нормальному закону, а в соответствии с одним из одномерных законов распределения с нулевым математическим ожиданием и единичной дисперсией. Затем заданная матрица Σ раскладывается по формуле (2), и осуществляется преобразование (1). Тогда на выходе мы имеем некоторый многомерный закон, отличный от нормального, с известным математическим ожиданием и неизвестной ковариационной матрицей.

Для моделирования различных совокупностей $\{Z_i\}, i = \overline{1,m}$, было решено использовать двустороннее экспоненциальное распределение (экспоненциальное семейство распределений) с плотностью

$$f(x) = \frac{\lambda}{2\sqrt{2}\theta_1\Gamma(\frac{1}{\lambda})} \exp\left(-\left(\frac{|x-\theta_0|}{\sqrt{2}\theta_1}\right)^{\lambda}\right),$$

где λ – параметр формы, так как оно охватывает целый класс симметричных распределений. Его частными случаями являются распределение Лапласа (параметр формы $\lambda = 1$), нормальное ($\lambda = 2$), предельными случаями – распределение Коши ($\lambda \to 0$) и равномерное ($\lambda \to +\infty$). Таким образом, с помощью параметра формы мы можем задавать непрерывное "удаление" наблюдаемого многомерного закона от нормального, как в сторону увеличения параметра ($\lambda > 2$), так и в сторону уменьшения параметра ($0 < \lambda < 2$). При параметре формы $\lambda = 2$ будут формироваться псевдослучайные векторы $\overline{X}$ в соответствии с нормальным законом.

Рис. 1 иллюстрирует изменение функции плотности семейства экспоненциальных распределений при изменении параметра формы от 0,5 до 10.

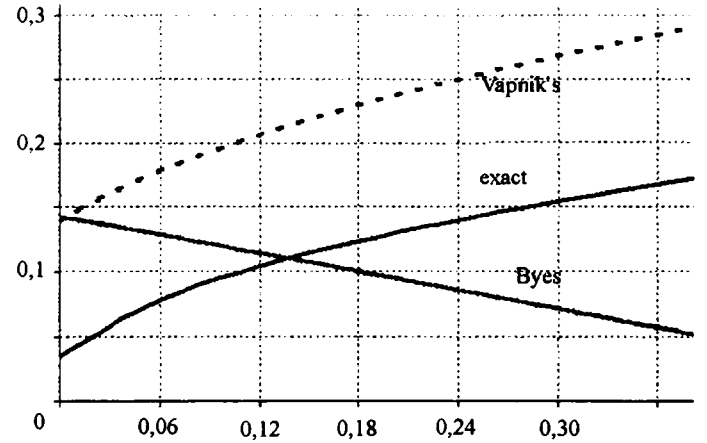


Рис. 1. Функции плотности двухстороннего экспоненциального распределения при различных параметрах формы

К сожалению, мы не можем моделировать многомерный закон с некоторой произвольной функцией распределения, с заданными математическим ожиданием и ковариационной матрицей, на "заданном" расстоянии (определяемом в смысле некоторой меры) от многомерного нормального. Однако мы можем построить распределение, отличающееся от нормального (в соответствии с процессом моделирования), с известными математическим ожиданием и ковариационной матрицей, определяемыми на основании исследования свойств полученного многомерного датчика (при заданных $\overline{M}$, Σ и λ). Для определения "истинной"

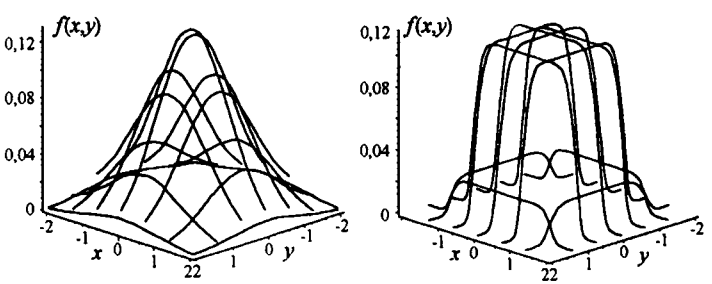


Рис. 2. Смоделированные плотности двумерного закона при $\lambda = 2$ (слева) и $\lambda = 10$ (справа)

ковариационной матрицы моделируемого многомерного закона использовались оценки максимального правдоподобия (ОМП), усредняемые по множеству проведенных экспериментов.

Таким образом решалась задача не по моделированию закона с заданными математическим ожиданием и ковариационной матрицей, а задача по моделированию закона с некоторыми параметрами $\overline{M}$ и Σ, истинные значения которых уточнялись в процессе исследования многомерного датчика. На рис. 2 приведен вид функций плотности, полученных при моделировании двумерных векторов при $\lambda = 2$ (плотность нормального закона, слева) и при $\lambda = 10$ (справа). Как видим, во втором случае полученное плосковершинное распределение существенно отличается от нормального.

3. Задачи проверки статистических гипотез классического корреляционного анализа

Пусть $\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_n$ – выборка m-мерного случайного вектора объема n; $\overline{M} = [M_i]_{i=1}^m$ – вектор математического ожидания случайного вектора $\overline{X}$; $\Sigma = [\sigma_{ij}]_{i,j=1}^m$ – ковариационная матрица случайного вектора $\overline{X}$; $\widehat{\overline{M}}$ и $\widehat{\Sigma}$ – ОМП вектора математического ожидания и ковариационной матрицы, вычисляемые по негруппированным данным:

$$\widehat{\overline{M}} = \frac{1}{n}\sum_{i=1}^{n}\overline{X}_i, \quad \widehat{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\overline{X}_i - \widehat{\overline{M}}\right)\left(\overline{X}_i - \widehat{\overline{M}}\right)^{\top}.$$

Основное множество задач проверки статистических гипотез в классическом корреляционном анализе касается проверки гипотез о векторе математического ожидания, ковариационной матрице, парных, частных и множественных коэффициентах корреляции.

Проверка гипотез о равенстве математического ожидания некоторому известному вектору. Проверяемая гипотеза имеет вид H_0 : $\overline{M} = \overline{M}_0$, где $\overline{M}_0$ – номинальное значение вектора математических ожиданий. Ковариационная матрица может быть известной или неизвестной.

а) Ковариационная матрица Σ известна. В этом случае статистика [2, 3]

$$X_m^2 = n\left(\widehat{\overline{M}} - \overline{M}_0\right)^{\top}\Sigma^{-1}\left(\widehat{\overline{M}} - \overline{M}_0\right) \qquad (3)$$

при справедливой гипотезе H_0 в качестве предельного распределения $G(X_m^2|H_0)$ имеет χ_m^2-распределение с числом степеней свободы m.

б) Ковариационная матрица Σ неизвестна. В этом случае используется статистика [2, 3]

$$T^2 = \frac{n(n-m)}{m(n-1)}(\widehat{M} - \overline{M}_0)^\top \widehat{\Sigma}^{-1}(\widehat{M} - \overline{M}_0), \tag{4}$$

которая при справедливости гипотезы H_0 имеет в пределе распределение Фишера с параметрами m и $n-m$: $G(T^2|H_0) = F_{m,n-m}$.

Проверка гипотез о коэффициентах парной корреляции. Взаимозависимость двух случайных величин характеризуется парным коэффициентом корреляции $r_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}$. Относительно парного коэффициента корреляции могут проверяться два вида гипотез: о значимости корреляции ($H_0 : r_{ij} = 0$) и равенстве коэффициента корреляции номинальному значению ($H_0 : r_{ij} = r_0$).

В случае проверки гипотезы $H_0 : r_{ij} = 0$ статистика [2, 3]

$$t = \frac{\sqrt{n-2}|\hat{r}_{ij}|}{\sqrt{1-\hat{r}_{ij}^2}}, \tag{5}$$

где $\hat{r}_{ij} = \hat{\sigma}_{ij}/\sqrt{\hat{\sigma}_{ii}\hat{\sigma}_{jj}}$ – оценка парного коэффициента, имеет в качестве предельного распределения $G(t|H_0)$ t_{n-2} – распределение Стьюдента с числом степеней свободы $n-2$.

При проверке гипотезы $H_0 : r_{ij} = r_0$ статистика [2, 3]

$$z_0 = \sqrt{n-3}\left[\frac{1}{2}\ln\left(\frac{1+\hat{r}_{ij}}{1-\hat{r}_{ij}}\right) - \frac{1}{2}\ln\left(\frac{1+r_0}{1-r_0}\right) - \left(\frac{r_0}{2(n-1)}\right)\right] \tag{6}$$

имеет в качестве предельного закона $G(z_0|H_0)$ стандартное нормальное распределение $N(0,1)$.

Проверка гипотез о коэффициентах частной корреляции. В случае частных корреляций рассматриваются условные корреляции между двумя величинами при фиксированных значениях некоторых других величин.

Представим случайный вектор $\overline{X}$ в следующем виде:

$$\overline{X} = \begin{bmatrix} \overline{X}_1 \\ \overline{X}_2 \end{bmatrix},$$

где $\overline{X}_1 = (X_1, X_2, \ldots, X_l)^\top$, $\overline{X}_2 = (X_{l+1}, X_{l+2}, \ldots, X_m)^\top$, а вектор математических ожиданий и ковариационную матрицу

$$\overline{M} = \begin{bmatrix} \overline{M}_1 \\ \overline{M}_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

Если случайный вектор $\overline{X}$ подчиняется нормальному закону с вектором средних $\overline{M}$ и ковариационной матрицей Σ, то условное распределение подвектора $\overline{X}_1$ при известном $\overline{X}_2$ является нормальным с математическим ожиданием $\overline{M}_1 + B(\overline{X}_2 - \overline{M}_2)$ и ковариационной матрицей $\Sigma_{11\bullet 2}$, где $B = \Sigma_{12}\Sigma_{22}^{-1}$, $\Sigma_{11\bullet 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

ОМП для частного коэффициента корреляции определяется соотношением

$$\hat{r}_{ij\bullet l+1,\ldots,m} = \frac{\hat{\sigma}_{ij\bullet l+1,\ldots,m}}{\sqrt{\hat{\sigma}_{ii\bullet l+1,\ldots,m}\hat{\sigma}_{jj\bullet l+1,\ldots,m}}},$$

где $\hat{\sigma}_{ij\bullet l+1,\ldots,m}$ – элемент i-й строки и j-го столбца матрицы $\Sigma_{11\bullet 2}$, l – число компонент в условном распределении, $2 \le l \le m$. В данном случае при оценке взаимозависимости между компонентами X_i и X_j случайной величины $\overline{X}$ исключается влияние компонент $X_{l+1}, X_{l+2}, \ldots, X_m$.

При проверке гипотез вида

$$H_0 : r_{ij\bullet l+1,\ldots,m} = 0 \text{ и } H_0 : r_{ij\bullet l+1,\ldots,m} = r_0$$

используются те же самые статистики, что и для парного коэффициента корреляции. Но в данном случае в соответствующих соотношениях n заменяется на $n - m + l$.

Для проверки гипотезы $H_0 : r_{ij\bullet l+1,\ldots,m} = 0$ вычисляется статистика

$$t = \frac{\sqrt{n - m + l - 2}|\hat{r}_{ij\bullet l+1,\ldots,m}|}{\sqrt{1 - \hat{r}_{ij\bullet l+1,\ldots,m}}}. \qquad (7)$$

При этом предельным распределением статистики $G(t|H_0)$ является $t_{n-m+l-2}$-распределение Стьюдента с числом степеней свободы $n - m + l - 2$.

При проверке гипотезы $H_0 : r_{ij\bullet l+1,\ldots,m} = r_0$ используется статистика

$$z_0 = \sqrt{n-3}\left[\frac{1}{2}\ln\left(\frac{1 + \hat{r}_{ij\bullet l+1,\ldots,m}}{1 - \hat{r}_{ij\bullet l+1,\ldots,m}}\right) - \frac{1}{2}\ln\left(\frac{1 + r_0}{1 - r_0}\right) - \left(\frac{r_0}{2(n-1)}\right)\right], \qquad (8)$$

предельным распределением $G(z_0|H_0)$ которой является стандартное нормальное распределение $N(0,1)$.

Проверка гипотезы о коэффициенте множественной корреляции. Множественный коэффициент корреляции является мерой зависимости компоненты многомерной случайной величины от некоторого множества компонент. Можно рассматривать корреляцию между одной компонентой случайного вектора и множеством всех остальных или каким-то подмножеством.

Если представить случайный вектор $\overline{X}$ в том виде, как это было показано выше, то ОМП множественного коэффициента корреляции между $X_i, i \le l$ и множеством компонент $X_{l+1}, X_{l+2}, \ldots, X_m$ определяется соотношением

$$\hat{r}_{i\bullet l+1,\ldots,m} = \sqrt{\frac{\hat{\sigma}_{(i)}\Sigma_{22}^{-1}\hat{\sigma}_{(i)}^{\top}}{\hat{\sigma}_{ii}}},$$

где $\sigma_{(i)}$ – i-я строка матрицы Σ_{12}, σ_{ii} – элемент матрицы Σ_{11}.

Для проверки гипотезы $H_0 : r_{i\bullet l+1,\ldots,m} = 0$ вычисляется статистика [2, 3]

$$F = \frac{n-m+l-1}{m-l}\,\frac{\hat{r}^2_{i\bullet l+1,\ldots,m}}{1-\hat{r}^2_{i\bullet l+1,\ldots,m}}, \tag{9}$$

предельным распределением $G(F|H_0)$ которой является $F_{m-l,n-m+l-1}$-распределение Фишера с параметрами $m-l$ и $n-m+l-1$.

4. Моделирование и исследование распределений статистик критериев, используемых в корреляционном анализе

Подчеркнем, что указанные предельные распределения рассмотренных выше статистик имеют место при наблюдении многомерного нормального закона. Что произойдет с предельными распределениями статистик, насколько могут быть справедливы выводы, формулируемые на основании решения задач классического корреляционного анализа, если наблюдаемый многомерный закон отличается от нормального, заранее сказать нельзя.

Исследование распределений статистик при принадлежности наблюдений нормальному закону. На первом этапе методами статистического моделирования исследовались распределения статистик корреляционного анализа при условии, что наблюдения принадлежат многомерному нормальному закону. Близость получаемых эмпирических распределений статистик известным в данном случае предельным законам является доводом в пользу надежности методики при анализе достоверности результатов последующих исследований.

Моделирование и исследование эмпирических распределений статистик классического корреляционного анализа показали, что они очень хорошо согласуются с соответствующими теоретическими предельными распределениями. Например, на рис. 3 представлены эмпирическое распределение статистики X^2_m (3) и соответствующее предельное χ^2_m-распределение при $m = 2$ и объеме выборки $n = 45$. В ходе проведенных исследований объемы выборок значений статистик, формируемых

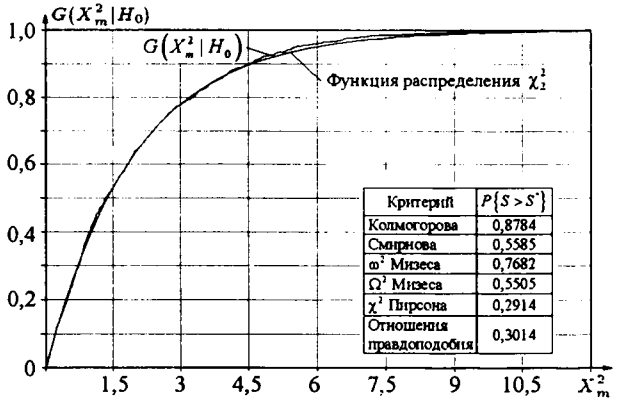


Рис. 3. Распределение статистики X_2^2 при нормальном законе ($m = 2$), $n = 45$

в результате моделирования, всегда задавались равными 1000. На рисунке отражены также результаты проверки согласия эмпирического распределения с теоретическим по критериям Колмогорова, Смирнова, ω^2 и Ω^2 Мизеса, χ^2 Пирсона и отношениям правдоподобия [4, 5]: по каждому из критериев приведен достигнутый уровень значимости $P\{S > S^*\} = 1 - G(S|H_0)$, где $G(S|H_0)$ – предельное распределений статистики S при справедливости проверяемой гипотезы H_0, S^* – значение статистики критерия согласия, вычисленное по анализируемой выборке. Исследование сходимости распределений статистик корреляционного анализа к предельным в зависимости от объема выборки n многомерного закона показало, что для тех статистик, параметры предельных распределений которых не зависят от объема выборки [статистики (3), (6) и (8)], эмпирические распределения статистик оказываются близки к предельным уже при выборках сравнительно небольшого объема. Так, у статистики X_m^2 высокий достигаемый уровень значимости по критериям согласия наблюдается с объемов выборки $n = 30 \div 45$, а для статистики z_0 (как для парного коэффициента корреляции, так и для частного) – с $n = 100 \div 150$. Распределения статистик T^2, t (для парного и частного коэффициента) и F, параметры предельных распределений которых зависят от объема выборки n, хорошо согласуются с предельными, начиная с объемов выборок $n = 15 \div 30$.

Существенного влияния размерности случайного вектора m на сходимость распределений соответствующих статистик к предельным при исследовании отмечено не было.

Исследование распределений статистик при многомерном законе, построенном по экспоненциальному семейству с параметром формы . В первую очередь отметим важную особенность, связанную с моделированием многомерного "ненормального" распределения, которую приходится учитывать при моделировании и исследовании рассматриваемых в данной работе статистик. Дело в том, что матрица Σ, задаваемая на этапе моделирования многомерного вектора, отличается от ковариационной матрицы получаемого псевдослучайного вектора. Это отличие является следствием того, что преобразуется случайный вектор, сформированный из компонент, распределенных по "ненормальному" закону (в нашем случае – по экспоненциальному семейству).

Но для исследования распределений статистик корреляционного анализа нам необходимы выборки псевдослучайных величин с известными (истинными) параметрами (математическим ожиданием и ковариационной матрицей), соответствующими проверяемой гипотезе. Для решения этой проблемы в качестве "истинной" ковариационной матрицы было предложено использовать среднее арифметическое ее оценок максимального правдоподобия, получаемых по ряду выборок большого размера при неизменном векторе математических ожиданий. Таким образом, мы (пока) не можем моделировать псевдослучайные векторы с "ненормальным" законом с заданной ковариационной матрицей, но можем моделировать с известной ковариационной матрицей. А этого достаточно для целей нашего исследования.

Эту особенность приходится учитывать только при исследовании статистики (3), при вычислении которой используется известная ковариационная матрица Σ. При вычислении остальных статистик, исследуемых в данной работе, используется оценка ковариационной матрицы.

Исследования распределений статистик проводились для многомерных законов, моделируемых на основании рассмотренной процедуры при параметрах $\lambda > 1$.

Это ограничение обусловлено тем, что предельным случаем экспоненциального семейства при $\lambda \to 0$ является распределение Коши, которое представляет собой пример "патологического" распределения: не существует математического ожидания и дисперсия расходится. Поэтому при параметре $\lambda < 1$ мы в результате моделирования получаем закон с вырожденной ковариационной матрицей.

Распределения статистик корреляционного анализа при многомерных законах, отличающихся от нормального и моделируемых в соответствии с описанной процедурой, базирующейся на экспоненциальном семействе с параметром формы λ, определяющим вид закона, иссле-

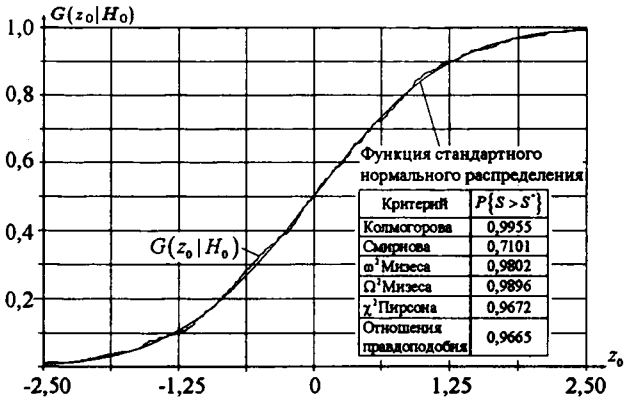
Рис. 4. Распределение статистики T^2 при многомерном законе ($m = 2$), $\lambda = 5$ и $n = 15$

довались при различных объемах выборок n и различной размерности случайных величин m. На рис. 4–8 приведены примеры эмпирических распределений исследуемых статистик с отражением соответствующих предельных распределений классических статистик. На рисунках представлены также значения достигнутых уровней значимости $P\{S > S^*\}$ по критериям Колмогорова, Смирнова, ω^2 и Ω^2 Мизеса, χ^2 Пирсона и отношения правдоподобия при проверке согласия, полученных в результате моделирования эмпирических распределений статистик предельным распределениям классических статистик.

На рис. 4 показан вид распределения статистики T^2 при законе, смоделированном по параметру $\lambda = 5$. Высокие достигнутые уровни значимости по всем критериям согласия и визуальная близость распределения статистики T^2 и предельного в случае многомерного нормального закона F-распределения Фишера позволяют утверждать, что вид предельного распределения статистики значимо не изменился.

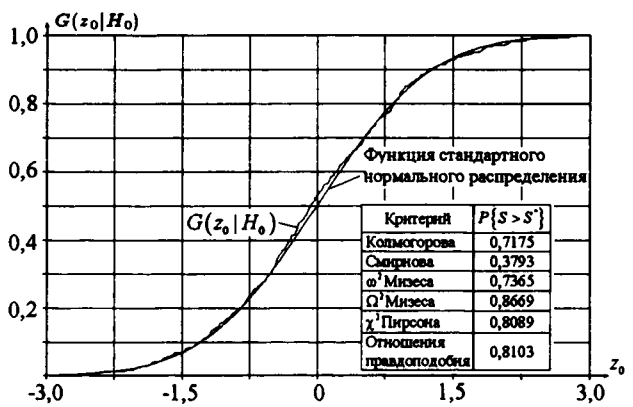
На рис. 5, 6 приведены результаты моделирования распределения статистики z_0, вычисляемой по формуле (6) при проверке гипотез о номинальном значении коэффициента парной корреляции, при многомерных законах, моделируемых соответственно при параметрах нашей процедуры $\lambda = 5$ и $\lambda = 10$. И в данных случаях очень высокие достигнутые уровни значимости по всем критериям согласия свидетельствуют в пользу того, что вид предельного распределения статистики тот же, что и в классическом случае.

Результаты моделирования распределения статистики z_0, вычисля-



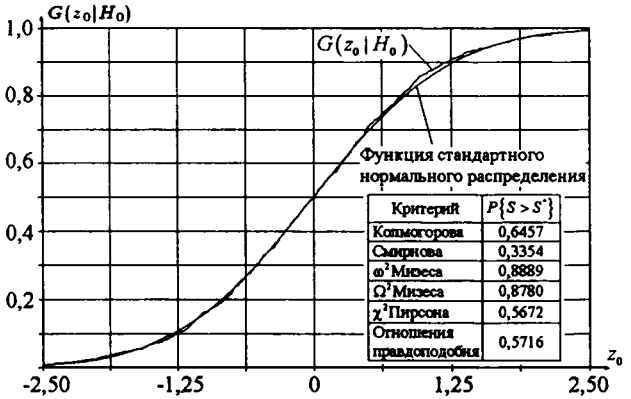
Критерий	$P\{S>S^*\}$
Колмогорова	0,9955
Смирнова	0,7101
ω^2 Мизеса	0,9802
Ω^2 Мизеса	0,9896
χ^2 Пирсона	0,9672
Отношения правдоподобия	0,9665

Рис. 5. Распределение статистики z_0 (для парного коэффициента корреляции) при многомерном законе ($m = 3$), $\lambda = 5$ и $n = 100$



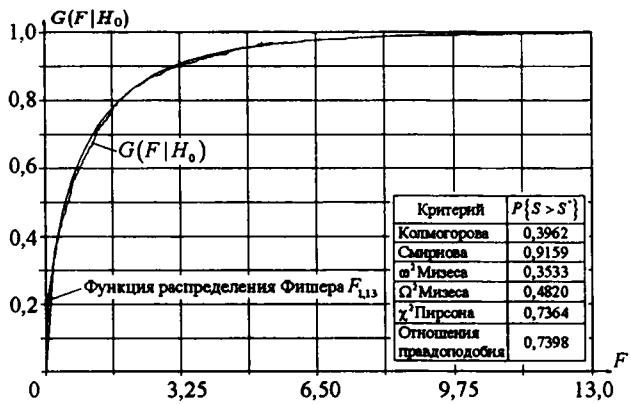
Критерий	$P\{S>S^*\}$
Колмогорова	0,7175
Смирнова	0,3793
ω^2 Мизеса	0,7365
Ω^2 Мизеса	0,8669
χ^2 Пирсона	0,8089
Отношения правдоподобия	0,8103

Рис. 6. Распределение статистики z_0 (для парного коэффициента корреляции) при многомерном законе ($m = 3$), $\lambda = 10$ и $n = 100$



Критерий	$P\{S > S^*\}$
Колмогорова	0,6457
Смирнова	0,3354
ω^2 Мизеса	0,8889
Ω^2 Мизеса	0,8780
χ^2 Пирсона	0,5672
Отношения правдоподобия	0,5716

Рис. 7. Распределение статистики z_0 (для частного коэффициента корреляции) при многомерном законе ($m = 3$), $\lambda = 5$ и $n = 100$



Критерий	$P\{S > S^*\}$
Колмогорова	0,3962
Смирнова	0,9159
ω^2 Мизеса	0,3533
Ω^2 Мизеса	0,4820
χ^2 Пирсона	0,7364
Отношения правдоподобия	0,7398

Рис. 8. Распределение статистики F при многомерном законе ($m = 3$), $\lambda = 10$ и $n = 15$

емой при проверке гипотез о частном коэффициенте корреляции, при законе, построенном по параметру $\lambda = 5$, отражены на рис. 7. На рис. 8 представлены результаты исследования распределения статистики F при многомерном законе, смоделированном по параметру $\lambda = 10$. И в этих случаях можно констатировать близость эмпирических распределений статистик и предельных законов, полученных в классическом корреляционном анализе.

Результаты проведенных исследований распределений рассмотренных статистик корреляционного анализа показывают, что в случае многомерных законов, достаточно существенно отличающихся от нормального (более островершинных или более плосковершинных, но симметричных), значимого изменения предельных распределений исследуемых статистик не происходит.

Это позволяет, по крайней мере, надеяться, что статистические выводы в задачах корреляционного анализа будут оставаться корректными и при нарушении предположений о нормальности наблюдаемого многомерного закона при условиях сохранения симметрии, существовании вектора математических ожиданий и невырожденности ковариационной матрицы.

5. Заключение

Основные результаты исследований могут быть сформулированы следующим образом.

Предложена универсальная процедура моделирования, позволяющая моделировать случайные векторы, подчиненные как многомерному нормальному (параметр формы $\lambda = 2$), так и отличному от него закону с "заданной" мерой отклонения.

Исследования эмпирических распределений статистик корреляционного анализа при наблюдении многомерного нормального закона показали, что они хорошо согласуются с теоретическими предельными распределениями, полученными в классическом корреляционном анализе, и подтвердили эффективность методики исследований.

Исследования распределений статистик корреляционного анализа в случае многомерных законов, отличающихся от нормального в достаточно широких пределах (смоделированных в соответствии с разработанной процедурой с параметром $\lambda \geq 1$), показали, что значимого изменения предельных распределений исследуемых статистик не происходит. Эмпирические распределения рассматриваемых статистик попрежнему хорошо описываются предельными распределениями, полученными в классическом корреляционном анализе в предположении о

нормальности наблюдаемого вектора. Это существенно расширяет сферу корректного применения методов классического корреляционного анализа в приложениях.

СПИСОК ЛИТЕРАТУРЫ

1. *Ермаков С. М., Михайлов Г. А.* Статистическое моделирование. Москва: Наука, 1982. 296 с.
2. *Андерсон Т.* Введение в многомерный статистический анализ. М.: Физматгиз, 1963. 500 с.
3. *Лемешко Б. Ю.* Корреляционный анализ многомерных наблюдений случайных величин: Программная система. Новосибирск: Изд-во НГТУ, 1995. 39 с.
4. *Денисов В. И., Лемешко Б. Ю., Постовалов С. Н.* Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Часть I. Критерии типа χ^2. Новосибирск: Изд-во НГТУ, 1998. 126 с.
5. *Лемешко Б. Ю., Постовалов С. Н.* Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Часть II. Непараметрические критерии. Новосибирск: Изд-во НГТУ, 1999. 85 с.

А.Г. Медведев

Робастная версия GMM-оценки

Рассматривается робастная версия оценок обобщенного метода моментов (GMM), удобная для оценивания параметров многофакторных моделей временной структуры процентных ставок. Приводятся условия состоятельности и асимптотической нормальности робастных оценок.

1. Введение

При принятии решений на финансовых рынках инвестор нуждается в информации, позволяющей ему представлять динамику рыночных показателей, важнейшим из которых является временная структура процентных ставок, показывающая доходность безрисковых инструментов в зависимости от срока их погашения. В последнее время в качестве моделей динамики процентных ставок чаще всего используются стохастические дифференциальные уравнения, а в многофакторных моделях

– их системы [1]. В простейших случаях эти системы являются линейными, что позволяет выписывать их решения в аналитическом виде, что может служить основой для построения разностных версий этих моделей [2]. Это в некотором смысле необходимо, с одной стороны, потому, что рыночные показатели котируются в дискретные моменты времени, а с другой стороны – для построения процедур оценивания неизвестных параметров модели, работающих с дискретно поступающей информацией. В последнее время популярным методом оценивания параметров стал так называемый обобщенный метод моментов [3]. Оценки по этому методу (в дальнейшем GMM-оценки) обладают многими полезными свойствами, однако характеризуются неустойчивостью по отношению к выбросам загрязняющих ошибок и аномальным наблюдениям. Здесь мы рассмотрим робастную версию GMM-оценок. Изложение в большей степени основывается на результатах работ [4–6]. После описания традиционного GMM вводятся модифицированные моментные условия со специальным образом определенной весовой матрицей, фильтрующей выбросы ошибок и аномальные наблюдения, формулируются условия регулярности, обеспечивающие GMM-оценкам состоятельность и асимптотическую нормальность. Особое внимание уделяется свойствам функции влияния для обеих версий GMM-оценок – обычной и робастной.

2. Обычный GMM-метод

Сначала рассмотрим оценку GMM, какой она обычно используется в литературе по обработке групповых данных. Не останавливаясь на финансовой интерпретации используемых переменных, рассмотрим линейную динамическую модель групповых данных

$$y_{it} = \alpha y_{i,t-1} + x'_{it}\beta + u_{it}, \tag{1}$$

для $i = \overline{1,N}$, $t = \overline{1,T}$, где N и T обозначают соответственно число инструментальных переменных и периодов времени. Мы предположим, что T/N мало, так что мы имеем прямоугольный массив с большим числом инструментальных переменных, которые наблюдаются в течение ограниченного числа периодов. (При исследовании состоятельности и асимптотической нормальности мы будем рассматривать асимптотический случай при $N \to \infty$ для фиксированного T.) K-вектор x_{it} содержит инструментальные переменные. Мы предполагаем, что возмущающее слагаемое u_{it} состоит из двух компонент: $u_{it} = \mu_i + \epsilon_{it}$, где μ_i и ϵ_{it} – случайные величины. Определим для удобства $z'_{it} = (y_{i,t-1}, x'_{it})$ и $\gamma' = (\alpha, \beta')$, тогда (1) можно записать в виде

$$y_{it} = z'_{it}\gamma + \mu_i + \epsilon_{it}. \tag{2}$$

Полезно ввести также следующие обозначения. Определим вектор $y_i' = (y_{i1}, \ldots, y_{iT})$ и матрицу $y' = (y_1, \ldots, y_N)$. Аналогично положим $Z_i' = (z_{i1}, \ldots, z_{iT})$ и $Z' = (Z_1, \ldots, Z_N)$. Вектор ϵ_i и матрица ϵ определяются аналогично y_i и y, а X_i и X – аналогично Z_i и Z. Кроме того, определим $(T-1) \times T$-матрицу D_T равенством

$$D_T = \begin{pmatrix} -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{pmatrix}.$$

Матрица D_T преобразовывает вектор y_i в вектор-столбец первых разностей $D_T y_i = (\Delta y_{i2}, \ldots, \Delta y_{iT})'$, где Δ является оператором первой разности $\Delta y_{it} = y_{it} - y_{i,t-1}$. Используя эти обозначения, можно переписать уравнение (2) в матричной форме:

$$D_T y_i = D_T Z_i \gamma + D_T \epsilon_i. \tag{3}$$

Вектор параметров γ можно состоятельно оценить при помощи GMM-метода [3]. Для того чтобы построить GMM-оценку, нужно знать корреляции между ϵ_{it}, μ_i и z_{it}. Поэтому здесь мы примем предположения о корреляционных свойствах случайных компонент для получения обычной GMM-оценки. В дальнейшем будем ссылаться на них как на предположения A1. В последующем мы заменим их модифицированными предположениями для того, чтобы определить робастную GMM-оценку.

*Предположения A*1: $\mathbf{E}(\epsilon_{it} y_{i0}) = 0$ для всех i и $t \geq 1$; $\mathbf{E}(\epsilon_{it} x_{js}) = 0$ для всех i, j, t и s; $\mathbf{E}(\epsilon_{it}\epsilon_{js}) = 0$, если $i \neq j$ или $t \neq s$; $\mathbf{E}(\epsilon_{it}\mu_j) = 0$ для всех i, j и t; $\mathbf{E}(\mu_i \mu_j) = 0$ для всех $i \neq j$.

Предположения A1 можно ослабить, модифицируя их определенным образом. Например, некоторые авторы ослабляют предположения о строгой экзогенности инструментальных переменных x_{it}, другие используют только подмножество условий, упомянутых в предположении A1. Однако такая модификация предположений не изменяет основной результат, касающийся чувствительности к выбросам GMM-оценок, получаемых на основе этих предположений. Такие оценки будем называть обычными GMM-оценками.

Используя предположения A1, легко получить равенство, которое принято называть моментным условием:

$$\mathbf{E}(W_i' D_T \epsilon_i) = 0, \tag{4}$$

где

$$W_i = \begin{pmatrix} y_i^0 & 0 & 0 & \dots & 0 & x_i^T & 0 & 0 & \dots & 0 \\ 0 & y_i^1 & 0 & \dots & 0 & 0 & x_i^T & 0 & \dots & 0 \\ 0 & 0 & y_i^2 & \dots & 0 & 0 & 0 & x_i^T & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & y_i^{T-2} & 0 & 0 & 0 & \dots & x_i^T \end{pmatrix},$$

$y_i^t = (y_{i0}, \dots, y_{it})$ и $x_i^T = (\mathrm{vec}(X_i'))'$. В условии (4) используется предположение о строгой экзогенности переменных x_{it}. Если такое предположение корректно, оно приводит к асимптотически более эффективной оценке. В этих условиях GMM-оценка определяется как

$$\hat{\gamma} = \arg\min_{\gamma} \mathbf{E}[W_i' D_T(y_i - Z_i\gamma)]' A_0 \mathbf{E}[W_i' D_T(y_i - Z_i\gamma)]. \qquad (5)$$

Матрица A_0 является положительно определенной весовой матрицей, которая определяется пользователем. В версии оценки с конечной выборкой оператор $\mathbf{E}$ обозначает оператор эмпирического усреднения, а матрица A_0 определяется как элемент сходящейся последовательности положительно определенных матриц $\{A_n, n \geq 1\}$ [3].

3. Функция влияния обычной GMM-оценки

Рассмотренная версия GMM-оценки часто используется в литературе по обработке данных, но оказывается чувствительной к выбросам данных. Как большие возмущения ϵ_{it}, так и аномальные значения инструментальных переменных W_i могут вызвать большие изменения реально получаемых оценок. Для того чтобы оценить чувствительность обычной GMM-оценки, используем понятие функции влияния (IF) [5]. Функция влияния оценки измеряет влияние малых возмущений в распределении данных на оценку. Для того чтобы формально ввести IF, мы сначала заметим, что $\hat{\gamma}$ зависит от совместного распределения y_i и Z_i. Обозначим это распределение символом F_0, так что $\hat{\gamma} = \hat{\gamma}(F_0)$. Теперь пусть G будет некоторой произвольной функцией распределения. Если функция распределения данных F_0 слабо изменяется, например на $F_\eta = (1-\eta)F_0 + \eta G$, где η является некоторым малым положительным числом, мы получим возмущенную оценку $\hat{\gamma} = \hat{\gamma}(F_\eta)$. Тогда IF определяется как

$$IF(G, \hat{\gamma}, F_0) = \lim_{\eta \downarrow 0} \frac{\hat{\gamma}(F_\eta) - \hat{\gamma}(F_0)}{\eta}, \qquad (6)$$

если этот предел существует. Желаемым свойством IF является ограниченность. Если IF является неограниченной, малые изменения в распределении данных могут привести к большим изменениям соответствующих оценок. Покажем, что IF обычной GMM-оценки является неограниченной, что показывает чувствительность оценки к выбросам данных. Позже мы получим робастную версию GMM-оценки, которая будет иметь ограниченную IF.

Предполагая, что перестановка операций интегрирования и дифференцирования допустима, из (5) можно получить условие первого порядка

$$\mathbf{E}[W_i' D_T Z_i]' A_0 \mathbf{E}[W_i' D_T (y_i - Z_i \hat{\gamma}(F_0))] = 0, \tag{7}$$

где математическое ожидание вычисляется по функции распределения F_0. Если F_0 заменить на F_η, равенство (7) можно продифференцировать по η и оценить полученную производную в точке $\eta = 0$. Заметим, что $d\hat{\gamma}(F_\eta)/d\eta$ равна функции влияния в точке $\eta = 0$. Мы получим

$$\mathbf{E}[W_i' D_T Z_i]' A_0 \mathbf{E}[\mathbf{E}_G(W_i' D_T \epsilon_i) - \mathbf{E}(W_i' D_T Z_i) IF(G, \hat{\gamma}, F_0)] = 0,$$

где $\mathbf{E}_G(.)$ обозначает математическое ожидание по функции распределения G. Из этого выражения IF легко определяется в виде

$$\begin{aligned} IF(G, \hat{\gamma}, F_0) = {} & [\mathbf{E}(W_i' D_T Z_i)' A_0 \mathbf{E}(W_i' D_T Z_i)]^{-1} \times \\ & \times \mathbf{E}(W_i' D_T Z_i)' A_0 \mathbf{E}_G(W_i' D_T \epsilon_i). \end{aligned} \tag{8}$$

Функция IF, определяемая формулой (8), является неограниченной, что легко выясняется, если G выбрать в виде распределения с точечной вероятностной массой.

4. Робастные GMM-оценки

Ключевым этапом для получения робастности к выбросам является замена моментных условий (4) на условие

$$\mathbf{E}(W_i' \Phi_i D_T \epsilon_i) = 0, \tag{9}$$

где Φ_i — диагональная весовая матрица, которая зависит от значений W_i и $D_T \epsilon_i$. Более детальное описание матрицы Φ_i приводится ниже. Здесь же упомянем, что влияние Φ_i состоит в назначении меньших весов наблюдениям с аномальными входными переменными. Моментное условие (9) отличается от обычного моментного условия (4). Они совпадают только в случае, когда $\Phi_i = I$. Вместе с тем в большинстве случаев условие (9) можно интерпретировать как несколько ослабленный

вариант (4), когда предполагается, что условию ортогональности (4) удовлетворяют не все наблюдения, а только большинство их. Введение Φ_i приводит к тому, что оценка автоматически реагирует на возможные выбросы. Основанная на моментном условии (9) версия выборочной GMM-оценки [6] определяется как

$$\hat{\gamma}_N = \arg\min_{\gamma} \left[\frac{1}{N}\sum_{i=1}^{N} W_i'\Phi_{iN}D_T(y_i - Z_i\gamma)\right]' \times \\ \times A_N \left[\frac{1}{N}\sum_{i=1}^{N} W_i'\Phi_{iN}D_T(y_i - Z_i\gamma)\right], \tag{10}$$

где $\{A_N, N \geq 1\}$ является последовательностью матриц, которая сходится почти наверняка к некоторой положительно определенной матрице A_0. Зависимость Φ_{iN} от объема выборки будет выяснена ниже. Заметим, что робастная GMM-оценка, определенная формулой (10), использует два типа весов. Во-первых, весовая матрица Φ_{iN} используется для ослабления влияния выбросов на значение оценки. Во-вторых, матрица A_N определяет веса, которые задаются модифицированными моментными условиями (9). Введем обозначения, необходимые для определения весовой матрицы Φ_{iN}. Пусть w_{it} обозначает t-й столбец матрицы W_i', φ_{itN} является t-м диагональным элементом матрицы Φ_{iN}, а $e_{it} = e_{it}(\gamma) = \Delta y_{i,t+1} - \Delta z_{i,t+1}'\gamma$. Структуру φ_{itN} определим следующим образом:

$$\varphi_{itN} = v_{tN}(w_{it})\hat{\sigma}_N\psi(e_{it}/\hat{\sigma}_N)/e_{it}, \tag{11}$$

если $e_{it} \neq 0$, и $\varphi_{itN} = v_{tN}(w_{it})$ в остальных случаях. Скалярная величина $\hat{\sigma}_N$ имеет смысл оценки масштабного коэффициента для $\{e_{it}\}$. Оценкой масштабного коэффициента $\hat{\sigma} = \hat{\sigma}(e_1, \ldots, e_N)$ выборки объема N одномерных наблюдений $e_1, \ldots, e_N$ является такая оценка, что выполняется равенство $\hat{\sigma}(ce_1, \ldots, ce_N) = |c|\hat{\sigma}(e_1, \ldots, e_N)$.

Оценка инвариантна к масштабу, если она обладает свойством $\hat{\sigma}(ce_1, \ldots, ce_N) = \hat{\sigma}(e_1, \ldots, e_N)$. Предположим, что эта оценка сходится почти наверное к положительной константе $\hat{\sigma}_0$. Наличие $\hat{\sigma}_N$ делает робастную GMM-оценку инвариантной к масштабу. Как можно видеть из (11), весовой множитель φ_{itN} расчленяется на два сомножителя. Первый зависит от w_{it} через функцию v_{tN}. Эта часть используется для ослабления влияния аномальных наблюдений в пространстве инструментальных переменных. Второй сомножитель зависит от e_{it} и служит для уменьшения веса наблюдений с большими невязками. Если функция ψ ограниченная, легко видеть, что наблюдения с большими

абсолютными значениями e_{it} получат малые веса. Такая декомпозиция φ_{itN} на две части, одна из которых зависит от w_{it}, а другая от e_{it}, следует из взвешивающей схемы для обычной регрессии [5]. Обычная GMM-оценка является частным случаем (10), когда принимается, что $\psi(e_{it}/\hat{\sigma}_N) = e_{it}/\hat{\sigma}_N$ и $v_{tN}(w_{it}) \equiv 1$.

Теперь обратимся к определению функционального вида весовой функции. Сначала сосредоточим внимание на выборе ψ, а затем определим v_{tN}. При выборе конкретного вида этих весовых функций, в первую очередь, предусмотрим защиту от выбросов, а затем обратим внимание на эффективность оценки. Наиболее важным условием для нечувствительности к выбросам является ограниченность функций $\psi(e_{it}/\hat{\sigma}_N)$ и $v_{tN}(w_{it})w_{it}$. Ряд спецификаций, которые удовлетворяют этому критерию можно найти в литературе [4]. Здесь мы используем функцию ψ, которую можно легко интерпретировать

$$\psi(e) = \begin{cases} e, & \text{если} \quad |e| \le c_1, \\ \tilde{\psi}(|e|)\mathrm{sgn}(e), & \text{если} \quad c_1 < |e| \le c_2, \\ 0, & \text{если} \quad |e| > c_2, \end{cases}$$

где $0 < c_1 < c_2$, $\mathrm{sgn}(.)$ – знаковая функция, а $\tilde{\psi}(|e|)$ является дважды непрерывно дифференцируемой функцией. В качестве этой функции можно взять полином с требуемыми свойствами. При использовании такой функции абсолютные значения нормированной невязки, меньшие c_1, получают вес, равный единице, в то время как наблюдения со значением нормированной невязки за пределами интервала $[-c_2, c_2]$ получают нулевой вес. Остальные наблюдения рассматриваются как промежуточные и учитываются лишь частично при определении оценки параметров.

Для конкретизации описанной выше функции ψ требуется задать две константы c_1 и c_2. Эти константы определяют как чувствительность, так и эффективность оценки. Между значениями этих констант имеется естественный компромисс. С одной стороны, мы хотим иметь достаточно эффективную оценку, если выборка не содержит выбросов. Для этого нужно установить значения $c_1 = c_2 = \infty$. С другой стороны, мы хотим, чтобы оценка была нечувствительной к выбросам. Это требует устанавливать константы c_1 и c_2 достаточно малыми положительными величинами. Мы установим их из следующих требований: $c_1^2 = \chi_1^{-2}(q_1)$ и $c_2^2 = \chi_1^{-2}(q_2)$, где χ_p^{-2} является обратной функцией χ^2-распределения с p степенями свободы. Величины q_1 и q_2, $0 < q_1 < q_2 < 1$, должны быть достаточно близкими к единице [6]. Такой выбор подразумевает, что получающаяся робастная GMM-оценка будет иметь вы-

сокую эффективность, если распределение ошибки близко к гауссовскому распределению. Кроме того, оценка будет робастной к выбросам, поскольку наблюдения с выбросами автоматически исключаются из выборки. Чтобы завершить описание робастной GMM-оценки, мы должны определить весовую функцию v_{tN}. Чтобы оценка была менее чувствительной к выбросам, $v_{tN}(w_{it})$ должна становиться малой, если w_{it} сильно отклоняется от кластера w_{it}-векторов. Для того чтобы формализовать это, используем расстояние Махалонобиса, которое определяется выражением

$$\delta(w_{it}, m_t, V_t) = \sqrt{(w_{it} - m_t)' V_t^{-1} (w_{it} - m_t)}, \tag{12}$$

где m_t и V_t представляют соответственно положение вектора и ковариационную матрицу w_{it}. Пусть для краткости $\delta_{it} = \delta(w_{it}, m_t, V_t)$ и $\hat{\delta}_{it} = \delta(w_{it}, \hat{m}_{tN}, \hat{V}_{tN})$, где $\hat{m}_{tN}$ и $\hat{V}_{tN}$ – оценки соответственно для m_t и V_t. Если V_t – вырожденная, в (12) следует использовать псевдоинверсию. С помощью расстояния Махалонобиса определим функцию $v_{tN}(w_{it}) = \psi_t(\hat{\delta}_{it}^2)/\hat{\delta}_{it}^2$, если $\hat{\delta}_{it} \neq 0$, и $v_{tN}(w_{it}) = 1$ – в противном случае. Здесь ψ_t является такой функцией, что аномальные векторы w_{it} получают меньший вес. Определим ψ_t так же, как $\psi(e)$, но с константами настройки c_{1t} и c_{2t}. Эти константы определим равенствами $c_{1t} = \chi_{pt}^{-2}(q_1)$ и $c_{2t} = \chi_{pt}^{-2}(q_2)$, где $\chi_{pt}^{-2}(q)$ является обратной функцией χ^2-распределения с pt степенями свободы и $pt = \mathrm{rank}(V_{tN})$. Эти значения констант настройки снова порождают веса, которые можно достаточно просто интерпретировать при практическом использовании.

В качестве оценок для m_t и V_t, конечно, можно принять соответственно выборочное среднее $\hat{m}_t$ и обычную выборочную матрицу ковариации $\hat{V}_{tN} = (N-1)^{-1} \sum_{i=1}^{N} (w_{it} - \hat{m}_{tN})(w_{it} - \hat{m}_{tN})'$. Однако при этом оказывается, что оценка будет чувствительной к появлению выбросов. Здесь лучше использовать аффинную эквивариантную S-оценку $(\hat{m}_{tN}, \hat{V}_{tN})$ [5]. Имеются литературные свидетельства, что S-оценка работает даже в том случае, когда среди данных имеется 50% выбросов. Это наиболее высокий процент загрязнения, которому может противостоять аффинная эквивариантная оценка. Кроме того, при некоторых слабых условиях регулярности S-оценка является состоятельной и асимптотически нормальной.

5. Состоятельность и асимптотическая нормальность робастных GMM-оценок

Как было упомянуто выше, мы рассматриваем только асимптотический случай $N \to \infty$ для фиксированного T. Для того чтобы доказать состоятельность и асимптотическую нормальность робастной GMM-оценк сделаем ряд предположений.

а) Модель (2) строго определена, т.е. имеются $(\gamma, \sigma_{\epsilon,0}, \sigma_{\mu,0})$ такие, что (2) порождает процесс данных с возмущениями ϵ_{it} и μ_{it}, имеющими соответственно дисперсии σ_ϵ^2 и σ_μ^2. Ошибки ϵ_{it} независимы и одинаково распределены с конечной дисперсией.

б) Векторы $\{p_i, i = \overline{1,N}\}$ независимы и одинаково распределены с плотностью вероятностей вида

$$|V|^{-1} f((p-m)'V^{-1}(p-m)),$$

где f – невозрастающая функция, определенная на интервале $[0,\infty)$ и V – положительно определенная матрица. Кроме того, функции f и ρ, которые используются при построении S-оценки [5], имеют общие точки убывания, т.е. существует такая константа c_1, что $f(c_0) > f(c_1) > f(c_2)$ и $\rho(c_0) > \rho(c_1) > \rho(c_2)$ для $c_0 > c_1 > c_2 > 0$.

в) Моментные условия (11) удовлетворяются. Кроме того,

$$\mathbf{E}(W_i \Phi_i D_T (y_i - Z_i \gamma)) \neq 0$$

для всех $(\gamma, \sigma) \neq (\gamma_0, \sigma_0)$.

г) Функция ψ_t непрерывно дифференцируема. Функция ψ дважды непрерывно дифференцируема.

д) Пусть $\gamma \in \Gamma, \sigma \in \Sigma, m \in \mathcal{M}$ и $v \in \mathcal{V}$, где $\Gamma \subset \mathbf{R}^{K+1}, \Sigma \subset (0,\infty)$ и $\mathcal{M} \subset \mathbf{R}^{T(K+1)+1}$, а $\mathcal{V}$ – подмножество пространства положительно определенных матриц. $\Gamma \times \Sigma \times \mathcal{M} \times \mathcal{V}$ является компактом. Последовательность $\{\hat{\sigma}_N\} \in \Sigma$ сходится почти наверное к σ_0, а последовательность весовых матриц для моментных условий (9) -- к положительно определенной матрице A_0.

е) Пусть $\theta = (\sigma, m', \mathrm{vec}(V)')'$. Тогда $\mathbf{E}[\partial(W_i \Phi_i D_T \epsilon_i)/\partial\gamma']$ имеет полный ранг, $\mathbf{E}[\partial(W_i \Phi_i D_T \epsilon_i)/\partial\theta'] = 0$ и

$$|\mathbf{E}[\partial(W_i \Phi_i D_T \epsilon_i)/\partial\gamma')(\partial(W_i \Phi_i D_T \epsilon_i)/\partial\gamma')']| \neq 0.$$

ж) Когда это необходимо, порядок интегрирования и дифференцирования можно изменять. Кроме того, все математические ожидания, которые используются в доказательствах, предполагаются конечными.

Теперь можно сформулировать следующую теорему.

Теорема. *Если предположения а) - ж) удовлетворяются, тогда робастная GMM-оценка, определяемая (10), является строго состоятельной. Кроме того, оценка $\hat{\gamma}$ асимптотически нормально распределена:*

$$N^{1/2}(\hat{\gamma}_N - \gamma_0) \xrightarrow{d} N(0, MM_2M'),$$

где

$$M = (M_1'A_0M_1)^{-1}M_1'A_0;$$

$$M_1 = \mathbf{E}\left(\sum_{t=1}^{T-1} w_{it}v_t(w_{it})\psi'(e_{it}/\sigma_0)\Delta z_{i,t+1}'\right);$$

$$M_2 = \mathbf{E}\left[\left(\sum_{t=1}^{T-1}\sigma_0 w_{it}v_t(w_{it})\psi(e_{it}/\sigma_0)\right)\left(\sum_{t=1}^{T-1}\sigma_0 w_{it}v_t(w_{it})\psi(e_{it}/\sigma_0)\right)'\right];$$

$$v_t(w_{it}) = \begin{cases} \psi_t(\delta_{it}^2)/\delta_{it}^2, & \text{если} \quad \delta_{it} \neq 0, \\ 1 & \text{в других случаях.} \end{cases}$$

Здесь δ_{it} определена формулой (12), а $\psi'(e_{it}) = d\psi(e_{it})/de_{it}$.

Доказательство теоремы состоит в том, чтобы показать, что при принятых условиях а) - ж) выполняются условия теоремы о строгой состоятельности и асимптотической нормальности [7, гл.3].

Асимптотическая матрица ковариации MM_2M' робастной GMM-оценки зависит от матрицы A_0. При выполнении моментного условия (9) матрицу A_0 можно выбрать так, чтобы минимизировать дисперсию робастной GMM-оценки. Следующее следствие представляет желаемый результат.

Следствие. *Пусть V_1 - асимптотическая матрица ковариации робастной GMM-оценки для $A_0 = M_2^{-1}$, и пусть V_2 - асимптотическая матрица ковариации робастной GMM-оценки для другого выбора A_0. Тогда $V_2 - V_1$ является неотрицательно определенной матрицей.*

Заметим, что оптимальность оценки понимается только в предельном смысле при существовании минимальной дисперсии, если выполняется набор моментных условий. Для того чтобы вычислить оптимальную робастную GMM-оценку, требуется знать матрицу M_2. Эту оценку можно получить путем вычисления робастной GMM-оценки, использующей неоптимальный выбор $A_N^{(0)}$. Невязки этой предварительной оценки затем можно применить для оценивания M_2, используя определение этой матрицы в условиях теоремы, заменив оператор математического

ожидания оператором эмпирического усреднения [3]. Оптимальная робастная GMM-оценка получается с помощью подстановки $A_N^{(1)} = M_2^{-1}$ и повторного расчета оценки.

6. Функция влияния робастной GMM-оценки

Для того чтобы выяснить свойства устойчивости робастной GMM-оценки снова обратимся к анализу функции влияния IF. В разделе 3 было показано, что обычная GMM-оценка имеет неограниченную функцию влияния. Здесь мы покажем, что робастная GMM-оценка обладает ограниченной IF. Поэтому можно ожидать, что эта оценка должна быть менее чувствительной к наблюдениям с выбросами. Так как IF определяется асимптотически, мы сначала дадим функциональное представление робастной GMM-оценки. Пусть функция распределения p_i имеет вид $F_\eta = (1-\eta)F_0 + \eta G$. Как и в разделе 3, функция F_0 является основным незагрязненным распределением, а G – распределение, порождающее выбросы. Используя моментные условия (9), робастную GMM-оценку определим как

$$\hat{\gamma}_N = \arg\min_{\gamma} \mathbf{E}(W_i'\Phi_i D_T(y_i - Z_i\gamma))' A_0 \mathbf{E}(W_i'\Phi_i D_T(y_i - Z_i\gamma)), \qquad (13)$$

где $\mathbf{E}$ обозначает оператор математического ожидания по распределению F_η. Дифференцируя правую часть (13) и используя условия теоремы, мы получим условия первого порядка:

$$\mathbf{E}\left[\sum_{t=1}^{T-1} w_{it} v_t(w_{it}) \psi'\left(\frac{e_{it}}{\sigma_0}\right) \Delta z_{i,t+1}'\right]' A_0 \mathbf{E}\left[\sum_{t=1}^{T-1} \sigma_0 w_{it} v_t(w_{it}) \psi\left(\frac{e_{it}}{\sigma_0}\right)\right] = 0, \qquad (14)$$

где $e_{it} = e_{it}(\hat{\gamma})$ и $\hat{\gamma} = \hat{\gamma}(F_\eta)$. Считается, что предположения а) – ж) удовлетворяются. Для того чтобы упростить обозначения, определим матрицу B_η и вектор C_η соответственно равенствами

$$B_\eta = \mathbf{E}_\eta\left(\sum_{t=1}^{T-1} w_{it} v_t(w_{it}) \psi'(e_{it}/\sigma_0) \Delta z_{i,t+1}'\right)$$

и

$$C_\eta = \mathbf{E}\left(\sum_{t=1}^{T-1} \sigma_0 w_{it} v_t(w_{it}) \psi(e_{it}/\sigma_0)\right).$$

Дифференцируя равенство (14) по η и вычисляя результат дифференцирования в точке $\eta = 0$, получим

$$\frac{dB'_\eta}{d\eta} A_0 C_0 - B'_0 A_0 B_0 IF(G, \hat{\gamma}, F_0) + \\ + B'_0 A_0 \mathbf{E}_G \left(\sum_{t=1}^{T-1} \sigma_0 w_{it} v_t(w_{it}) \psi(e_{it}/\sigma_0) \right) = 0, \qquad (15)$$

где $\mathbf{E}_G$ обозначает оператор математического ожидания по распределению G. Согласно условиям (9), C_0 равно нулю. Кроме того, мы предполагаем, что производная B_η по η конечна при $\eta = 0$. Это тривиально выполняется, если ψ и ψ_t задаются так, как в разделе 4. Тогда равенство (15) можно переписать в виде

$$IF(G, \hat{\gamma}, F_0) = (B'_0 A_0 B_0)^{-1} B'_0 A_0 \mathbf{E}_G \left(\sum_{t=1}^{T-1} \sigma_0 w_{it} v_t(w_{it}) \psi(e_{it}/\sigma_0) \right).$$

При заданных функциях ψ и v_t ограниченность функции влияния робастной GMM-оценки следует из того факта, что аномальные наблюдения получают меньшие веса. Большие невязки ограничиваются функцией ψ, а веса аномальных векторов инструментальных переменных w_{it} уменьшаются с помощью функции v_t.

Метод, используемый здесь для получения функции влияния GMM-оценки, взят из [5]. Он опирается на предположение о независимости и одинаковой распределенности векторов p_i. Тот факт, что модель (1) порождает стохастический процесс для каждого источника независимых переменных, подразумевает ограничения на форму загрязняющего распределения G, которое можно выбирать реалистически.

Итак, можно построить альтернативу обычной GMM-оценке, которая менее чувствительна к выбросам. Эта альтернативная робастная оценка использует весовые функции для наблюдений при моментных условиях. Эти весовые функции уменьшают влияние выбросов и аномальных наблюдений на оценку. Как традиционная GMM-оценка, так и ее робастная к выбросам версия состоятельны и асимптотически нормально распределены при определенных условиях регулярности. Состоятельность и асимптотическая нормальность этих оценок в общем случае имеют место при различных условиях. Условия могут совпасть в частных случаях, например, если ошибки нормально распределены и если вектора p_i имеют сферическое распределение с нулевым средним.

Наконец, робастность альтернативной GMM-оценки формально демонстрируется тем фактом, что она имеет ограниченную IF. В этом отношении оценка отличается от обычной GMM-оценки, имеющей неограниченную функцию влияния.

СПИСОК ЛИТЕРАТУРЫ

1. *Langetieg T. C.* A multivariate model of the term structure // J. Finance. 1980. Vol. 35. P. 71-97, 1980.
2. *Ilieva N. G.* On the some of yield interest rate models of the affine class term structures. EURO Working Group on Financial Modelling, 26th Meeting in Trondheim, 2000.
3. *Hansen L. P.* Large sample properties of generalised method of moments estimators // Econometrica. 1982. Vol. 50. P. 1029-1054.
4. *Хьюбер П.* Робастность в статистике. М.: Мир, 1984.
5. *Хампель Ф. и др.* Робастность в статистике. Подход на основе функций влияния. М.: Мир, 1989.
6. *Lucas A., van Dijk R., Kloek T.* Outlier robust GMM estimation of leverage determinants in linear dynamic panel data models. Working paper. Free University Amsterdam and Erasmus University Rotterdam, 1996.
7. *Gallant A. R.* Nonlinear statistical models. NY: Wiley, 1987.

В.И. Смагин, Е.В. Иванова

Синтез робастных регуляторов для систем со случайными скачкообразными параметрами

Рассматривается алгоритм синтеза робастных регуляторов по интегральному квадратичному критерию для непрерывных линейных систем со случайными скачкообразными параметрами, описываемыми цепью Маркова с конечным числом состояний. Управление формируется в виде регулятора с обратной связью по выходу.

1. Введение

Актуальной является задача синтеза управлений для класса систем со случайными скачкообразно изменяющимися параметрами, которые могут использоваться в качестве моделей реальных физических объектов. Такой класс систем имеет две составляющие состояния. Первая составляющая изменяется непрерывно, ее моделью является дифференциальное уравнение, вторая изменяется дискретно, и для ее описания

часто используется дискретная цепь Маркова с конечным числом состояний. К данному классу могут быть отнесены системы с возможными нарушениями (например вследствие внезапных отказов), многорежимные системы, в которых смена режима функционирования осуществляется скачкообразно и определяется в соответствии с измерением некоторого случайного процесса. Методы и алгоритмы синтеза систем управления объектами со случайными скачкообразно изменяющимися параметрами изучались в работах [1–8], в которых разработаны алгоритмы синтеза регуляторов, требующих измерения дискретной составляющей состояния или включения в контур управления блоков, вычисляющих оценки этих составляющих.

В настоящей работе разрабатывается метод синтеза робастных регуляторов по интегральным квадратичным критериям для непрерывных объектов со случайными скачкообразными параметрами. При этом система управления синтезируется непосредственно по вектору наблюдаемого выхода (в виде обратной связи по выходу), и не используется информация о дискретной составляющей состояния.

2. Робастный регулятор для нестационарной системы

Поведение непрерывного объекта со случайными скачкообразными параметрами описывается следующими уравнениями:

$$\dot{x}(t) = A(\gamma)x(t) + B(\gamma)u(t) + \sum_{s=1}^{m} A_s(\gamma)x(t)\theta_s(t) + q(\gamma,t), \qquad (1)$$

$$x(0) = x_0,$$

где $x(t) \in \mathbf{R}^n$ – вектор состояния; $u(t) \in \mathbf{R}^l$ – управление; $\gamma(t) \in \mathbf{R}$ – марковская цепь с дискретным множеством состояний $\gamma_1, \gamma_2, \ldots, \gamma_r$ (вероятность перехода из i-го состояния в j-е ($i \neq j$) за время Δt равна $\lambda_{ij}\Delta t + o(\Delta t)$); x_0 – начальные условия ($\mathbf{E}(x_0 x_0^{\mathrm{T}} | \gamma(0) = \gamma_0 = \gamma_i) = N_{0i}$, γ_0 – начальное состояние переменной γ); $q(\gamma,t) \in R^n, \theta(t) \in R^m$ – белые гауссовские шумы с характеристиками: $\mathbf{E}q(\gamma,t) = 0$, $\mathbf{E}(q(\gamma,t)q^{\mathrm{T}}(\gamma,\tau)) = Q(\gamma)\delta(t-\tau)$, $\mathbf{E}\theta(t) = 0$, $\mathbf{E}(\theta(t)\theta^{\mathrm{T}}(\tau)) = I\delta(t-\tau)$, $\mathbf{E}(q(\gamma,t)\theta^{\mathrm{T}}(t)) = 0$; $A(\gamma)$, $B(\gamma)$ и $A_s(\gamma)$ ($s = \overline{1,m}$) – матрицы порядка $n \times n$, $n \times l$, $n \times n$ соответственно; матрицы $A_s(\gamma)$ определяют интенсивность мультипликативной составляющей возмущений.

Здесь приняты следующие обозначения: $^{<\mathrm{T}>}$ – операция транспонирования, $\delta(\cdot)$ – дельта-функция Дирака, $Q(\gamma) = Q^{\mathrm{T}}(\gamma) \geq 0$ – неотрицательно определенная матрица, I – единичная матрица.

Процесс $\gamma(t)$ задается уравнением

$$d\gamma(t) = \int v\Omega(dt, dv); \quad \gamma(0) = \gamma_0, \tag{2}$$

где пуассоновская случайная мера $\Omega(dt, dv)$ характеризуется функцией

$$\Pi_{t,\gamma_i} = \sum_{\substack{j=1 \\ j \neq i}}^{r} \lambda_{ij} \delta(v + \gamma_i - \gamma_j) dv. \tag{3}$$

Вектор вероятностей $p(t) = (p_1(t), p_2(t), \cdots, p_r(t))^\top$, где $p_i(t) = P\{\gamma(t) = \gamma_i\}$ $(i = \overline{1,r})$, удовлетворяет прямому уравнению Колмогорова

$$\frac{dp(t)}{dt} = \Lambda^\top p(t), \quad p(0) = \overline{p}_0, \quad t \in [0, T], \tag{4}$$

где матрица интенсивностей переходов Λ с элементами λ_{ij}, $\overline{p}_0$ – вектор начальных вероятностей. Предполагается, что наблюдению доступен вектор $y(t) = S(\gamma)x(t)$, где $y(t) \in \mathbf{R}^{n_1}$, $(n_1 \leq n)$, $S(\gamma)$ – матрица полного ранга. Зададим структуру робастного регулятора в виде

$$u(t) = K(t)y(t), \tag{5}$$

где $K(t)$ – подлежащая определению матрица коэффициентов передачи. Отметим, что структура закона управления вида (5) не требует измерения скачкообразного процесса $\gamma(t)$.

Для объекта (1) необходимо найти управление $u(t)$, $t \in [0, T]$, вида (5), при котором был бы минимальным следующий квадратичный интегральный критерий:

$$J[0,T] = \mathbf{E}\left(\int_0^T \sum_{i=1}^{r} p_i(\tau)(x^\top(\tau) C_i x(\tau) + u^\top(\tau) D_i u(\tau)) d\tau + \right.$$
$$\left. + \sum_{i=1}^{r} p_i(T) x^\top(T) E_i x(T) \,\middle|\, x(0) = x_0, \gamma(0) = \gamma_0 \right), \tag{6}$$

где $C_i = C_i^\top \geq 0$, $D_i = D_i^\top > 0$, $E_i = E_i^\top \geq 0$ – весовые матрицы.

Решение задачи синтеза робастного регулятора для нестационарной системы дает следующая теорема.

Теорема 1. *Оптимальные коэффициенты передачи робастного регулятора* (5) *определяются по формуле*

$$\mathrm{ct}(K(t)) = -\left[\sum_{i=1}^{r} p_i (D_i \otimes S_i N_i S_i^{\top})\right]^{-1} \mathrm{ct}\left(\sum_{i=1}^{r} p_i B_i^{\top} L_i N_i S_i^{\top}\right), \quad (7)$$

если существуют матрицы $N_i(t) > 0$, $L_i(t) > 0$, *удовлетворяющие следующей двухточечной краевой задаче:*

$$\dot{N}_i = \widetilde{A}_i N_i + N_i \widetilde{A}_i^{\top} + \sum_{\substack{j=1 \\ j \neq i}}^{r} \lambda_{ij}(N_j - N_i) + Q_i + \\ + \sum_{s=1}^{m} A_s^{(i)} N_i A_s^{(i)\top}, \quad N_i(0) = N_{0i}, \quad (8)$$

$$-\dot{L}_i = L_i \widetilde{A}_i + \widetilde{A}_i^{\top} L_i + S_i^{\top} K^{\top} D_i K S_i + C_i + \\ + \sum_{s=1}^{m} A_s^{(i)} L_i A_s^{(i)\top} + \sum_{\substack{j=1 \\ j \neq i}}^{r} \lambda_{ij}(L_j - L_i), \; L_i(T) = E_i, \; i = \overline{1,r}, \quad (9)$$

где $\widetilde{A}_i = A_i + B_i K S_i$. *При этом объект* (1) *с регулятором по выходу* (5) *устойчив в среднеквадратическом. (В* (7)–(9) $A_i = A(\gamma_i)$, $B_i = B(\gamma_i)$, $Q_i = Q(\gamma_i)$, $S_i = S(\gamma_i)$, *символ* $\otimes$ *обозначает кронекеровское произведение,* $\mathrm{ct}(\cdot)$ – *вектор-столбец, составленный из столбцов матрицы.)*

Теорема доказывается на основе объединения принципа минимума и метода функций Ляпунова. Метод доказательства близок к методу доказательства теоремы, приведенной в работе [8, *с.* 66].

3. Робастный регулятор для стационарной системы

В этом случае вместо критерия (7) необходимо минимизировать критерий

$$\limsup_{T \to \infty} \frac{1}{T} J[0, T]. \quad (10)$$

Предполагается, что пары матриц A_i, B_i стабилизируемы и в критерии (6) $E_i = 0$ $(i = \overline{1,r})$. Задача синтеза робастного регулятора при этом упрощается, так как уравнения (8)–(9) становятся алгебраическими:

$$\widetilde{A}_i N_i + N_i \widetilde{A}_i^{\top} + Q_i + \sum_{s=1}^{m} A_s^{(i)} N_i A_s^{(i)\top} + \sum_{\substack{j=1 \\ j \neq i}}^{r} \lambda_{ij}(N_j - N_i) = 0, \quad (11)$$

$$L_i\widetilde{A}_i + \widetilde{A}_i^{\top}L_i + S_i^{\top}K^{\top}D_iKS_i + C_i + \sum_{s=1}^{m} A_s^{(i)}L_iA_s^{(i)\top} +$$

$$+\sum_{\substack{j=1\\ j\neq i}}^{r}\lambda_{ij}(L_j - L_i) = 0,\ i = \overline{1,r}. \tag{12}$$

Постоянные оптимальные коэффициенты передачи определятся по формуле

$$\mathrm{ct}(K) = -\left[\sum_{i=1}^{r}\overline{p}_i(D_i\otimes S_iN_iS_i^{\top})\right]^{-1}\mathrm{ct}\left(\sum_{i=1}^{r}\overline{p}_iB_i^{\top}L_iN_iS_i^{\top}\right), \tag{13}$$

где $\overline{p}_i$ установившиеся значения вероятностей состояний скачкообразного процеса $\gamma(t)$.

Теорема 2. *Если существует решение уравнений* (11), (12) *и существуют числа* β_i $(0 < \beta_i < 1)$, *такие, что*

$$L_i > 0,\quad M_i = \beta_iC_i + \sum_{\substack{j=1\\ j\neq i}}^{r}\lambda_{ij}(L_j - L_i) \geq 0, \tag{14}$$

пара матриц

$$\sqrt{C_i}, A_i \tag{15}$$

детектируема, то матрица $A_i + B_iKS_i$ *- асимптотически устойчива* $(i = \overline{1,r})$.

Доказательство. Из условия детектируемости пары матриц (15) следует детектируемость пары матриц $\sqrt{(1-\beta_i)C_i}, A_i$, а из условия (14) и в силу теоремы 3.6 работы [9] следует детектируемость пары матриц

$$\sqrt{(1-\beta)C_i + M_i + S_i^{\top}K^{\top}D_iKS_i}, A_i + B_iKS_i. \tag{16}$$

Уравнение (12) эквивалентно следующему уравнению Ляпунова:

$$L_i(A_i + B_iKS_i) + (A_i + B_iKS_i)^{\top} +$$
$$+S_i^{\top}K^{\top}D_iKS_i + (1-\beta_i)C_i + M_i = 0. \tag{17}$$

Тогда, учитывая (17) и лемму 12.2 [9], при $L_i > 0$ и условии детектируемости пары матриц (16) получим, что матрица $A_i + B_iKS_i$ асимптотически устойчива.

Отметим, что при выполнении условий теоремы 2 замкнутая робастная система управления с коэффициентами передачи (13) будет устойчивой в среднеквадратическом.

СПИСОК ЛИТЕРАТУРЫ

1. *Красовский Н.Н., Лидский Э.А.* Аналитическое конструирование регуляторов в системах со случайными свойствами. I // Автоматика и телемеханика. 1961. № 9. С. 732-745; II // Там же. 1961. № 11. С. 1273-1278.
2. *Wonham W.M.* Random differential equation in control theory // Probabilistic methods in applied mathematics / Ed. A.T. Bharucha-Reid. NY: Academ. Press, 1971. P. 131-213.
3. *Параев Ю.И.* Введение в статистическую динамику процессов управления и фильтрации. М.: Сов. радио, 1976.
4. *Mariton M.* Jump linear system in automatic control. NY: Marcel Dekker Inc., 1990.
5. *Boukas E.K.* Control of systems with controlled jump Markov disturbances // Control Theory and Advanced Technology. 1993. Vol. 9. No. 2. P. 577-595.
6. *Смагин В.И., Параев Ю.И.* Синтез следящих систем управления по квадратичным критериям. Томск: Изд-во Том. ун-та, 1996.
7. *Do Val G.B., Costa O.L.* Output feedback control of Markov jump linear systems in continuos time // IEEE Trans. Automatic Control. 2000. Vol. AC-45. No. 5. P. 658-663.
8. *Смагин В.И., Поползухина Е.В.* Синтез следящих регуляторов с обратной связью по выходу для систем со случайными скачкообразными параметрами // Автоматика и вычислительная техника. 2001. № 6. С. 62-72.
9. *Уонем М.* Линейные многомерные системы управления. М.: Наука, 1980.

Ю.С. Харин, А.В. Акинфин

Робастность в планировании регрессионных экспериментов при наличии функциональных искажений

1. Введение

Непараметрическая и робастная статистика — одно из наиболее интенсивно разрабатываемых направлений математической и прикладной статистики в последние три десятилетия [1].

Для статистического оценивания параметров исследуемых стохастических систем, описываемых гипотетической моделью множественной линейной регрессии, широко используются оптимальные планы экспериментов, в частности D- и A-оптимальные планы [2, 3]. Однако на практике часто оказывается, что гипотетическая регрессионная модель

экспериментальных данных допускает некоторые искажения: непараметрические искажения функции регрессии, наличие "выбросов", неконтролируемая гетероскедастичность, коррелированность ошибок наблюдений. Поэтому всё бóльшее внимание уделяется исследованию устойчивости (робастности) классических (оптимальных в некотором смысле) планов экспериментов при отклонениях от гипотетических модельных предположений, а также поиску новых устойчивых (робастных) к заданным искажениям планов [4–8].

В данной статье представлены результаты исследования устойчивости (робастности) классических D- и A-оптимальных планов для множественной линейной регрессии, подверженной функциональным искажениям.

2. Математическая модель и постановки задач

Рассмотрим множественную линейную регрессию при наличии функциональных искажений:

$$y_i = \theta^{0T} x_i + \varepsilon\lambda(x_i) + \xi_i, \quad i = \overline{1,n}. \tag{1}$$

Здесь $\theta^0 = (\theta^0_1, \ldots, \theta^0_m)^T \in \mathbf{R}^m$ – подлежащий оцениванию m-вектор (столбец) параметров, $x_i = (x_{i1}, \ldots, x_{im})^T \in \mathbf{U}^m \subset \mathbf{R}^m$ – m-вектор значений контролируемых переменных (регрессоров) в i-м эксперименте, каждая компонента которого изменяется в $\mathbf{U} = [-1, +1]$, $y_i \in \mathbf{R}$ – значение наблюдаемой переменной в i-м эксперименте, $0 \le \varepsilon \le \varepsilon_+$ – уровень функциональных искажений, ограниченный сверху заданным допустимым уровнем искажений $\varepsilon_+ \ge 0$, $\xi = (\xi_1, \ldots, \xi_n)^T$ – n-вектор некоррелированных случайных ошибок наблюдений, таких что $\mathbf{E}\{\xi_i\} = 0$, $i = \overline{1,n}$, $E\xi\xi^T = \sigma^2 \mathbf{I}_n$, где $\mathbf{E}\{\cdot\}$ – символ математического ожидания, $\mathbf{I}_n$ – единичная $(n \times n)$-матрица; $\lambda(x) : \mathbf{U}^m \to \mathbf{R}$ – неизвестная детерминированная ограниченная функция, описывающая функциональные искажения. Заметим, что без потери общности область планирования экспериментов принято задавать в виде гиперкуба $\mathbf{U}^m$ [2,3].

Будем рассматривать два типа T_1, T_2 функциональных искажений, наиболее часто встречающихся в приложениях; T_1 определяет искажения в l_2-метрике: $\sum\limits_{i=1}^{n} \lambda^2(x_i) \le 1$; T_2 — это искажения в c-метрике: $0 \le |\lambda(x_i)| \le 1$, $i = \overline{1,n}$.

Представим модель (1) в матричном виде

$$Y = X\theta^0 + \varepsilon\Lambda(X) + \xi, \tag{2}$$

где $\Lambda(X) = (\lambda(x_1), \ldots, \lambda(x_n))^T$ – $(n \times 1)$-вектор, $X = (x_{ij}) = (x_1, \ldots, x_n)^T$ – $(n \times m)$-матрица плана экспериментов, $Y = (y_1, \ldots, y_n)^T$ – $(n \times 1)$-вектор наблюдаемых переменных. Предполагается, что план экспериментов — невырожденный: $|X^T X| \neq 0$.

Классическая МНК-оценка $\hat{\theta} = (\hat{\theta}_i) \in \mathbf{R}^m$ вектора параметров θ^0 для модели (1) при отсутствии искажений ($\varepsilon_+ = 0$) является несмещенной и имеет вид [2]

$$\hat{\theta} = (X^T X)^{-1} X^T Y; \tag{3}$$

её матрица вариаций равна

$$V_0(X) = \mathbf{E}\left\{(\hat{\theta} - \theta^0)(\hat{\theta} - \theta^0)^T\right\} = \sigma^2 (X^T X)^{-1};$$

определитель этой матрицы, который согласно [9] условимся называть *обобщенной дисперсией* оценки $\hat{\theta}$, равен

$$D_0(X) = |V_0(X)| = |\sigma^2 (X^T X)^{-1}| > 0,$$

а след этой матрицы, который условимся называть *суммарной дисперсией* оценки $\hat{\theta}$, равен

$$A_0(X) = tr\ V_0(X) = \sigma^2 tr\ \left((X^T X)^{-1}\right) > 0.$$

D-оптимальным точным планом X_0 называют план [2, 3], минимизирующий обобщенную дисперсию $D_0(X)$:

$$X_0 = \arg \min_{x_1, \ldots, x_n \in \mathbf{U}^m} D_0(X).$$

A-оптимальным точным планом X_0 называют план [2, 3], минимизирующий суммарную дисперсию $A_0(X)$:

$$X_0 = \arg \min_{x_1, \ldots, x_n \in \mathbf{U}^m} A_0(X).$$

Из [2] известно, что для гипотетической модели (1) при отсутствии искажений ($\varepsilon_+ = 0$) D- и A-оптимальным планом является план X_0, столбцы которого взаимноортогональны и компоненты $x_{ij}^0 = \pm 1$, $i = \overline{1, n}$, $j = \overline{1, m}$.

Актуальны следующие задачи, которым и посвящена данная статья: 1) анализ робастности классических D- и A-оптимальных планов экспериментов при наличии функциональных искажений; 2) построение робастных планов; 3) сравнительный анализ классических и робастных планов.

3. Анализ робастности D-оптимальных планов

Прежде всего исследуем статистические свойства МНК-оценки (3), когда наблюдения Y подвержены функциональным искажениям (1), (2) типов T_1, T_2.

Теорема 1. *Если множественная линейная регрессия подвержена функциональным искажениям (1), (2), то смещение, матрица вариаций, обобщенная и суммарная дисперсии МНК-оценки (3) равны соответственно*

$$b\{\hat{\theta}\} = \varepsilon(X^TX)^{-1}X^T\Lambda(X); \tag{4}$$

$$\mathbf{V}_\varepsilon\{\hat{\theta}\} = (X^TX)^{-1}(\varepsilon^2X^T\Lambda(X)\Lambda^T(X)X + \sigma^2X^TX)(X^TX)^{-1};$$

$$D_\varepsilon(X) = |\mathbf{V}_\varepsilon\{\hat{\theta}\}| = D_0(X)(1 + \frac{\varepsilon^2}{\sigma^2}\Lambda^T(X)X(X^TX)^{-1}X^T\Lambda(X)),$$

$$A_\varepsilon(X) = tr\ \mathbf{V}_\varepsilon\{\hat{\theta}\} = A_0(X) + \varepsilon^2\Lambda^T(X)X\left((X^TX)^{-1}\right)^2X^T\Lambda(X)).$$

Доказательство. Смещение оценки (3) для модели (1), (2) есть m-вектор

$$b\{\hat{\theta}\} = \mathbf{E}\left\{\hat{\theta} - \theta^0\right\} =$$
$$= \mathbf{E}\left\{(X^TX)^{-1}X^T(X\theta^0 + \varepsilon\Lambda(X) + \xi)\right\} - \theta^0 = \varepsilon(X^TX)^{-1}X^T\Lambda(X).$$

Таким образом, при появлении искажений оценка (3) становится смещенной оценкой.

Матрица вариаций оценки $\hat{\theta}$ имеет вид

$$\mathbf{V}_\varepsilon\{\hat{\theta}\} = \mathbf{E}\left\{(\hat{\theta} - \theta^0)(\hat{\theta} - \theta^0)^T\right\} =$$
$$= (X^TX)^{-1}(\varepsilon^2X^T\Lambda(X)\Lambda^T(X)X + \sigma^2X^TX)(X^TX)^{-1},$$

что совпадает с (4). Для вычисления обобщенной дисперсии $\left|\mathbf{V}_\varepsilon\{\hat{\theta}\}\right|$ воспользуемся свойством, следующим из известной формулы блочного определителя [10]: $|A + BB^T| = |A||I_k + B^TA^{-1}B|$, где A — невырожденная $(m \times m)$-матрица, B — $(m \times k)$-матрица. Полагая $k = 1$, $A = \sigma^2X^TX$, $B = \varepsilon X^T\Lambda(X)$, получаем

$$\left|\mathbf{V}_\varepsilon\{\hat{\theta}\}\right| = |(X^TX)^{-1}|^2|\sigma^2X^TX| \times$$
$$\times\left(1 + \frac{\varepsilon^2}{\sigma^2}\Lambda^T(X)X(X^TX)^{-1}X^T\Lambda(X)\right). \tag{5}$$

С учетом введенных обозначений приходим к выражению для обобщенной дисперсии в (4).

Простым вычислением несложно показать, что верна формула (4) для суммарной дисперсии. Теорема доказана.

Используя [6], введем следующие количественные характеристики робастности D-оптимальных планов:

- *коэффициент неустойчивости плана* X для искажений (1), (2)

$$æ_\varepsilon(X) = (D_\varepsilon(X) - D_0(X))/D_0(X); \tag{6}$$

- *верхняя граница обобщенной дисперсии* $D_+(X)$, для которой

$$\forall \Lambda(X),\ 0 \le \varepsilon \le \varepsilon_+, \qquad D_\varepsilon(X) \le D_+(X); \tag{7}$$

- *верхняя граница коэффициента неустойчивости плана* X

$$æ_+(X) = (D_+(X) - D_0(X))/D_0(X) \ge 0; \tag{8}$$

- *точные верхние границы обобщенной дисперсии и коэффициента неустойчивости плана* X

$$D_+^*(X) = \sup_{\substack{\Lambda(X),\\ 0\le\varepsilon\le\varepsilon_+}} D_\varepsilon(X),\ D_0(X) \le D_+^*(X) \le D_+(X), \tag{9}$$

$$æ_+^*(X) = (D_+^*(X) - D_0(X))/D_0(X) \ge 0,\ æ_+^*(X) \le æ_+(X);$$

- *δ-допустимый уровень искажений* ($\delta > 0$)

$$\varepsilon_+^* = \varepsilon_+^*(\delta, X) = \max\{\varepsilon_+ \ge 0 : æ_+^*(X) \le \delta\}. \tag{10}$$

Как видно из (6)–(9), величины D_+^* (или D_+), $æ_+^*$ (или $æ_+$) характеризуют гарантированное значение обобщенной дисперсии оценки (3) соответственно в смысле абсолютных и относительных значений. Согласно (10), величина ε_+^* характеризует максимально допустимый уровень функциональных искажений в (1), (2), при котором гарантируется заданный уровень δ коэффициента неустойчивости $æ_+^*$ регрессионного плана X. Чем меньше значения характеристик $D_+^*(X)$, $D_+(X)$, $æ_+^*(X)$, $æ_+(X)$ и чем больше $\varepsilon_+^*(\delta, X)$, тем устойчивее план X к функциональным искажениям в (1), (2).

Для модели (1) точный план X_* будем называть D_+^*-*робастным планом*, если

$$D_+^*(X_*) = \min_{x_1,\ldots,x_n \in \mathbf{U}^m} D_+^*(X). \tag{11}$$

Если не удается найти точную верхнюю границу, в (9) вместо $D_+^*(\cdot)$ будем использовать некоторую найденную верхнюю границу $D_+(\cdot)$ и аналогично (11) определим D_+-*робастный план* X_*; аналогично в (10) вместо $æ_+^*$ может использоваться $æ_+$ для нахождения $\varepsilon_+(\delta, X)$.

Для вычисления $D_\varepsilon(X)$, $æ_\varepsilon(X)$, $D_+^*(X)$, $æ_+^*(X)$, $\varepsilon_+^*(X)$ и построения X_* сформулируем вспомогательный результат.

Лемма 1 [10]. *Если C — идемпотентная матрица, то все ее характеристические числа равны* 1 *или* 0.

Теорема 2. *Если множественная линейная регрессия подвержена функциональным искажениям* (1), (2) *типа* T_1, *то точная верхняя граница обобщенной дисперсии*

$$D_+^*(X) = D_0(X)(1 + \varepsilon_+^2/\sigma^2), \tag{12}$$

а точная верхняя граница коэффициента неустойчивости плана X

$$æ_+^*(X) = \varepsilon_+^2/\sigma^2. \tag{13}$$

Доказательство. Из теоремы 1, (9) и описания искажений T_1 следует, что

$$D_+^*(X) = \max_{\substack{0\le|\Lambda(X)|\le 1,\\ 0\le\varepsilon\le\varepsilon_+}} \left(D_0(X)(1 + \frac{\varepsilon^2}{\sigma^2}\Lambda^T(X)X(X^TX)^{-1}X^T\Lambda(X))\right) =$$

$$= D_0(X)\left(1 + \max_{0\le\varepsilon\le\varepsilon_+}\left(\frac{\varepsilon^2}{\sigma^2}\max_{0\le|\Lambda(X)|\le 1}\Lambda^T(X)C\Lambda(X)\right)\right), \tag{14}$$

где $C = X(X^TX)^{-1}X^T$ — неотрицательно определенная симметрическая идемпотентная $(n\times n)$-матрица. Известно [10], что максимум квадратичной формы в единичном шаре равен

$$\max_{0\le|\Lambda|\le 1}\Lambda^TC\Lambda = \nu_{\max}(C),$$

где $\nu_{\max}(C)$ – максимальное собственное значение матрицы C. Так как $|X^TX| > 0$, то $\operatorname{rank}(C) = m \le n$ и с учетом леммы $\nu_{\max}(C) \equiv 1$. Теперь, максимизируя в (14) по $\varepsilon \in [0, \varepsilon_+]$, получаем (12). Соотношение (13) следует из (12) с учетом (9). Теорема доказана.

Следствие 1. *Для любого $\delta > 0$ δ-допустимый уровень функциональных искажений типа T_1 регрессионной модели* (1), (2) *не зависит от плана X и равен $\varepsilon_+^*(\delta, X) = \sigma\sqrt{\delta}$.*

Доказательство. Искомое выражение следует из (10), (13). Следствие доказано.

Следствие 2. *Если множественная линейная регрессионная модель* (1), (2) *подвержена функциональным искажениям типа T_1, то D_+^*-робастный план экспериментов совпадает с точным D-оптимальным планом: $X_* = X_0$.*

Доказательство. Из теоремы 2 следует, что величина $D_+^*(X)$ пропорциональна $D_0(X)$, минимум которой достигается на плане X_0, следовательно $X_* = X_0$. Следствие доказано.

Теорема 3. *Если множественная линейная регрессия подвержена функциональным искажениям* (1), (2) *типа T_2, то для верхних границ обобщенной дисперсии и коэффициента неустойчивости регрессионного плана X справедливы выражения*

$$D_+(X) = D_0(X)(1 + \varepsilon^2 n/\sigma^2), \quad æ_+(X) = \varepsilon^2 n/\sigma^2. \tag{15}$$

Доказательство. Обозначим $\Lambda = (\lambda_1, \ldots, \lambda_n)^T$ — n-вектор, $Q(\Lambda) = \Lambda^T C \Lambda$ — квадратичная форма для неотрицательно определенной симметрической идемпотентной $(n \times n)$-матрицы $C = X(X^TX)^{-1}X^T$. Тогда из (4), (9) и описания искажений T_2 имеем аналогично (14)

$$D_+^*(X) = D_0(X)(1 + \max_{0 \le \varepsilon \le \varepsilon_+} (\varepsilon^2/\sigma^2) \max_{\Lambda \in \mathbf{U}^n} Q(\Lambda)). \tag{16}$$

Впишем гиперкуб $\mathbf{U}^n$ в n-мерную гиперсферу $S_n(\sqrt{n})$ радиуса $\sqrt{n}$ с центром в точке $z = \mathbf{0}$: $S_n(\sqrt{n}) = \{z = (z_i) \in \mathbf{R}^n : |z| = \sqrt{n}\}$. Так как $\mathbf{U}^n$ содержится в n-мерном гипершаре, ограниченном гиперсферой $S_n(\sqrt{n})$, то для искомого максимума в (16) справедлива оценка

$$\max_{\Lambda \in \mathbf{U}^n} Q(\Lambda) \le \max_{0 \le |\Lambda| \le \sqrt{n}} Q(\Lambda). \tag{17}$$

Переходя к новым переменным

$$\mu = \Lambda/\sqrt{n} \in \mathbf{R}^n, \; Q(\Lambda) = Q(\sqrt{n}\mu) = nQ(\mu),$$

получаем

$$\max_{0 \le |\Lambda| \le \sqrt{n}} Q(\Lambda) = n \cdot \max_{0 \le |\mu| \le 1} Q(\mu). \tag{18}$$

Как уже отмечалось при доказательстве теоремы 2, $\max\limits_{0\le|\mu|\le 1} Q(\mu) = \nu_{\max}(C) = 1$, поэтому из (17), (16) получаем первое выражение в (15): $D_+^*(X) \le D_+(X)$. Тогда второе выражение в (15) следует из (8). Теорема доказана.

Следствие 3. *Для любого $\delta > 0$ δ-допустимый (в смысле критерия $æ_+$) уровень функциональных искажений типа T_2 регрессионной модели* (1), (2) *не зависит от плана X и равен*

$$\varepsilon_+(\delta, X) = \sigma\sqrt{\delta/n}.$$

Следствие 4. *Если множественная линейная регрессионная модель* (1), (2) *подвержена функциональным искажениям типа T_2, то D_+-робастный план экспериментов совпадает с точным D-оптимальным планом: $X_* = X_0$.*

Следствия 3, 4 доказываются аналогично следствиям теоремы 3 с использованием верхних границ (15).

Представляет интерес выяснить условия, при которых верхние границы в (15) совпадают с точными верхними границами

$$D_+^*(X) = D_+(X), \quad æ_+^*(X) = æ_+(X). \tag{19}$$

Следствие 5. *Если план экспериментов $X = (X_1 \vdots \ldots \vdots X_m)$ является ортогональным планом и для некоторого $k \in \{1, \ldots, m\}$ k-й столбец X_k является биполярным, то есть $x_{kj} = \pm a_k$ ($j = \overline{1,m}$) для некоторого $a_k > 0$, то достигаются равенства* (19).

Доказательство. В силу условия ортогональности $X_i^T X_{i'} = 0$ $(i \ne i')$, идемпотентная матрица $C = X(X^TX)^{-1}X^T$ имеет вид $C = \sum\limits_{i=1}^{m} \tilde{X}_i \tilde{X}_i^T$, где $\{\tilde{X}_i = (\tilde{x}_{ij}) = |X_i|^{-1}X_i : i = \overline{1,m}\}$ — ортонормированная подсистема m векторов. Согласно [10] отсюда следует, что $\{\tilde{X}_i : i = \overline{1,m}\}$ — подсистема m главных собственных векторов, соответствующих максимальному собственному значению $\nu_{\max} = 1$, на которых достигается максимум квадратичной формы $Q(\mu)$ в (18): $\max\limits_{0\le|\mu|\le 1} Q(\mu) = Q(\tilde{X}_1) = \cdots = Q(\tilde{X}_m) = 1$. Тогда неравенство в (17) обратится в равенство, если хотя бы один из собственных векторов системы $\{\sqrt{n}\cdot\tilde{X}_i : i = \overline{1,m}\}$ совпадет с одной из вершин гиперкуба $\mathbf{U}^n$. Для $i = k$ имеем в силу условия биполярности: $|X_k| = \sqrt{n}a_k$, $\tilde{x}_{kj} = \pm 1/\sqrt{n}$, $j = \overline{1,m}$. Следовательно, вектор $\sqrt{n}\tilde{X}_k$ имеет координаты ± 1 и является одной из вершин гиперкуба $\mathbf{U}^n$. При этом (17), (19) обращаются в равенства. Следствие доказано.

Отметим, что условия ортогональности плана X и биполярности его столбцов являются часто используемыми условиями в теории оптимального планирования экспериментов, в частности D-оптимальный план X_0 удовлетворяет этим свойствам [2, 3].

Следствие 6. *В случае искажений T_2 для точного D-оптимального плана X_0*

$$D_+^*(X_0) = D_0(X)(1 + \varepsilon^2 n/\sigma^2), \quad æ_+^*(X_0) = \varepsilon^2 n/\sigma^2.$$

Исследуем асимптотическое поведение коэффициента неустойчивости $æ_+^*(X)$ при увеличении числа экспериментов n.

Следствие 7. *В условиях следствия 5 при $n \to \infty$ в зависимости от асимптотического поведения уровня искажений $\varepsilon_+ = \varepsilon_+(n)$ справедливы следующие утверждения:*

1) *если $\varepsilon_+(n) = o(1/\sqrt{n})$, то план X — асимптотически робастный:*

$$æ_+^*(X) \to 0;$$

2) *если $\varepsilon_+(n) = c/\sqrt{n}$ для некоторого $0 < c < +\infty$, то*

$$æ_+^*(X) \to c^2/\sigma^2 > 0;$$

3) *если $\sqrt{n}\varepsilon_+(n) \to +\infty$, то план X — неустойчивый:*

$$æ_+^*(X) \to +\infty.$$

Доказательство. Достаточно провести анализ трех случаев асимптотики с учетом (15), (19). Следствие доказано.

4. Анализ робастности A-оптимальных планов

Аналогично предыдущему разделу, введем следующие количественные характеристики робастности A-оптимальных планов:

- *коэффициент неустойчивости регрессионного плана* X для искажений (1), (2)

$$æ_\varepsilon(X) = (A_\varepsilon(X) - A_0(X))/A_0(X); \tag{20}$$

- *верхняя граница суммарной дисперсии* $A_+(X)$, для которой

$$\forall \Lambda(X),\ 0 \le \varepsilon \le \varepsilon_+, \qquad A_\varepsilon(X) \le A_+(X); \tag{21}$$

- *верхняя граница коэффициента неустойчивости плана* X

$$æ_+(X) = (A_+(X) - A_0(X))/A_0(X) \geq 0; \tag{22}$$

- *точные верхние границы суммарной дисперсии и коэффициента неустойчивости плана* X

$$A_+^*(X) = \sup_{\substack{\Lambda(X),\\ 0\leq\varepsilon\leq\varepsilon_+}} A_\varepsilon(X), \quad A_0(X) \leq A_+^*(X) \leq A_+(X); \tag{23}$$

$$æ_+^*(X) = (A_+^*(X) - A_0(X))/A_0(X) \geq 0, \quad æ_+^*(X) \leq æ_+(X);$$

- *δ-допустимый уровень искажений* $(\delta > 0)$

$$\varepsilon_+^* = \varepsilon_+^*(\delta, X) = \max\{\varepsilon_+ \geq 0 : æ_+^*(X) \leq \delta\}. \tag{24}$$

Как видно из (20)–(23), величины A_+^* (или A_+), $æ_+^*$ (или $æ_+$) характеризуют гарантированное значение суммарной дисперсии оценки (3) соответственно в смысле абсолютных и относительных значений. Согласно (24), величина ε_+^* характеризует максимально допустимый уровень функциональных искажений в (1), (2), при котором гарантируется заданный уровень δ коэффициента неустойчивости $æ_+^*$ регрессионного плана X. Чем меньше значения характеристик $A_+^*(X)$, $A_+(X)$, $æ_+^*(X)$, $æ_+(X)$ и чем больше $\varepsilon_+^*(\delta, X)$, тем устойчивее план X к функциональным искажениям в (1), (2).

Для модели (1) точный план X_* будем называть *A_+^*-робастным планом,* если

$$A_+^*(X_*) = \min_{x_1,\ldots,x_n\in\mathbf{U}^m} A_+^*(X). \tag{25}$$

Если не удается найти точную верхнюю границу, в (23) вместо $A_+^*(\cdot)$ будем использовать некоторую найденную верхнюю границу $A_+(\cdot)$ и аналогично (25) определим *A_+-робастный план X_*;* аналогично в (24) вместо $æ_+^*$ может использоваться $æ_+$ для нахождения $\varepsilon_+(\delta, X)$.

Теорема 4. *Если множественная линейная регрессия подвержена функциональным искажениям* (1), (2) *типа* T_1, *то верхняя граница суммарной дисперсии*

$$A_+(X) = A_0(X)(1 + \varepsilon_+^2/\sigma^2), \tag{26}$$

а верхняя граница коэффициента неустойчивости плана X

$$æ_+(X) = \varepsilon_+^2/\sigma^2. \tag{27}$$

Доказательство. Из теоремы 1, (23) и описания искажений T_1 следует, что

$$A_+^*(X) = \max_{\substack{\Lambda(X)\in\Lambda_1,\\ 0\le\varepsilon\le\varepsilon_+}} \left(A_0(X)\varepsilon^2\Lambda^T(X)X((X^TX)^{-1})^2X^T\Lambda(X))\right) =$$

$$= A_0(X) + \max_{0\le\varepsilon\le\varepsilon_+}\left(\varepsilon^2 \max_{\Lambda(X)\in\Lambda_1}\Lambda^T(X)C\Lambda(X)\right), \qquad (28)$$

где $C = X\left((X^TX)^{-1}\right)^2X^T$ — неотрицательно определенная симметрическая $(n\times n)$-матрица. Известно [10], что максимум квадратичной формы в единичном шаре равен

$$\max_{0\le|\Lambda|\le1}\Lambda^TC\Lambda = \nu_{\max}(C),$$

где $\nu_{\max}(C)$ – максимальное собственное значение матрицы C.

Максимизируя в (28) по $\varepsilon\in[0,\varepsilon_+]$, получаем

$$A_+^*(X) = A_0(X) + \varepsilon_+^2\nu_{\max}(C).$$

Заменим точную верхнюю границу $A_+^*(X)$ на верхнюю границу $A_+(X)$ $(A_+^*(X)\le A_+(X))$:

$$A_+(X) = A_0(X) + \varepsilon_+^2\operatorname{tr}C.$$

Отметим, что

$$\operatorname{tr}C = \operatorname{tr}aX\left((X^TX)^{-1}\right)^2X^T = \operatorname{tr}a(X^TX)^{-1} = \frac{1}{\sigma^2}A_0(X).$$

Таким образом, получаем

$$A_+(X) = A_0(X)(1+\varepsilon_+/\sigma^2),$$

т. е. соотношение (26). Соотношение (27) следует из (26) с учетом (22). Теорема доказана.

Следствие 8. *Для любого $\delta>0$ δ-допустимый (в смысле критерия $æ_+$) уровень функциональных искажений типа T_1 регрессионной модели* (1), (2) *не зависит от плана X и равен $\varepsilon_+(\delta,X) = \sigma\sqrt{\delta}$.*

Доказательство. Искомое выражение следует из (24), (27). Следствие доказано.

Следствие 9. *Если множественная линейная регрессионная модель* (1), (2) *подвержена функциональным искажениям типа T_1, то*

A_+-робастный план экспериментов совпадает с точным A-оптимальным планом: $X_* = X_0$.

Доказательство. Из теоремы 4 следует, что величина $A_+(X)$ пропорциональна $A_0(X)$, минимум которой достигается на плане X_0, следовательно, $X_* = X_0$. Следствие доказано.

Как видно их доказательства теоремы 4, точная верхняя граница $A_+^*(X)$ имеет вид

$$A_+^*(X) = \sigma^2 \left(\mathbf{tr}(X^T X)^{-1} + \frac{\varepsilon_+^2}{\sigma^2} \nu_{\max}(C) \right). \tag{29}$$

Лемма 2. *Для любой вещественной $(n \times m)$-матрицы X $(n \geq m)$ ранга m верно, что все ненулевые собственные значения матрицы $C = X\left(X(X^TX)^{-1}X^T\right)$ являются собственными значениями матрицы $(X^TX)^{-1}$.*

Доказательство. Из [10] известно, что

$$\mathrm{rank}(X^TX) = \mathrm{rank}X,$$

следовательно, матрица X^TX — невырожденная.

В случае, когда $n = m$, результат леммы очевиден, так как $C = (X^TX)^{-1}$.

Когда $n > m$, несложно показать, что у матрицы C ровно m ненулевых собственных значений.

Обозначим $\{\nu_1, \ldots, \nu_n\}$ — собственные значения матрицы C, причем $\nu_1 \geq \nu_2 \geq \ldots \nu_m$, $\nu_{m+1} = \cdots = \nu_n = 0$; $\{y_1, \ldots, y_n\}$ — система линейно независимых собственных векторов матрицы C, причем вектор y_i соответствует собственному значению ν_i.

Тогда $\forall i$, $i = \overline{1,m}$, $\exists y_i$ — собственный вектор C, т.е.

$$X\left((X^TX)^{-1}\right)X^Ty_i = \nu_i y_i.$$

Домножим слева данное соотношение на X^T и получим, что вектор $y_i' = X^Ty_i$ является собственным вектором матрицы $(X^TX)^{-1}$, соответствующим собственному значению ν_i:

$$(X^TX)^{-1}X^Ty_i = \nu_i X^Ty_i, \quad \text{для всех} \quad i = \overline{1,m}.$$

Отметим, что ни один из собственных векторов y_i, $i = \overline{1,m}$, матрицы C не лежит в ядре матрицы X^T и никакая линейная комбинация из этих векторов, за исключением линейной комбинации с нулевыми коэффициентами, не лежит в ядре матрицы X^T. Следовательно, система векторов $y_1', \ldots, y_m'$ — линейно независимая.

Таким образом, на основании собственных векторов и значений матрицы C можно указать все собственные значения и собственные векторы матрицы $(X^TX)^{-1}$. Лемма доказана.

Следствие 10. *Точная верхняя граница (29) допускает представление*

$$A_+^*(X) = \sigma^2 \sum_{i=1}^{m} w_i \nu_i \left((X^TX)^{-1}\right), \tag{30}$$

где $\nu_i\left((X^TX)^{-1}\right)$ *— i-е собственное значение матрицы* $(X^TX)^{-1}$, $\nu_1\left((X^TX)^{-1}\right) \geq \cdots \geq \nu_m\left((X^TX)^{-1}\right)$, w_i *— вес i-го собственного значения,*

$$w_1 = 1 + \varepsilon_+^2/\sigma^2,\ w_2 = \cdots = w_m = 1. \tag{31}$$

Доказательство. Согласно лемме 2,

$$\nu_{\max}(C) = \nu_1\left((X^TX)^{-1}\right).$$

Подставляя $\nu_{\max}(C)$ в (29) и учитывая, что

$$\operatorname{tr}\left((X^TX)^{-1}\right) = \sum_{i=1}^{m} \nu_i\left((X^TX)^{-1}\right),$$

приходим к (30), (31). Следствие доказано.

Интересно отметить, что функционал (30) приводит к новому семейству критериев оптимальности регрессионных планов экспериментов.

Определение. wA-оптимальным планом будем называть план, минимизирующий следующий функционал:

$$F(X) = \sum_{i=1}^{m} w_i \nu_i\left((X^TX)^{-1}\right),$$

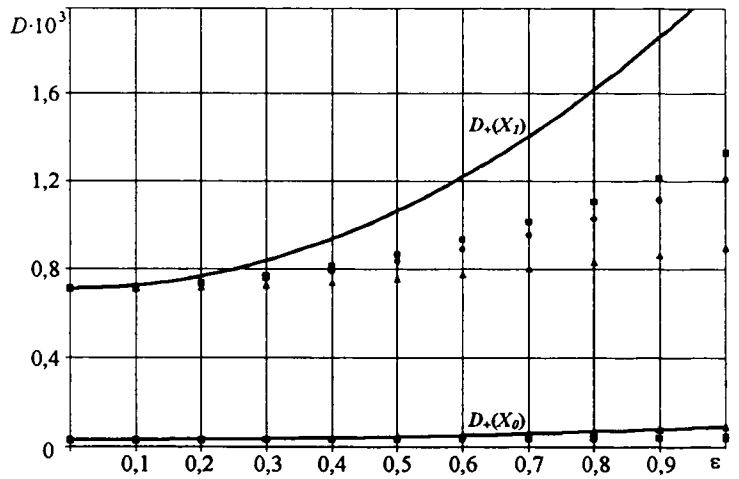
где $\{w_i\}$ — некоторые неотрицательные веса.

Частными случаями wA-оптимального плана, очевидно, являются:

- A-оптимальный план [2,3] при $w_i = 1$, $i = \overline{1,m}$, для случая $\varepsilon_+ = 0$;
- E-оптимальный план [2,3] при $w_1 = 1$, $w_2 = \cdots = w_m = 0$ для случая $\varepsilon_+ = 0$;
- A_+^*-робастный план при выборе коэффициентов в виде (31) для случая $\varepsilon_+ > 0$.

5. Численные результаты

Было проведено две серии численных экспериментов. В первой серии при $n = 16$, $m = 3$, $\sigma^2 = 0{,}5$ исследовалась устойчивость двух планов экспериментов (D-оптимального плана X_0 с компонентами $x_{ij} = \pm 1$,



$i = \overline{1,n}$, $j = \overline{1,m}$, построенного с помощью матрицы Адамара [2], и случайного равномерного плана X_1, компоненты которого $\{x_{ij}\}$ — независимые равномерно распределенные на $[-1,1]$ случайные величины) при гармонических искажениях типа T_1:

$$u(x_i) = \cos(\omega x_{i1} + \pi/3)\cos(\omega x_{i2} + \pi/3)\cos(\omega x_{i3} + \pi/3),$$

$$\lambda(x_i) = u(x_i)/\sqrt{u^2(x_1) + \cdots + u^2(x_n)},\ \omega = 2\pi f,\ f \in \{1/5, 1/3, 1\};$$

уровень искажений $\varepsilon = \varepsilon_+ \in \{0; 0.1; 0.2; \ldots; 1.0\}$. Значения обобщенных дисперсий $D_\varepsilon(X_0)$, $D_\varepsilon(X_1)$ вычислялись согласно (4), а их верхние границы — согласно (12).

На рисунке изображены графики зависимостей $D_+(X_0)$, $D_+(X_1)$. Здесь же незакрашенными кружками ($f = 1/2$), квадратиками ($f = 1/3$) и треугольниками ($f = 1$) изображены значения $D_\varepsilon(X_0)$, а закрашенными фигурами — $D_\varepsilon(X_1)$. Из рисунка виден значительный выигрыш в точности оценивания (по величине обобщенной дисперсии) при использовании D_+^*-робастного плана X_0 по сравнению с планом X_1.

Во второй серии экспериментов рассматривались функциональные искажения типа T_2 и для различных сочетаний

$$(n, m) \in \{(3,2), (4,3), (5,3)\}$$

минимаксная задача (11), (9): $\max\limits_{\Lambda \in \mathbf{U}^n} D_{\varepsilon_+}(X) \to \min\limits_{x_1, \ldots, x_n \in \mathbf{U}^m}$ решалась численным перебором всевозможных допустимых значений λ, $x_1, \ldots, x_n$ с

шагом Δ (вычислительная сложность имеет порядок $(2/\Delta)^{n(m+1)}$) для $\Delta = 0{,}1; 0{,}2; 0{,}25; 0{,}5$. Во всех этих экспериментах найденная численная оценка D_+^*-робастного плана X_* совпала с X_0, что согласуется с теоремой 3 и ее следствиями.

СПИСОК ЛИТЕРАТУРЫ

1. *Тарасенко Ф.П.* Непараметрическая статистика. Томск: Изд-во Том. ун-та, 1976.
2. *Федоров В.В.* Теория оптимального планирования. М.: Мир, 1968.
3. *Ермаков С.М., Жиглявский А.А.* Математическая теория оптимального эксперимента. М.: Наука, 1987.
4. *Muller Ch.H.* Robust planning and analysis of experiments // Lecture Notes in Statistics. NY: Springer, 1997. Vol. 124.
5. *Chang Y-J., Notz W.I.* Model robust desings // Handbook of Statistics. NY: Wiley, 1996. Vol. 13.
6. *Kharin Yu.* Robustness in statistical pattern recognition. Dordrecht/ Boston / London: Kluwer Academic Publishers, 1996.
7. *Wiens D.P.* Minimum variance unbiased weights and designs for biased regression model // 3rd St. Petersburg Workshop on Simulation. St.Petersburg: SPSU, 1998.
8. *Харин Ю.С., Акинфин А.В.* Об устойчивости D-оптимальных планов регрессионных экспериментов при наличии функциональных искажений // Вести НАН Беларуси. Серия физ.-мат. наук. 2000. № 2. С. 46-51.
9. *Андерсон Т.* Введение в многомерный статистический анализ. М.: Наука, 1963.
10. *Хорн Н., Джонсон Ч.* Матричный анализ. М.: Мир, 1989.

В.П. Шуленин

Робастные статистические процедуры

Развитие научных исследований в настоящее время характеризуется все более возрастающей ролью статистических процедур и методов, их широким применением в различных областях практики и требованиями качественной обработки экспериментальных данных. Однако их успешное применение сдерживается тем фактом, что многие стандартные (классические) процедуры, разработанные для конкретных типов статистических моделей, обеспечивают заданное качество решений лишь при справедливости всех предположений модели (т.е. при адекватности принятой статистической модели). На практике часто и неизбежно возникают ситуации, в которых реальные характеристики

данных отличаются от предполагаемых в модели (т.е. принятая модель не является адекватной). Применение стандартных процедур в таких ситуациях в лучших случаях может оказаться малоэффективным и даже может приводить к грубым ошибкам в статистических выводах. В связи с этим в математической статистике в последнее время интенсивно развивается новое направление, названное робастной статистикой. Это направление появилось как средство преодоления отрицательных последствий возможной неадекватности статистических моделей. Общая характеристика этого раздела статистики, основные понятия, задачи и подходы к их решению обсуждаются в [1].

В данной работе приводятся уже устоявшиеся результаты по основным типам робастных процедур на примере базовой задачи оценивания параметра положения случайной величины. Результаты, полученные при решении этой задачи, являются основой для обобщений и построения робастных процедур проверки статистических гипотез и оценивания параметров линейных моделей (эти обобщения могут быть найдены в научной литературе, см., например, [2] и [3]). Приведенные ниже результаты по исследованию свойств основных классов робастных процедур являются асимптотическими (т.е. полученными в предположении, что объем выборки n (n – число экспериментов) достаточно велик). Однако, как показывают многочисленные исследования, часто асимптотические результаты могут быть приемлемы для практических целей аппроксимацией уже при $n \geq 20$.

Основные классы робастных оценок параметров

Проиллюстрируем применение обсужденных в работе понятий робастности на задаче оценивания параметра положения случайной величины. Для наглядности рассмотрим случай, когда функция распределения (ф.р.) F принадлежит классу симметричных распределений вероятностей, т.е.

$$F \in \Im_S, \text{ где } \Im_S = \{F : F(x) = 1 - F(-x), \forall x \in R\}.$$

Итак, пусть задана последовательность $X_1, \ldots, X_n$ независимых и одинаково распределенных (н.о.р.) случайных величин с ф.р. $F(x - \theta)$, функциональный вид которой не известен, но предполагается, что $F \in \Im_S$ и $-\infty < \theta < +\infty$. При этих предположениях под параметром положения естественно понимать функционал $T(F)$, который соответствует точке симметрии ф.р. вероятностей, т.е. задача сводится к построению оценки функционала, характеризующего параметр θ, по наблюдениям $X_1, \ldots, X_n$. Для $F \in \Im_S$ параметр θ может быть представлен различными функционалами, которые условно можно разделить на четыре

типа (класса). Оценки функционалов строятся методом подстановки, путем замены ф.р. F на эмпирическую ф.р. F_n, построенную по наблюдениям $X_1, \ldots, X_n$. Рассматриваемые классы функционалов соответствуют М-оценкам (оценкам типа максимального правдоподобия); L-оценкам (оценкам в виде линейных комбинаций порядковых статистик); R-оценкам (оценкам, основанным на линейных ранговых статистиках); MD-оценкам (оценкам, построенным методом минимума расстояний). Конкретизируем перечисленные классы оценок и обсудим их свойства. Мы приведем результаты, полученные методом Мизеса, детали опускаем (они могут быть найдены в [1]). Приведенные ниже результаты являются основой для сравнения рассматриваемых классов робастных оценок и выбора подходящей оценки в зависимости от целей их пользователей.

М-оценки. Оценки параметра положения, предложенные Хьюбером в 1964 году, (см. [2]) были названы М-оценками, поскольку они включают в частном случае оценки максимального правдоподобия. М-оценки определяются как решение относительно T_n уравнения

$$\sum_{i=1}^{n} \psi(x_i - T_n) = 0, \tag{1}$$

где ψ – заданная функция, конкретизирующая М-оценку.

Пусть $\psi \in C_M$, где C_M – класс нечетных неубывающих дифференцируемых функций ψ, причем ψ' имеет компактный носитель, ограниченную вариацию и $\int \psi'(x)dF(x) \neq 0$. Отметим, что М-оценка, определяемая как решение (1) относительно T_n, может быть представлена в виде функционала от эмпирической ф.р., то есть в виде $T_n = T(F_n)$, где функционал задан неявно соотношением вида

$$\int \psi(x - T(F))dF(x) = 0. \tag{2}$$

Асимптотические свойства М-оценок подробно исследованы в литературе. В частности, для $\psi \in C_M$ и $F \in \Im_S$ М-оценки асимптотически нормальны, т.е.

$$L\left\{\frac{\sqrt{n}(T_n - T(F))}{\sigma(\psi, F)}\right\} = N(0;1), \tag{3}$$

где $\sigma^2(\psi, F)$ обозначает асимптотическую дисперсию М-оценки, определяемую в виде

$$\sigma^2(\psi, F) = \int IF^2(x; F, \psi)dF(x) = \frac{\int \psi^2(x)dF(x)}{(\int \psi'(x)dF(x))^2}. \tag{4}$$

Здесь $IF(x; F, \psi)$ обозначает функцию влияния М-оценки с функцией $\psi \in C_M$, которая равна

$$IF(x;F,\psi)=\frac{\psi(x)}{\int\psi'(x)dF(x)}. \tag{5}$$

Замечание. Отметим, что функция влияния М-оценки пропорциональна функции ψ, что позволяет конструировать М-оценки с нужными числовыми характеристиками робастности путем целенаправленного выбора функции ψ. Отметим также, что в параметрическом случае, когда семейство функций распределения определяется в виде $\Im_\theta=(F: F(x)=F(x,\theta),\theta\in\Theta)$, функционал $T(F)$, определяющий параметр θ и являющийся состоятельным по Фишеру, т.е. $T(F(x,\theta))=\theta$ для всех $\theta\in\Theta$, задается неявно уравнением $\int\psi(x,T(F))dF(x,\theta)=0$. При этом, если функция $\psi(x,\theta)$ согласована с плотностью $f(x,\theta)$ ф.р. $F(x,\theta)$ соотношением вида $\psi(x,\theta)=-dlnf(x,\theta)/d\theta$, то М-оценка параметра θ, полученная как решение последнего уравнения после замены ф.р. F на эмпирическую ф.р. F_n, является оценкой максимального правдоподобия.

L-оценки. Оценки параметра положения в виде линейной комбинации порядковых статистик определяются в виде

$$T_n=\sum_{i=1}^{n}a_{ni}X_{(i)}, \tag{6}$$

где весовые коэффициенты a_{ni} конкретизируются путем задания функции $h(t)\in C_L$, с помощью выражения

$$a_{ni}=\int\limits_{(i-1)/n}^{i/n}h(t)dt. \tag{7}$$

Здесь через C_L обозначен класс функций $h(t),0\le t\le 1$, которые непрерывны, имеют ограниченную вариацию, причем $h(t)=h(1-t)$ и $\int\limits_0^1 h(t)dt=1$. Отметим, что L-оценки из (6) также могут быть представлены в виде функционала от эмпирической ф.р., то есть в виде $T_n=T(F_n)$, где функционал $T(F)$ определяется следующим образом:

$$T(F)=\int\limits_0^1 h(t)F^{-1}(t)dt. \tag{8}$$

Для $F\in\Im_S$ и $h(t)\in C_L$ L-оценки асимптотически нормальны, т.е.

$$L\left\{\frac{\sqrt{n}(T_n-T(F))}{\sigma(h,F)}\right\}=N(0;1), \tag{9}$$

где $\sigma^2(h, F)$ – асимптотическая дисперсия $\sqrt{n}T_n$-оценки, определяемая как

$$\sigma^2(h, F) = \int_0^1 U^2(t)dt. \tag{10}$$

Здесь через $U(t)$ обозначен интеграл с переменным верхним пределом в виде

$$U(t) = \int_{1/2}^{t} \frac{h(t)}{f(F^{-1}(t))}dt. \tag{11}$$

Отметим, что функция влияния L-оценки с весовыми коэффициентами, определяемыми функцией $h(t) \in C_L$, равна

$$IF(x; F, h) = U(F(x))$$

и может быть сделана ограниченой функцией путем выбора подходящей функции $h(t) \in C_L$. Например, выбор функции $h(t)$ в виде $h(t) = 1/(1-2\alpha)$, $t \in [\alpha, (1-\alpha)]$, и $h(t) = 0$ в других случаях приводит к α-урезанному среднему с ограниченной функцией влияния при $0 < \alpha \leq 1/2$.

R-оценки. Оценки параметра положения, основанные на линейных ранговых статистиках, были предложены Ходжесом и Леманом в 1963 году (см., например, [1]). Поясним суть процедуры Ходжеса–Лемана. Пусть $V(X_1, \ldots, X_n)$ – статистика критерия проверки статистических гипотез

$$H_0 : \theta = 0, \qquad H_1 : \theta > 0.$$

Предположим, что статистика $V(X_1, \ldots, X_n)$ удовлетворяет следующим условиям:

1. Статистика $V(X_1, \ldots, X_n)$ правосторонняя, т.е. гипотеза H_0 отвергается для "больших" значений статистики.

2. Статистика $V(X_1 + l, \ldots, X_n + l)$ – неубывающая функция от l для фиксированных значений $X_1, \ldots, X_n$.

3. Распределение статистики $V(X_1, \ldots, X_n)$ при справедливости гипотезы H_0 является симметричным относительно некоторой точки ξ, т.е. плотность распределения вероятностей статистики V удовлетворяет соотношению $f_V(x \mid H_0) = f_V(2\xi - x \mid H_0)$.

Оценка Ходжеса–Лемана параметра θ определяется в виде $\theta_n = (\theta_n^* + \theta_n^{**})/2$, где $\theta_n^* = \mathrm{Sup}\{\theta : V(X_1 - \theta, \ldots, X_n - \theta) > \xi\}$, $\theta_n^{**} = \mathrm{Inf}\{\theta : V(X_1 - \theta, \ldots, X_n - \theta) < \xi\}$.

Замечание. Если статистика $V(X_1-\theta,\ldots,X_n-\theta)$ строго убывает по θ и является непрерывной функцией, тогда $\theta_n=\theta_n^{**}=\theta_n^*$ и при этом $V(X_1-\theta,\ldots,X_n-\theta)=\xi$.

Замечание. Процедура Ходжеса–Лемана первоначально была реализована в виде, когда статистика V соответствовала статистике рангового критерия для двувыборочной задачи с альтернативой сдвига. Эта задача формулируется следующим образом. Имеются две выборки $X_1,\ldots,X_n$ н.о.р. случайных величин с ф.р. $F(x)$ и $Y_1,\ldots,Y_m$ н.о.р. случайных величин с ф.р. $G(x)=F(x-\theta)$. Требуется проверить гипотезы

$$H_0^*:\ \theta=0,\qquad H_1^*:\ \theta>0.$$

Обозначим $N=n+m$ и $R_1,\ldots,R_m$ – ранги $Y_1,\ldots,Y_m$ в смешанной выборке, состоящей из $X_1,\ldots,X_n$ и $Y_1,\ldots,Y_m$. Статистики двувыборочных ранговых критериев для проверки гипотезы H_0^* против альтернативы H_1^* записываются в виде

$$S_n=m^{-1}\sum_{i=1}^{m}\varphi\Big(\frac{R_i}{N+1}\Big),$$

где $\varphi(t)$, $0\le t\le 1$ – заданная функция меток, конкретизирующая статистику рангового критерия. Обозначим через C_R класс непрерывных возрастающих функций меток $\varphi(t)$, $0\le t\le 1$, которые дифференцируемы, причем $\varphi(t)=-\varphi(1-t)$, $\int\limits_0^1\varphi(t)dt=0$ и для $F\in\Im_S$ $\varphi'(t)f(F^{-1})$ имеет ограниченную вариацию и $\int\limits_0^1\varphi'(t)f(F^{-1})dt\ne 0$.

Замечание. В случае, когда проверяются гипотезы H_0 и H_1, имеется лишь одна выборка $X_1,\ldots,X_n$, поэтому, чтобы воспользоваться статистикой S_N двувыборочного критерия, поступают следующим образом. Учитывая, что $F\in\Im_S$, исходную выборку $X_1,\ldots,X_n$ "симметризуют" относительно оценки T_N параметра θ, рассматривая в качестве второй выборки выборку вида $2T_n-X_1,\ldots,2T_n-X_n$. Окончательно, в соответствии с процедурой Ходжеса–Лемана, в качестве R-оценки T_N параметра положения θ выбирают такое значение T_N, которое является решением уравнения

$$\sum_{i=1}^{n}\varphi\Big(\frac{R_i}{2n+1}\Big)=0, \tag{12}$$

где R_i – ранг $2T_n-X_i$, $i=\overline{1,n}$, в смешанной выборке, состоящей из $X_1,\ldots,X_n$ и $2T_n-X_1,\ldots,2T_n-X_n$. Отметим, что R-оценка T_n, являющаяся решением уравнением (12), может быть представлена в виде

функционала от эмпирической ф.р., т. е. в виде $T_n = T(F_n)$, где функционал $T(F)$ задается неявно соотношением вида

$$\int \varphi\{[F(x) + 1 - F(2T(F) - x)]/2\} dF(x) = 0. \tag{13}$$

Для $F \in \Im_S$ и $\varphi(t) \in C_R$ R-оценки параметра θ асимптотически нормальны, т. е.

$$L\left\{\frac{\sqrt{n}(T_n - T(F))}{\sigma(\varphi, F)}\right\} = N(0;1), \tag{14}$$

где $\sigma^2(\varphi, F)$ – асимптотическая дисперсия $\sqrt{n}T_n$-оценки, определяемая в виде

$$\sigma^2(\varphi, F) = \frac{\int\limits_0^1 \varphi^2(t)dt}{\left\{\int\limits_0^1 \varphi'(t) f(F^{-1}(t))dt\right\}^2}. \tag{15}$$

Отметим, что функция влияния $IF(x; F, \varphi)$ R-оценки, метки которой определяются конкретизацией функции $\varphi(t) \in C_R$, имеет вид

$$IF(x; F, \varphi) = \frac{\varphi(F(x))}{\int\limits_0^1 \varphi'(t) f(F^{-1}(t))dt}. \tag{16}$$

Пример 1. Пусть функция меток $\varphi(t) = \Phi^{-1}(t)$, $0 \le t \le 1$, где $\Phi^{-1}(t)$ – квантильная функция стандартного нормального распределения. В этом случае R-оценка соответствует ранговому критерию Ван-дер-Вардена и называется R-оценкой с нормальными метками.

Пример 2. Пусть функция меток $\varphi(t) = t - 1/24$, $0 \le t \le 1$. В этом случае R-оценке, которая носит название оценки Ходжеса–Лемана, соответствует функционал $T(F)$, заданный выражением $\int F(2T(F) - x)dF(x) = 1/2$, и оценка выписывается в явном виде через исходные наблюдения $X_1, \ldots, X_n$ в следующей форме:

$$T_n = HL = \text{med}\{(X_i + X_j)/2; 1 \le i, j \le n\}, \tag{17}$$

где символ med обозначает выборочную медиану. Обобщение этой оценки на случай урезанной выборки обсуждается в [1].

MD-оценки. Оценки, построенные методом минимума расстояний, называемые в литературе MD-оценками, были предложены Вольфовитцем в 1957 году. Кратко обсудим суть метода минимума расстояний.

Пусть задана последовательность $X_1, \ldots, X_n$ н.о.р. случайных величин с ф.р. F. Предполагаем, что ф.р. F принадлежит заданному параметрическому семейству $\Im_\theta = (F : F(x) = F(x,\theta), \theta \in \Theta)$. На множестве функций распределений для пары ф.р. F, G введем расстояние $d(F,G)$. MD-оценка параметра θ, построенная по методу минимума расстояния d, определяется в виде

$$\theta_n = \arg\min_\theta d(F_n, F(x,\theta)).$$

Рассмотрим MD-оценки параметра θ для случая, когда $F(x,\theta) = F(x-\theta)$, основаны на взвешенном расстоянии Крамера–Мизеса, определяемом в виде

$$\int [F_n(x) - F_0(x-\theta)]^2 \omega(x-\theta)dx,$$

где F_0 – выбранная из некоторых соображений опорная ф.р., F_n – эмпирическая ф. р. и ω – заданная весовая функция. Отметим, что выбор весовой функции в виде $\omega(x) = f_0(x)$, где f_0 – плотность ф.р. F_0, соответствует расстоянию Крамера–Мизеса, и MD-оценка находится путем минимизации по θ следующего выражения:

$$\frac{1}{n}\sum_{i=1}^{n}\left[F_0(X_{(i)} - \theta) - \frac{2i-1}{2n}\right]^2 + \frac{1}{12n^2}.$$

При выборе весовой функции в виде $\omega(x) \equiv 1$ получаем MD-оценку, которая совпадает с M-оценкой, т.е. является решением уравнения

$$\sum_{i=1}^{n} \psi(X_{(i)} - \theta) = 0 \tag{18}$$

для $\psi(x) = F_0(x) - 1/2$. При использовании взвешенного расстояния Крамера–Мизеса для заданной весовой функции ω, MD-оценки параметра находятся путем решения уравнения

$$\sum_{i=1}^{n}\left[\frac{i-1/2}{n} - F_0(X_{(i)} - \theta)\right]\omega(X_{(i)} - \theta) = 0. \tag{19}$$

Отметим, что MD-оценки также могут быть представлены в виде функционала от эмпирической ф.р., т.е. в виде $T(F_n)$, причем функци-

онал $T(F) = \theta$ задается неявно уравнением вида

$$2\int [F(x+T(F)) - F_0(x)]f_0(x)\omega(x)dx - \\ - \int [F(x+T(F)) - F_0(x)]^2\omega'(x)dx = 0. \tag{20}$$

Для нахождения асимптотического распределения MD-оценок воспользуемся методом Мизеса. Проиллюстрируем вычисление функций влияния оценок на примере MD-оценок. Обозначим через $F_{\lambda,u}$ ф.р. вида $F_{\lambda,u} = F(x) + \lambda[C(u-x) - F(x)]$, где $C(u)$ – единичная функция. Подставив в (20) ф.р. $F_{\lambda,u}$ вместо F, дифференцируя по λ и полагая $\lambda = 0$, получим выражение для функции влияния MD-оценки с весовой функцией ω. Полученное выражение упростим, рассмотрев дополнительные предположения. С учетом инвариантности MD-оценок относительно сдвига без потери общности положим $\theta = 0$.

Рассмотрим четные весовые функции, т.е. $\omega(x) = \omega(-x)$, $\forall x \in R$, и пусть F, $F_0 \in \Im_S$. При этих предположениях интегрирование по частям числителя полученного выражения позволяет записать функцию влияния в виде

$$IF(u; F, \omega) = A_F(u)/B_F, \tag{21}$$

$$A_F(u) = \int_0^u \omega(x)dF(x) - \omega(u)[F(u) - F_0(u)], \tag{22}$$

$$B_F = \int \{f_0(x)\omega(x) - [F(x) - F_0(x)]\omega'(x)\}dF(x). \tag{23}$$

Замечание. Отметим, что при определенном выборе весовой функции ω, функцию влияния MD-оценки можно сделать ограниченной. Например, при $F = F_0$ и $\omega(x) = f_0(x)$, используя формулу (21), можно показать, что

$$IF(x; F, \omega) = \frac{\int_0^x f^2(y)dy}{\int f^3(x)dx} \tag{24}$$

и, в частности, при нормальном распределении, т.е. при $F = \Phi$, получаем ограниченную функцию влияния в виде

$$IF(x; F, \omega) = \frac{\sqrt{3\pi}}{2}(2\Phi(x\sqrt{2}) - 1), \tag{25}$$

где Φ – стандартное нормальное распределение, а φ – его плотность. Обозначим через C_{MD} класс неотрицательных, четных, непрерывных и дифференцируемых весовых функций ω. Для $\omega \in C_{\mathrm{MD}}$ и $F, F_0 \in \Im_S$ MD-оценки, являющиеся решением (18), асимптотически нормальны (детали см. в [1]), т.е.

$$L\left\{\frac{\sqrt{n}(\theta_n - \theta)}{\sigma(\omega, F)}\right\} = N(0;1), \tag{26}$$

где $\sigma^2(\omega, F)$ – асимптотическая дисперсия $\sqrt{n}T_n$-оценки, определяемая в виде

$$\sigma^2(\omega, F) = \frac{2}{B_F^2}\int\limits_0^\infty A_F^2(x)dF(x) = \frac{2}{B_F^2}\int\limits_{1/2}^1 A_F^2(F^{-1}(t))dt. \tag{27}$$

Отметим, что для случая, когда опорная ф.р. F_0 удачно выбрана и совпадает с ф.р. F наблюдений (т.е. при $F = F_0$), получаем

$$\sigma^2(\omega, F) = \int\limits_{-\infty}^\infty \left[\int\limits_0^x \omega(y)dF(y)\right]^2 dF(x)\left[\int\limits_{-\infty}^\infty f^2(x)\omega(x)dx\right]^{-2}. \tag{28}$$

Замечание. Из выражения (28) следует, что выбор весовой функции ω в виде

$$\omega(x) = a\frac{d^2(-lnf_0(x))}{dx^2}\Big/ f_0(x) \tag{29}$$

приводит к асимптотически эффективным MD-оценкам (т.е. при $F = F_0$ и при выполнении (29)), асимптотическая дисперсия MD-оценки равна обратной величине количества информации Фишера.

Связи между классами робастных оценок

Для установления связей между рассмотренными классами робастных оценок параметра положения θ введем следующие отображения

между классами функций C_{M}, C_{L}, C_{R}:

$$C_{\mathrm{M}} \to C_{\mathrm{L}} : h(t) = \psi'(F^{-1}(t)) \Big/ \int \psi'(x)dF(x),\ 0 < t < 1; \tag{30}$$

$$C_{\mathrm{L}} \to C_{\mathrm{M}} : \psi(x) = \int_0^x h(F(y)dy),\ -\infty < x < +\infty; \tag{31}$$

$$C_{\mathrm{M}} \to C_{\mathrm{R}} : \varphi(t) = \psi(F^{-1}(t)),\ 0 < t < 1; \tag{32}$$

$$C_{\mathrm{R}} \to C_{\mathrm{M}} : \psi(x) = \varphi(F(x)),\ -\infty < x < +\infty; \tag{33}$$

$$C_{\mathrm{R}} \to C_{M} : \psi(x) = \varphi(F(x)),\ -\infty < x < +\infty; \tag{34}$$

$$C_{\mathrm{R}} \to C_{\mathrm{L}} : h(t) = \varphi'(t)f(F^{-1}(t)) \Big/ \int_0^1 \varphi'(t)f(F^{-1}(t)),\ 0<t<1. \tag{35}$$

Отображения (30)–(35) обеспечивают равенство асимптотических дисперсий оценок, т.е. $\sigma^2(\psi, F) = \sigma^2(h, F) = \sigma^2(\varphi, F)$ при $F \in \Im_s$. Этот факт, который может быть проверен непосредственно с использованием приведенных формул (4), (10), (15), впервые был установлен Джекелем в 1971 году, известен в литературе как теорема Джекеля и играет важную роль в получении робастных оценок с оптимальными свойствами.

Например, для случая $F(x, \theta) = F(x - \theta)$, $-\infty < \theta < +\infty$, выбор функции $\psi(x)$ в виде $\psi_F(x) = -f'(x)/f(x)$ приводит к асимптотически эффективной M-оценке параметра θ, т.е. $\sigma^2(\psi_F, F) = 1/I(F)$, где $I(F)$ – количество информации Фишера.

Используя (30), получаем асимптотически эффективную L-оценку при $h_F(t) = \psi'_F(F^{-1}(t))/I(F)$, и, согласно (33), асимптотически эффективная R-оценка соответствует выбору функции $\varphi(t)$ в виде $\varphi_F(t) = -f'(F^{-1}(t))/f(F^{-1}(t))$.

Пример 3. Рассмотрим супермодель в виде гауссовской модели с симметричным засорением, т.е. предполагаем, что $F \in \Im_\epsilon(\Phi)$, где множество $\Im_\epsilon(\Phi)$ определено в виде

$$\Im_\epsilon(\Phi) = \{F : F(x) = (1 - \epsilon)\Phi(x) + \epsilon H(x)\}, 0 \le \epsilon \le 1. \tag{36}$$

Здесь Φ – стандартная нормальная ф.р., $H \in \Im_S$ – произволное распределение и ϵ – заданная пропорция засорения. Построим минимаксно-робастную R-оценку в супермодели $\Im_\epsilon(\Phi)$. Для этого необходимо конкретизировать функцию меток R-оценки. Обозначив ее через φ_0, используя отображение (32) и формулу (29), получаем

$$\varphi_0(t) = \Phi^{-1}\left\{\frac{\alpha_0 - \epsilon/2}{1 - \epsilon}\right\},\ 0 < t < \alpha_0;$$

$$\varphi_0(t) = \Phi^{-1}\left\{\frac{t-\epsilon/2}{1-\epsilon}\right\},\ \alpha_0 \le t \le 1-\alpha_0;$$

$$\varphi_0(t) = \Phi^{-1}\left\{\frac{1-\alpha_0-\epsilon/2}{1-\epsilon}\right\},\ 1-\alpha_0 < t \le 1, \qquad (37)$$

где параметр α_0 при заданной пропорции засорения ϵ вычисляется по формуле

$$\alpha_0 = \epsilon/2 + (1-\epsilon)\alpha', \qquad (38)$$

а α' является решением уравнения

$$2\varphi\{\Phi^{-1}(\alpha')\}/\Phi^{-1}(1-\alpha') - 2\alpha' = \epsilon/(1-\epsilon). \qquad (39)$$

Здесь Φ^{-1} обозначает обратную функцию стандартного нормального распределения. Отметим, что в отличие от R-оценки с нормальными метками, соответствующей функции $\varphi(t) = \Phi^{-1}(t)$, $0 \le t \le 1$, функция влияния минимаксно-робастной R-оценки, для которой функция φ_0 определена в (37), является ограниченной и, следовательно, R-оценка будет подвержена лишь ограниченному влиянию грубых ошибок (выбросов) в выборке.

Пример 4. Построим минимаксно-робастную в оценку в классе L-оценок. Для этого необходимо конкретизировать функцию h, определяющую весовые коэффициенты L-оценки. Обозначим ее через $h_0(t)$. Используя отображение (30), с учетом того, что $\psi_0'(F^{-1}(t)) = 1$ для $t \in [\alpha_0, 1-\alpha_0]$ и равна нулю вне этого интервала, а также, что $\int \psi_0'(F_o^{-1}(t))dt = 1-2\alpha_0$, получаем функцию h_0 в виде $h_0(t) = 1/(1-2\alpha_0)$ для $t \in [\alpha_0, 1-\alpha_0]$ и равную нулю в других случаях. Таким образом, минимаксно-робастной L-оценкой в супермодели $\Im_\epsilon(\Phi)$ является α_0-урезанное среднее, где α_0 определено в (38).

Замечание. Изложенные результаты остаются справедливыми и для более общей супермодели вида

$$\Im_\epsilon(G) = \{F:\ F(x) = (1-\epsilon)G(x) + \epsilon H(x)\},\ 0 \le \epsilon \le 1, \qquad (40)$$

где $G(x)$ – заданная непрерывная симметричная ф.р. с абсолютно непрерывной плотностью $g(x)$ и $H(x)$ – произвольная непрерывная симметричная ф.р.

Замечание. Обсудим кратко связи между MD-оценками и классами M-, L- и R-оценок . Пусть $\omega(x) \in C_{\mathrm{MD}}$, определим отображения между классами функций C_{MD}, C_{M}, C_{L} и C_{R} в следующем виде:

$$C_{\mathrm{MD}} \to C_{\mathrm{M}}:\ \psi(x) = \int_0^x \omega(y)dF(y),\ -\infty < x < +\infty; \qquad (41)$$

$$C_{\mathrm{M}} \to C_{\mathrm{MD}} : \omega(x) = \psi'(x)/f(x),\ -\infty < x < +\infty; \tag{42}$$

$$C_{\mathrm{MD}} \to C_{\mathrm{L}} : h(t) = \omega(F^{-1}(t))f(F^{-1}(t)),\ 0 < t < 1; \tag{43}$$

$$C_{\mathrm{L}} \to C_{\mathrm{MD}} : \omega(x) = h(F(x))/f(x),\ -\infty < x < +\infty; \tag{44}$$

$$C_{\mathrm{MD}} \to C_{\mathrm{R}} : \omega(x) = \varphi'(F(x)),\ -\infty < x < +\infty; \tag{45}$$

$$C_{\mathrm{R}} \to C_{\mathrm{MD}} : \varphi(t) = \int_0^{F^{-1}(t)} \omega(x)dF(x),\ 0 < t < 1. \tag{46}$$

Для отображений (41)–(46) между классами функций, конкретизирующими MD-, M-, L-, и R-оценки, обобщение теоремы Джекеля можно сформулировать в следующем виде. Пусть $F \in \Im_S$, тогда для случая, когда ф.р. наблюдений совпадает с выбранной опорной ф.р. (т.е. при $F = F_0$) и при справедливости отображений (41)–(46), асимптотическая дисперсия MD-оценки равна соответственно асимптотическим дисперсиям M-, L- и R-оценок. Справедливость этого факта проверяется непосредственно с использованием формул (4), (10), (15), (27) при выполнении (41)–(46).

Пример 5. Пусть $F \in \Im_S$ и $\omega(x) \in C_{\mathrm{MD}}$. Тогда при $F = F_0$ MD-оценка является минимаксно-робастной в супермодели $\Im_S$ при выборе весовой функции $\omega(x)$ в виде

$$\omega(x) = I\{|x| \le a\}\frac{d^2(lng(x))}{dx^2}/(1-\epsilon)g(x),$$

где $I\{x\}$ – единичная функция и параметр a при заданном ϵ находится из уравнения $\epsilon/2(1-\epsilon) = G(a) - (g'(a)g^2(a))/g'(a)$.

Заключение

В данной работе мы обсудили основные типы робастных статистических процедур на примере задачи оценивания параметра положения случайной величины. В методологическом отношении эта задача является базовой и служит основой для обобщения и построения робастных процедур в более общих статистических моделях (см., например, [2] и [3]). Изложенные робастные процедуры находят широкое практическое применение. Многие из них уже включены в современные зарубежные пакеты программ, и их успешное и осознанное применение на практике определяется знанием свойств этих процедур. Обсуждению этих свойств и была посвящена данная работа.

СПИСОК ЛИТЕРАТУРЫ

1. *Шуленин В. П.* Введение в робастную статистику. Томск: Изд-во Том. ун-та, 1993. 227 с.
2. *Хьюбер П.* Робастность в статистике / Пер. с англ. под ред. И.Г. Журбенко. М.: Мир, 1984. 303 с.
3. *Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В.* Робастность в статистике. Подход на основе функций влияния / Пер. с анг. под ред. В.М. Золотарчва. М.: Мир, 1989. 512 с.

III. Статистические приложения

Н.С. Демин, С.В. Рожкова, О.В. Рожкова

Оценивание и распознавание в стохастических системах при непрерывно-дискретных наблюдениях со скользящей памятью

Рассматриваются задачи оценивания и распознавания стохастических процессов по непрерывно-дискретным наблюдениям со скользящей памятью. Используемая в постановке задачи параметрическая неопределенность позволяет рассмотреть обе задачи в рамках единого подхода.

1. Постановка задачи

Для случая непрерывно-дискретных наблюдений с памятью в [1] рассмотрена задача фильтрации, в [2] – задача экстраполяции, в [3] – задача распознавания с фиксированной памятью. В данной работе результаты [3] обобщаются на случай наблюдений со скользящей памятью. Используемые обозначения: $P\{\cdot\}$ – вероятность события; $M\{\cdot\}$ – математическое ожидание; $\mathcal{N}\{\cdot\}$ – гауссова плотность с параметрами a и B; $|\cdot|$ – определитель матрицы; $<T>$ – транспонирование как верхний правый индекс.

На вероятностном пространстве [4] $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ ненаблюдаемый n-мерный процесс x_t и наблюдаемый l-мерный процесс z_t определяются дифференциальными уравнениями (в смысле Ито):

$$dx_t = f(t, x_t, z, \theta)dt + \Phi_1(t, x_t, z, \theta)dw_t, \quad t \geq 0; \tag{1}$$

$$dz_t = h(t, x_t, x_{t-t_1^*}, \cdots, x_{t-t_N^*}, z, \theta)dt + \Phi_2(t, z)dv_t, \tag{2}$$

а наблюдаемый q-мерный процесс с дискретным временем $\eta(t_m)$ имеет вид

$$\eta(t_m) = g(t_m, x_{t_m}, x_{t_m-t_1^*}, \cdots, x_{t_m-t_N^*}, z, \theta) + \Phi_3(t_m, z, \theta)\xi(t_m), \tag{3}$$

где: $\Phi_1(\cdot)\Phi_1^T(\cdot) = Q(\cdot)$, $\Phi_2(\cdot)\Phi_2^T(\cdot) = R(\cdot)$, $\Phi_3(\cdot)\Phi_3^T(\cdot) = V(\cdot)$; $m \geq 0$; $0 \leq t_0 \leq t - t_N^* < \cdots < t - t_1^* < t_m \leq t$, $t_k^* = \text{const}$, $k = \overline{1, N}$. Остальные предположения те же, что и в [3].

Подобным образом заданная параметрическая неопределенность позволяет в рамках единого подхода рассмотреть задачи оценивания и распознавания.

З а д а ч а 1 (оценивание). По совокупности реализаций

$$\mathcal{F}_t^{z,\eta} = \{z_0^t; \eta_0^m\}, \text{ где } z_0^t = \{z_s; 0 \le s \le t\},$$

$$\eta_0^m = \{\eta(t_0), \eta(t_1), \cdots, \eta(t_m);\ 0 \le t_0 < \cdots < t_m \le t\}$$

найти оптимальные в среднеквадратическом смысле оценки $\mu(t)$ и $\mu(t - t_k^*, t)$ соответственно для значений x_t и $x_{t-t_k^*}$, $k = \overline{1; N}$.

З а д а ч а 2 (распознавание). По совокупности реализаций $\mathcal{F}_t^{z,\eta} = \{z_0^t; \eta_0^m\}$ найти отношение правдоподобия $\Lambda_t(\theta_j : \theta_\alpha)$ в задаче распознавания гипотез $\mathcal{H}_j\{\theta = \theta_j\}$ и $\mathcal{H}_\alpha\{\theta = \theta_\alpha\}$.

2. Основные результаты

Введем рассширенные процессы ($k = \overline{1; N}$)

$$\widetilde{x}_{t-t^*}^{N-k+1} = \begin{bmatrix} x_{t-t_k^*} \\ x_{t-t_{k+1}^*} \\ \vdots \\ x_{t-t_N^*} \end{bmatrix}, \qquad \widetilde{x}_{t,t-t^*}^{N+1} = \begin{bmatrix} x_t \\ \widetilde{x}_{t-t^*}^N \end{bmatrix} \tag{4}$$

и операторы

$$L_{s,y}[\varphi(s,y,\cdot)] = -\sum_{i=1}^n \frac{\partial[f_i(s,y,\cdot)\varphi(s,y,\cdot)]}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2[Q_{ij}(s,y,\cdot)\varphi(s,y,\cdot)]}{\partial y_i \partial y_j};$$

$$L_{s,y}^*[\varphi(s,y,\cdot)] = \sum_{i=1}^n f_i(s,y,\cdot)\frac{\partial\varphi(s,y,\cdot)}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^n Q_{ij}(s,y,\cdot)\frac{\partial^2\varphi(s,y,\cdot)}{\partial y_i \partial y_j};$$

$$\mathcal{L}_{s,y}[\varphi_1(s,y,\cdot); \varphi_2(s,y,\cdot)] =$$

$$= \frac{\varphi_1(s,y,\cdot)}{\varphi_2(s,y,\cdot)} L_{s,y}[\varphi_2(s,y,\cdot)] - \varphi_2(s,y,\cdot) L_{s,y}^*\left[\frac{\varphi_1(s,y,\cdot)}{\varphi_2(s,y,\cdot)}\right]. \tag{5}$$

Пусть для $j = \overline{0; r}$, $k = \overline{1; N}$, $t_{m_k} \le t - t_k^* \le t_{m_k+1}$;

$$p_t(\theta_j) = P\{\theta = \theta_j | z_0^t, \eta_0^m\}; \tag{6}$$

$$p_t(x, \widetilde{x}_N | \theta_j) = \frac{\partial^{N+1} P\{x_t \le x; \widetilde{x}_{t-t^*}^N \le \widetilde{x}_N | \theta_j, z_0^t, \eta_0^m\}}{\partial x \partial \widetilde{x}_N}; \tag{7}$$

$$p_{t-t_k^*}(\widetilde{x}_{N-k+1}|\theta_j) = \frac{\partial^{N-k+1} P\{\widetilde{x}_{t-t^*}^{N-k+1} \le \widetilde{x}_{N-k+1}|\theta_j, z_0^{t-t_k^*}, \eta_0^{m_k}\}}{\partial \widetilde{x}_{N-k+1}}. \quad (8)$$

Теорема 1. *Апостериорные вероятности (6) гипотез* $\mathcal{H}_j\{\theta=\theta_j\}$ *и апостериорные условные плотности вероятностей (7) на интервалах времени* $t_m \le t < t_{m+1}$ *определяются уравнениями*

$$d_t p_t(\theta_j) = p_t(\theta_j)[\overline{h(t,z|\theta j)} - \overline{h(t,z)}]^T R^{-1}(t,z)[dz_t - \overline{h(t,z)}dt]; \quad (9)$$

$$d_t p_t(x;\widetilde{x}_N|\theta_j) = \{L_{t,x}[p_t(x;\widetilde{x}_N|\theta_j)] +$$

$$+\sum_{k=1}^{N} \mathcal{L}_{t-t_k^*,x_k}[p_t(x;\widetilde{x}_N|\theta_j); p_{t-t_k^*}(\widetilde{x}_{N-k+1}|\theta_j)]\}dt + p_t(x;\widetilde{x}_N|\theta_j) \times$$

$$\times[h(t,x,\widetilde{x}_N,z,\theta j) - \overline{h(t,z|\theta_j)}]^T R^{-1}(t,z)[dz_t - \overline{h(t,z|\theta_j)}dt] \quad (10)$$

с начальными условиями

$$p_{t_m}(\theta_j) = \frac{C(\eta(t_m),z|\theta_j)}{C(\eta(t_m),z)} p_{t_m-0}(\theta_j); \quad (11)$$

$$p_{t_m}(x;\widetilde{x}_N|\theta_j) = \frac{C(\eta(t_m),z,x,\widetilde{x}_N,\theta_j)}{C(\eta(t_m),z|\theta_j)} p_{t_m-0}(x;\widetilde{x}_N|\theta_j); \quad (12)$$

$$\overline{h(t,z)} = M\{h(t,x_t,\widetilde{x}_{t-t^*}^N,z,\theta)|z_0^t,\eta_0^m\}; \quad (13)$$

$$\overline{h(t,z|\theta_j)} = M\{h(t,x_t,\widetilde{x}_{t-t^*}^N,z,\theta)|\theta=\theta_j,z_0^t,\eta_0^m\}; \quad (14)$$

$$C(\eta(t_m),z) = M\{C(\eta(t_m),z,x_{t_m},\widetilde{x}_{t_m-t^*}^N,\theta)|z_0^{t_m},\eta_0^{m-1}\}; \quad (15)$$

$$C(\eta(t_m),z|\theta_j) = M\{C(\eta(t_m),z,x_{t_m},\widetilde{x}_{t_m-t^*}^N,\theta)|\theta=\theta_j,z_0^{t_m},\eta_0^{m-1}\}; \quad (16)$$

$$C(\eta(t_m),z,x,\widetilde{x}_N,\theta_j) = |V(t_m,z,\theta_j)|^{-1/2}\exp\{(-1/2)[\eta(t_m) -$$

$$-g(t_m,x,\widetilde{x}_N,z,\theta_j)]^T V^{-1}(t_m,z,\theta_j)[\eta(t_m)-g(t_m,x,\widetilde{x}_N,z,\theta_j)]\}. \quad (17)$$

Доказательство. Для

$$p_t(x;\widetilde{x}_N;\theta_j) = \frac{\partial^{N+1} P\{x_t \le x;\widetilde{x}_{t-t^*}^N \le \widetilde{x}_N, \theta=\theta_j|z_0^t,\eta_0^m\}}{\partial x \partial \widetilde{x}_N} \quad (18)$$

на интервалах $t_m \le t < t_{m+1}$ справедливо уравнение

$$d_t p_t(x;\widetilde{x}_N;\theta_j) = \Big\{L_{t,x}[p_t(x;\widetilde{x}_N;\theta_j)] +$$

$$+\sum_{k=1}^{N} \mathcal{L}_{t-t_k^*,x_k}[p_t(x;\widetilde{x}_N;\theta_j); p_{t-t_k^*}(\widetilde{x}_{N-k+1};\theta_j)]\Big\}dt + p_t(x;\widetilde{x}_N;\theta_j) \times$$

$$\times[h(t,x,\widetilde{x}_N,z,\theta_j) - \overline{h(t,z)}]^T R^{-1}(t,z)[dz_t - \overline{h(t,z)}dt] \quad (19)$$

с начальным условием

$$p_{t_m}(x;\widetilde{x}_N;\theta_j)=\frac{C(\eta(t_m),z,x,\widetilde{x}_N,\theta_j)}{C(\eta(t_m),z)}p_{t_m-0}(x;\widetilde{x};\theta_j), \tag{20}$$

которое следует из теоремы 1 в [1; I]. Так как

$$p_t(x;\widetilde{x}_N;\theta_j)=p_t(x;\widetilde{x}_N|\theta_j)p_t(\theta_j), \tag{21}$$

то интегрирование (19), (20) по $\{x;\widetilde{x}_N\}$ с учетом (14), (16) дает (9), (11), а (12) следует из (11), (20), (21). Так как инновационный процесс

$$\widetilde{z}_t=\int\limits_0^t[dz_s-\overline{h(s,z)}ds] \tag{22}$$

такой, что $\widetilde{Z}=(\widetilde{z}_t,\mathcal{F}_t^z)$ является винеровским процессом с

$$M\{\widetilde{z}_t\widetilde{z}_\tau^T|\mathcal{F}^z_{\max(t,\tau)}\}=\int\limits_0^{\min(t,\tau)}R(s,z)ds, \tag{23}$$

то дифференцирование по формуле Ито

$$p_t(x;\widetilde{x}_N|\theta_j)=p_t(x;\widetilde{x}_N;\theta_j)/p_t(\theta_j) \tag{24}$$

с учетом (9), (19) дает (10).

Теорема 2. *Оптимальные в среднеквадратическом смысле оценки фильтрации $\mu(t)$ и интерполяции $\mu(t-t_k^*,t)$, $k=\overline{1;N}$, определяются формулами*

$$\mu(t)=\sum_{j=0}^{r}\mu(t|\theta_j)p_t(\theta_j),\quad \mu(t-t_k^*,t)=\sum_{j=0}^{r}\mu(t-t_k^*,t|\theta_j)p_t(\theta_j), \tag{25}$$

где (см. (4), (7))

$$\mu(t|\theta_j)=\int\cdots\int xp_t(x;\widetilde{x}_N|\theta_j)dxd\widetilde{x}_N;$$
$$\mu(t-t_k^*,t|\theta_j)=\int\cdots\int x_kp_t(x;\widetilde{x}_N|\theta_j)dx\,d\widetilde{x}_N, \tag{26}$$

а $p_t(\theta_j)$, $p_t(x;\widetilde{x}_N|\theta_j)$ определены в теореме 1.

Доказательство. Поскольку оптимальной в среднеквадратическом смысле оценкой является апостериорное среднее [4], то

$$\mu(t) = M\{x_t|z_0^t, \eta_0^m\}; \quad \mu(t-t_k^*, t) = M\{x_{t-t_k^*}|z_0^t, \eta_0^m\}. \tag{27}$$

Так как $\mathcal{F}_t^{z,\eta} \subseteq \mathcal{F}_t^{z,\eta,\theta}$, то, согласно (27),

$$\mu(t) = M\{M\{x_t|\mathcal{F}_t^{z,\eta,\theta_j}\}|\mathcal{F}_t^{z,\eta}\} = M\{\mu(t|\theta_j)|\mathcal{F}_t^{z,\eta}\};$$
$$\mu(t-t_k^*, t) = M\{M\{x_{t-t_k^*}|\mathcal{F}_t^{z,\eta,\theta_j}\}|\mathcal{F}_t^{z,\eta}\} = M\{\mu(t-t_k^*|\theta_j)|\mathcal{F}_t^{z,\eta}\}. \tag{28}$$

Формулы (25) следуют из (6), (7), (26), (28).

Теорема 3. *Отношение правдоподобия* $\Lambda_t(\theta_j : \theta_\alpha)$ *в задаче распознавания гипотез* $\mathcal{H}_j\{\theta = \theta_j\}$ *и* $\mathcal{H}_\alpha\{\theta = \theta_\alpha\}$, $j = \overline{0;r}$, $\alpha = \overline{0;r}$, *определяется формулой*

$$\Lambda_t(\theta_j : \theta_\alpha) = \Lambda_{t_N^*-0}(\theta_j : \theta_\alpha) \exp\Bigg\{ \sum_{t_N^* \le t_i \le t} \ln\left[\frac{C(\eta(t_i), z|\theta_j)}{C(\eta(t_i), z|\theta_\alpha)}\right] + \tag{29}$$

$$+ \int_{t_N^*}^{t} [\overline{h(s,z|\theta_j)} - \overline{h(s,z|\theta_\alpha)}]^T R^{-1}(s,z)[dz_s - \frac{1}{2}(\overline{h(s,z|\theta_j)} + \overline{h(s,z|\theta_\alpha)})ds]\Bigg\}.$$

Доказательство проводится по методике доказательства утверждения 2 в [3].

Из теорем 1–3 следует, что эффективное вычисление $\mu(t)$, $\mu(t-t_k^*, t)$, $\Lambda_t(\theta_j : \theta_\alpha)$ реализуется при возможности эффективного вычисления $\mu(t|\theta_j)$, $\mu(t-t_k^*, t|\theta_j)$, $\overline{h(t,z|\theta_j)}$, $C(\eta(t_m), z|\theta_j)$. В следующем пункте рассматривается частный случай процессов x_t, z_t, $\eta(t_m)$, допускающий такую возможность.

3. Условно-гауссовский случай

Лемма. *Пусть (см. (4))*

$$f(\cdot) = f(t,z,\theta) + F(t,z,\theta)x_t, \quad \Phi_1(\cdot) = \Phi_1(t,z,\theta);$$

$$h(\cdot) = h(t,z,\theta) + H_{0,N}(t,z,\theta)\widetilde{x}^{N+1}_{t,t-t^*};$$

$$g(\cdot) = g(t_m,z,\theta) + G_{0,N}(t_m,z,\theta)\widetilde{x}^{N+1}_{t_m,t_m-t^*};$$

$$p_0(x|\theta_j) = \mathcal{N}\{x;\mu_0^j,\Gamma_0^j\}, \quad j=\overline{0;r}; \tag{30}$$

$$H_{0,N}(\cdot) = [H_0(\cdot)\vdots H_1(\cdot)\vdots\cdots\vdots H_N(\cdot)],\ G_{0,N}(\cdot) = [G_0(\cdot)\vdots G_1(\cdot)\vdots\cdots\vdots G_N(\cdot)]. \tag{31}$$

Тогда

$$(t-\widetilde{t}^*_N = [t-t^*_1,\cdots,t-t^*_N]);$$

$$p_t(x;\widetilde{x}_N|\theta_j) = p_t(\widetilde{x}_{N+1}|\theta_j) =$$

$$= \mathcal{N}\{\widetilde{x}_{N+1};\widetilde{\mu}_{N+1}(t-\widetilde{t}^*_N,t|\theta_j),\widetilde{\Gamma}_{N+1}(t-\widetilde{t}^*_N,t|\theta_j)\}; \tag{32}$$

$$\beta_{N+1}(\cdot|\theta_j) = \begin{matrix}\mu(t|\theta_j)\\ \widetilde{\mu}_N(t-\widetilde{t}^*_N,t|\theta_j)\end{matrix} = \begin{matrix}\mu(t|\theta_j)\\ \mu(t-t^*_k,t|\theta_j)\end{matrix}, \ k=\overline{1;N}; \tag{33}$$

$$\widetilde{\Gamma}_{N+1}(\cdot|\theta_j) = \begin{bmatrix}\Gamma(t|\theta_j) & \widetilde{\Gamma}_{0N}(t-\widetilde{t}^*_N,t|\theta_j)\\ \widetilde{\Gamma}^T_{0N}(\cdot) & \widetilde{\Gamma}_N(t-\widetilde{t}^*_N,t|\theta_j)\end{bmatrix} = \tag{34}$$

$$= \begin{bmatrix}\Gamma(t|\theta_j) & \Gamma_{01}(t-t^*_1,t|\theta_j) & \Gamma_{0k}(t-t^*_k,t|\theta_j)\\ \Gamma^T_{01}(\cdot) & \Gamma_{11}(t-t^*_1,t|\theta_j) & \Gamma_{lk}(t-t^*_l,t-t^*_k,t|\theta_j)\\ \Gamma^T_{0k}(\cdot) & \Gamma^T_{lk}(\cdot) & \Gamma_{kk}(t-t^*_k,t|\theta_j)\end{bmatrix} \begin{matrix} l=\overline{1;N-1},\\ k=\overline{2;N},\\ k>l,\end{matrix}$$

а блочные составляющие параметров $\widetilde{\mu}_{N+1}(\cdot)$, $\widetilde{\Gamma}_{N+1}(\cdot)$, $k=\overline{1;N}$, распределения (36) определяются дифференциально-рекуррентными уравнениями типа (11)–(20), (24)–(29) в [1;II].

Сформулированная лемма следует непосредственно из результатов [1;II].

Теорема 4. *Пусть (34) выполнено. Тогда справедливы теоремы 1–3, где $\mu(t|\theta_j)$, $\mu(t-t^*_k|\theta_j)$ определяются соотношениями леммы*

$$\overline{h(t,z|\theta_j)} = h(t,z,\theta_j) + H_{0,N}(t,z,\theta_j)\widetilde{\mu}_{N+1}(t-\widetilde{t}^*_N,t|\theta_j), \tag{35}$$

$$C(\eta(t_m),z|\theta_j) = |V(t_m,z,\theta_j)|^{-1/2}\exp\Big\{-\frac{1}{2}[\eta(t_m)-g(t_m,z,\theta_j)]^T\times$$

$$\times V^{-1}(t_m, z, \theta_j)[\eta(t_m) - g(t_m, z, \theta_j)]\Big\} \frac{|\widetilde{\Gamma}_{N+1}(t_m - \widetilde{t}^*_N, t_m|\theta_j)|^{1/2}}{|\widetilde{\Gamma}_{N+1}(t_m - \widetilde{t}^*_N, t_m - 0|\theta_j)|^{1/2}} \times$$

$$\times \frac{\exp\left\{\frac{1}{2}\widetilde{\mu}^T_{N+1}(t_m - \widetilde{t}^*_N, t_m|\theta_j)\widetilde{\Gamma}^{-1}_{N+1}(t_m - \widetilde{t}^*_N, t_m|\theta_j)\widetilde{\mu}_{N+1}(t_m - \widetilde{t}^*_N, t_m|\theta_j)\right\}}{\exp\left\{\frac{1}{2}\widetilde{\mu}^T_{N+1}(\cdot, t_m - 0|\theta_j)\widetilde{\Gamma}^{-1}_{N+1}(\cdot, t_m - 0|\theta_j)\widetilde{\mu}_{N+1}(\cdot, t_m - 0|\theta_j)\right\}}. \quad (36)$$

Формула (35) следует непосредственно из (14), (30)–(34), а (36) получается аналогично (4.26) в [3].

СПИСОК ЛИТЕРАТУРЫ

1. *Абакумова О. Л., Демин Н. С., Сушко Т. В.* Фильтрация стохастических процессов по совокупности непрерывных и дискретных наблюдений с памятью. I // АиТ. 1995. Т. 9; II // Там же. Т. 10.
2. *Демин Н. С., Рожкова С. В., Рожкова О. В.* Обобщенная скользящая экстраполяция стохастических процессов по совокупности непрерывных и дискретных наблюдений с памятью // Изв. РАН. ТиСУ. 2000. Т. 4.
3. *Dyomin N. S., Rozhkova S. V., Rozhkova O.V.* Likelihood ratio determination for stochastic processes recognition problem with respect to the set continuous and discrete memory observations // Informatica. 2001. Vol. 12. No 2.
4. *Липцер Р. Ш., Ширяев А. Н.* Статистика случайных процессов. М.: Наука. 1974.

Yu.G. Dmitriev, F.P. Tarasenko, Yu.K. Ustinov

On statistical structure of growing systems

A specific peculiarity of growing systems is described. It is that the parts of a system have their masses as equidistanted quantiles of certain probability distribution. Theoretical and experimental considerations are presented and they are in agreement.

1. Introduction

The main purpose of any modeling is prognostification of certain unknown characteristics of the modeled system from available observations of other its characteristics. It is evident that this is possible only when there are connections between observed and desired characteristics, i.e. when ones are containing information about others, when there is a dependence between them.

Our concern is directed to problems of abovementioned type for a special class of systems, namely growing systems. It is along story, which has started from studying interdependencies between volumes of neighboring oil and gas deposits at a certain territory, made by a geologist F.N. Alexeev [1]. He has found empirically that the common genetic origin of such deposits makes them interconnected, and suggested a hypothesis that sizes of resources in deposits of a separate "natural set" (term introduced by Alexeev for a group of neighboring and genetically united deposits), are the order statistics from certain probability distribution: on probabilistic paper they lay very closely to a straight line. Well proved retrospectively on data from fully explored fields in various continents, it opens a possibility to us this knowledge for working with data from poorly, as yet researched areas, trying to estimate how many deposits are not yet discovered, and how big are volumes of resources in them.

After several years of further studies, we came to a conclusion that this idea can be improved and advanced in several directions.

1. Mineral resources are only one example of systems of more general type – growing, proliferating systems. Processes of growth are typical to many physical, biological, economic, social, et. al. phenomena. It is a challengi task to find common regularities between all of them.

2. Alexeev's empirical finding must be backed by rigorous mathematical modeling. Building of such model [2,3] and its further development show that several more exact statements can be made about the nature of growing systems.

3. There are many difficulties and pitfalls in the process of statistical prognosis of unknown characteristic of a system from obtained partial data. This enforces one to introduce into algorithms of estimation many sophisticate techniques for improving quality of decision, i.e. its accuracy and reliability. This moves us far from just technique of Alexeev.

4. It is interesting to see whether theoretical findings are supported by experimental data for systems of different nature.

We five in this paper an outline of some results obtained in these directions

2. Basic Premises of the Theory

There are several background propositions for the theory.

I) Inany growing system one can distinguish its parts: deposits in the oil field; trunk, branches, twigs and leaves in a tree; groups of trees in a forest; cities in a country; bacterial colonies in a bullion; and so on.

II) Our interest will be focused on such characteristics of a system as: the number of its parts; the proportion of the whole mass of the system, belonging to each part; mutual disposition of the parts in space.

III) A formation of parts in a system is a result of interaction between numerous and various external and internal factors; they are so variable in space and time, so complicated,and we are so ignorant of their intrinsic properties, that the only appropriate approach is to consider them as random phenomena.

IV) We will consider only growing systems; this means that each part is growing, gradually enlarging its mass by collecting the substance, by adding it to the previously gained one. Various alternatives of growth are possible (for instance, the increments can be proportional to the mass of the part, or to the surface of it).

V) The next peculiarity of a growing system is that its parts are growing in time, taking the substance from the common source with limited resources and/or limited rate of supply. This makes them interdependent in a sort of competition. We can think of such system as developing gradually, in an evolution, "from the level achieved". And random factors make all parts to be of different sizes, as if there are "priorities" in the system.

3. An Outline of the Probabilistic Theory of Growing Systems

Let us consider the following model of growth of N elements of the system during L time periods:

$$Q_k^{j+1} = Q_k^j + z_k^{j+1} g_k(Q_k^j), j = \overline{1, L}, k = \overline{1, N}, \tag{1}$$

where Q_k^j is a "mass" of k-th element of the system after j periods; z_k^j are independent (re j) random variables distributed uniformly at [0,1]; $g_k(x)$ are some functions chosen in the wake of certain "physical" considerations of the system's nature; and we will assume them following the "priorities condition":

$$0 \leq g_1(x) \leq g_2(x) \leq \ldots \leq g_N(x). \tag{2}$$

The model (1) plus (2) allows to find the probability distribution for Q_k^L; it turns out to be the Captain's distribution [4, p. 121]. Various definitions of

functions $g_k(x)$ lead to the different options of this distribution [2] (among them normal, lognormal, and others).

Equations (1) can be rewritten in the vector form as

$$Q^{j+1} = Q^j + z^{j+1} g(Q^j), \quad j = \overline{1, L},$$

where

$$Q^s = (Q_1^s, \ldots, Q_N^s),\ z^s = (z_1^s, \ldots, z_N^s),\ g(Q^s) = (g_1(Q_1^s), \ldots, g_N(Q_N^s)).$$

$$Mz^s = a = \left(\frac{1}{2}, \ldots, \frac{1}{2}\right) \quad \text{and} \quad \operatorname{cov} z^s = \| \sigma_{il} \|_{i,l=\overline{1,N}} = C,$$

$$z^{j+1} g\left(Q^j\right) = \left(z_1^{j+1} g_1 \leq \left(Q_1^j\right), \ldots, z_N^{j+1} g_N \left(Q_N^j\right)\right).$$

If vectors z^s are i.i.d.r.v's then, according to the multidimensional theorem of Levi–Lindeberg [4, p. 78], we have as $L \to \infty$:

$$\frac{1}{\sqrt{L|C|}} \left(\sum_{j=1}^{L} \frac{Q^{j+1} - Q^j}{g(Q^j)} - aL \right) = \frac{1}{\sqrt{L|C|}} \sum_{j=1}^{L} \left(z^j - a\right) \xrightarrow{d} \mathcal{N}(\vec{0}, \vec{I}),$$

where $\mathcal{N}(\vec{0}, \vec{I})$ is the normal distribution with the null vector of averages and unit matrix covariances, and division of vectors is component-wise.

Denotation

$$G_k(Q_k) = \int_{Q_k^0}^{Q_k^L} \frac{du}{g_k(u)}, \quad G(Q) = (G_1(Q_1), \ldots, G_N(Q_N)),$$

makes it possible to write:

$$S = \frac{G(Q^{L+1}) - aL}{\sqrt{L|C|}} \xrightarrow{d} \mathcal{N}(\vec{0}, \vec{I}),\ L \to \infty. \tag{3}$$

Hence, coordinates of vector S can be considered in asimptotic as indentionship (3) describes probability distribution for normalized "masses" of elements embraced by the system.

Non-normalized "masses" Q^{L+1} are following the Captain's distribution [4]. The transformation $G_k(Q_k^{L+1})$ gives the normal distribution with the unknown (in general case) parameters.

The important finding was that "mass" of system's elements are equally distanced quantiles of the Captain's distribution. Levels of quantiles can be found by the minimization of the distances between e.d.f. F_N of observed "masses" and the Captain's distribution F:

$$d_1 = \sup_x |F_N(x) - F(x)| \quad \text{or} \quad d_2 = \int (F_N(x) - F(x))^2 dF(x).$$

Both lead to quantiles of the levels

$$p_k = (2k-1)/2N, k = \overline{1,N}.$$

This means that discrete distribution of masses $1/N$ condensed at quantiles

$$Q_k = F^{-1}((2k-1)/2N),\ k = \overline{1,N},$$

gives the best approximation to the true distribution governing the evolution of the system.

4. Checking the Theory on Various Systems

An attempt was made to see whether the theory is valid for growing systems of very different origins [5]. The empirical data were collected from sources available about the following systems:

- a pure colony of bacteria
- a plot of natural birch-tree forest
- a single tree
- populations of Armenia, France and England
- populations of various Departments of France
- Solar system. "Masses" of parts of bacterial colony were areas of its spots on the surface of bullion. Data were obtained in the laboratory of Biology Department of TSU.

The Chair of Forestry of TSU kindly gave its results of studying birch-tree forests and separate trees. The observations of the forest were used in two ways. First, a large enough area of the forest was divided in natural clustered groups of trees ("biogroups", in terminology of the Chair). Biomasses of these groups were considered as elements of the system. In the second case a single biogroup was taken as the system, and biomasses of separate trees in it were considered as its elements. When a single tree was considered as the system, biomasses of its branches and twigs, and leaves were taken as its elements.

When population of a country is taken as the system, its parts are either populations of its cities or populations of its historically emerged

The system	The probability	
	W-test	χ^2-test
Colonies of bacteria	0.81	–
Departments of France	0.9975	0.983
Cities of France	0.9998	0.9635
Cities of Armenia	0.977	0.6216
Cities of England	0.999	0.8487
Tree No. 20	0.8175	0.801
Tree No. 42	0.935	0.7546
Tree No. 47	0.965	0.7276
Tree No. 61	0.955	0.9164
Tree No. 65	0.0844	–
Tree No. 67	0.719	0.9387
Trees in a biogroups	0.86	0.9655
Biogroups in a forest	0.999	0.9163
Solar System	0.76	–

provinces. Statistical data, necessary for studying, were taken from relevant handbooks.

For the Solar system its elements were masses of planets.

The results can be presented briefly in the following way.

1. In all cases the best approximation of the distribution of data was the lognormal distribution. It follows from the theory that in these systems the sizes of increments are proportional to the mass before.

2. For checking the reliability of our statistical inferences, W (Shapiro–Wilk) and X^2 (Pearson) tests of goodness-of-fit were applied. Here are their results

From these figures can draw a few conclusions:

a) Usually W-test is more powerful, i.e. it rejects an untrue model with the greater probability than X^2-test. So, we can consider the left column as more reliable.

b) The relatively low indication for Solar system needs futher consideration: there is a hope that including masses of satellites of planets and the mass of the asteroid belt as an additional elements could improve the result.

c) It is very important fact inherent to our calculations that the same regulations are true for a system and for its subsystems (like forest and vs. biogroups, trees in a biogroup, separate trees in the biogroup, each branch in a tree; or a state vs its departments, a state vs its cities; etc.). It looks

like the Nature follows the same laws at all of its levels of organization (here seems to be a nection with the theory of fractals in mathematics or theory of multinodes of S.Beer in management, etc.).

d) It looks like finding a general law for a wide class of growing systems.

e) It is evident that there are many other systems that follow other regulations, and it is very challenging to study them.

REFERENCES (all in Russian)

1. *Alexeev F. N.* Statistical Methods of Prognostification of Mineral Resources. Chita, 1986. 259 p.
2. *Alexeev F. N., Ustinov Yu. K.* Probabilistic Theory of Formation and Prognostification of Mineral Resources, in "Non-Parametric and Robust Statistical Methods in Cybernetics and Informatics" / Ed. by F.P. Tarassenko, Tomsk: TSU-Press, 1990. Pt. I, P. 32-44.
3. *Alexeev F. N., Dmitriev Yu. G., Ustinov Yu. K.* Statistical Structure of Natural Sets of Objects Emerged by accumulance // Applications of statistical Methods in Industry and Managment. Perm, 1990. Vol. 1. P. 1.
4. *Skorokhod A. V., et al.* Handbook on Probability Theory and Math. Statistics. Moskow: Nauka, 1985.
5. *Makleuchuk C. M.* A Class of Growing systems following Lognormal Distribution, A Diploma Work, Dept. of Applied Maths. and Cybernetics, TSU 1998, Supervisiel by F.P. Tarassenko.

В.В. Домбровский, В.А. Гальперин

Управление инвестиционным портфелем в непрерывном времени при квадратической функции риска

Задача управления портфелем ценных бумаг, состоящим из рисковых (обыкновенных акций) и безрисковых (банковского счета, надежных облигаций) вложений, формулируется как динамическая задача слежения за эталонным портфелем, имеющим заданную желаемую эффективность. Предлагается подход к определению оптимальной стратегии управления при квадратичной функции риска.

1. Введение

Проблема выбора инвестиционного портфеля (ИП) является одной из основных в финансовом менеджменте [1]. Классический подход к решению данной проблемы заключается в том, что определяется структура вложений (доли вложений) в рисковые и безрисковые ценные бумаги так, чтобы минимизировать функцию риска (обычно дисперсию портфеля) и получить желаемую эффективность портфеля. При этом задача оптимизации структуры портфеля решается в статической постановке (однопериодные модели) и сводится к решению задач нелинейного или линейного программирования [2]. В динамической постановке определяют стратегию управления, максимизирующую некоторую интегральную функцию полезности, которая имеет достаточно условный характер и обычно подбирается из класса функций, для которых можно получить приемлемое аналитическое решение [3].

В данной работе задача управления ИП формулируется как динамическая задача оптимального слежения за эталонным портфелем, имеющим заданную (желаемую) эффективность.

2. Динамическая модель инвестиционного портфеля

Рассмотрим ИП, состоящий из n видов рисковых вложений и безрискового вклада с неслучайной доходностью. Капитал, помещенный в i-й рисковый актив в момент времени t равен $u_i(t)$, $i = \overline{1,n}$; в безрисковый — $u_0(t)$. Тогда общий объем вложений (портфель) в момент времени t будет равен

$$V(t) = \sum_{i=1}^{n} u_i(t) + u_0(t). \tag{1}$$

Динамика портфеля определяется уравнением

$$dV(t) = \sum_{i=1}^{n} d\nu_i(t)u_i(t) + r(t)u_0(t)dt, \tag{2}$$

где $d\nu_i(t)$ – ставка доходности i-го рискового вложения в момент времени t – случайная величина, $r(t)$ — неслучайная доходность безрисковых вложений. Используя (1), уравнение (2) преобразуем к виду

$$dV(t) = r(t)V(t)dt + b(t)u(t),$$

где $b(t) = [d\nu_1(t) - r(t)dt, d\nu_2(t) - r(t)dt, \ldots, d\nu_n(t) - r(t)dt]$, $u(t) = [u_1(t), u_2(t), \ldots, u_n(t)]$; капитал, вкладываемый в безрисковый актив $u_0(t) = V(t) - \sum\limits_{i=1}^{n} u_i(t)$.

Необходимо определить стратегию управления портфелем путем перераспределения капитала между различными видами инвестиций так, чтобы капитал реального портфеля с минимально возможными отклонениями (с минимально возможным риском) следовал капиталу некоторого определяемого инвестором эталонного портфеля с доходностью $\mu_0(t) > r(t)$, то есть чтобы его капитал рос по закону

$$dV^0(t) = \mu_0(t)V^0(t)dt.$$

В начальный момент времени $V(0) = V^0(0)$.

В качестве меры риска выберем квадратичный функционал

$$J = \mathbf{M}\Big\{\int_0^T \Big(\big[V(t) - V^0(t)\big]^2 + u^{\mathrm{T}}(t)R(t)u(t)\Big)\,dt + \\ +[V(T) - V^0(T)]^2\Big\}, \tag{3}$$

где $R(t) > 0$ – диагональная матрица соответствующей размерности. Второе слагаемое в (3) косвенно учитывает издержки, связанные с управлением портфелем.

Для описания эволюции цен рисковых активов используем модель

$$dS_i(t) = S_i(t)\left[\mu_i(t)dt + \sum_{j=1}^{n}\sigma_{ij}(t)dw_j(t)\right], \tag{4}$$

где $w_j(t)$ – стандартные независимые винеровские процессы, $\sigma_{ij}(t)$ – элементы матрицы волатильности $\sigma(t)$, $\sigma(t)\sigma^{\mathrm{T}}(t) \geq 0$. Данные уравнения представляют собой геометрическое (экономическое) броуновское движение, модели которого обычно используются в финансовой математике [4]. Определим вектор-столбец $x(t) = \big[V(t), V^0(t)\big]^{\mathrm{T}}$, и с учетом (4) представим уравнения динамики реального и эталонного портфелей в виде

$$dx(t) = A(t)x(t)dt + D(t)u(t)dt + \sum_{i=1}^{n} B_i(t)u(t)dw_i(t),$$

где $A(t) = \begin{bmatrix} r(t) & 0 \\ 0 & \mu_0(t) \end{bmatrix}$; $B_i(t) = \begin{bmatrix} \sigma_{1i}(t) & \sigma_{2i}(t) & \dots & \sigma_{ni}(t) \\ 0 & 0 & \dots & 0 \end{bmatrix}$;

$$D(t) = \begin{bmatrix} \mu_1(t) - r(t) & \mu_2(t) - r(t) & \dots & \mu_n(t) - r(t) \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Функционал (3) запишем следующим образом:

$$J = \mathbf{M}\Big\{\int\limits_0^T \left[x^{\mathrm{T}}(t)h^{\mathrm{T}}hx(t) + u^{\mathrm{T}}(t)R(t)u(t)\right]dt + \\ + x^{\mathrm{T}}(T)h^{\mathrm{T}}hx(T)\Big\}, \quad h = [1, -1]. \tag{5}$$

Оптимальный закон управления определим в виде

$$u(t) = K_1(t)V(t) + K_2(t)V^0(t) = K(t)x(t), \tag{6}$$

где $K(t)$ – матрица коэффициентов обратной связи – выбирается из условия минимума функционала (5). Функционал (5) можно записать в виде

$$J = \mathbf{tr}\Big\{\int\limits_0^T \left[h^{\mathrm{T}}hP(t) + R(t)K(t)P(t)K^{\mathrm{T}}(t)\right] + h^{\mathrm{T}}hP(T)\Big\}, \tag{7}$$

где матрица вторых моментов $P(t) = \mathbf{M}\left[x(t)x^{\mathrm{T}}(t)\right]$ с учетом (6) удовлетворяет уравнению

$$\dot{P}(t) = [A(t) + D(t)K(t)]P(t) + P(t)[A(t) + D(t)K(t)]^{\mathrm{T}} + \\ + \sum_{i=1}^{n} B_i(t)K(t)P(t)K^{\mathrm{T}}(t)B_i^{\mathrm{T}}(t). \tag{8}$$

Применяя принцип максимума в матричной формулировке [5], получим решение задачи минимизации функционала (7) на траекториях системы (8):

$$K(t) = -\left[R(t) + \sum_{i=1}^{n} B_i^{\mathrm{T}}(t)Q(t)B_i(t)\right]^{-1} D^{\mathrm{T}}(t)Q(t),$$

$$\dot{Q}(t) = -[A(t) + D(t)K(t)]^{\mathrm{T}}Q(t) - Q(t)\left[A(t) + D(t)K(t)\right] - \\ - \sum_{i=1}^{n} K^{\mathrm{T}}(t)B_i^{\mathrm{T}}(t)Q(t)B_i(t)K(t) - K^{\mathrm{T}}(t)R(t)K(t) - h^{\mathrm{T}}h,$$

$$Q(T) = h^{\mathrm{T}}h.$$

Блоки матрицы $K(k) = [K_1(k), K_2(k)]$ и элементы матрицы $Q(k)$ определяются уравнениями

$$K_1(t) = -\left[R(t) + Q_{11}(t)\sigma(t)\sigma^{\mathrm{T}}(t)\right]^{-1}[\mu(t) - \overline{r}(t)]^{\mathrm{T}}Q_{11}(t);$$

$$K_2(t) = -\left[R(t) + Q_{11}(t)\sigma(t)\sigma^{\mathrm{T}}(t)\right]^{-1}[\mu(t) - \overline{r}(t)]^{\mathrm{T}}Q_{12}(t);$$

$$\dot{Q}_{11}(t) = -K_1^{\mathrm{T}}(t)R(t)K_1(t) - [2r(t) + 2[\mu(t) - \overline{r}(t)]K_1(t) + K_1^{\mathrm{T}}(t)\sigma(t)\sigma^{\mathrm{T}}(t)K_1(t)]\,Q_{11}(t) - 1;$$

$$\dot{Q}_{12}(t) = -K_1^{\mathrm{T}}(t)R(t)K_2(t) - [\mu_0(t) + r(t) + [\mu(t) - \overline{r}(t)]K_1(t)]\,Q_{12}(t) - \left[[\mu(t) - \overline{r}(t)]K_2(t) + K_1^{\mathrm{T}}(t)\sigma(t)\sigma^{\mathrm{T}}(t)K_2(t)\right]Q_{11}(t) + 1;$$

$$Q_{21}(t) = Q_{12}(t),\ Q_{11}(T) = 1,\ Q_{12}(T) = -1,$$

где $\mu(t) = [\mu_1(t), \mu_2(t), ..., \mu_n(t)]^{\mathrm{T}}$, $\overline{r} = [r(t), r(t), ..., r(t)]^{\mathrm{T}}$ – n-мерный вектор-столбец. Заметим, что решение задачи не зависит от Q_{22}.

3. Анализ частного случая

Рассмотрим оптимальное управление при $R(t) \equiv 0$ и критерии вида

$$J = \mathbf{M}\left\{\left[V(T) - V^0(T)\right]^2\right\} = tr\ \left\{h^{\mathrm{T}}hP(T)\right\}. \tag{9}$$

Будем предполагать, что $\sigma(t)\sigma^{\mathrm{T}}(t) > 0$. В данном случае уравнения, определяющие оптимальное управление, запишутся следующим образом:

$$K_1(t) = -\gamma(t)[\mu(t) - \overline{r}(t)]^{\mathrm{T}};\quad K_2(t) = -\gamma(t)[\mu(t) - \overline{r}(t)]^{\mathrm{T}}\frac{Q_{12}(t)}{Q_{11}(t)};$$

$$\dot{Q}_{11}(t) = [\rho(t) - 2r(t)]Q_{11}(t),\ \dot{Q}_{12}(t) = [\rho(t) - \mu_0(t) - r(t)]Q_{12}(t); \tag{10}$$

$$Q_{11}(T) = 1,\ Q_{12}(T) = -1,$$

где $\gamma(t) = \left[\sigma(t)\sigma^{\mathrm{T}}(t)\right]^{-1}$, $\rho(t) = [\mu(t) - \overline{r}(t)]\gamma(t)[\mu(t) - \overline{r}(t)]^{\mathrm{T}}$. Решения дифференциальных уравнений (10) имеют вид

$$Q_{11}(t) = \exp\left[\int_t^T \{2r(\tau) - \rho(\tau)\}d\tau\right];$$

$$Q_{12}(t) = -\exp\left[\int_t^T \{\mu_0(\tau) + r(\tau) - \rho(\tau)\}\,d\tau\right].$$

С учетом этого оптимальные коэффициенты $K_1(t)$ и $K_2(t)$ определяются по формулам

$$K_1(t) = -\gamma(t)[\mu(t) - \overline{r}(t)]^{\mathrm{T}};$$
$$K_2(t) = \gamma(t)[\mu(t) - \overline{r}(t)]^{\mathrm{T}} e^{\int_t^{\mathrm{T}} \{\mu_0(\tau) - r(\tau)\} d\tau}.$$

Уравнения для элементов матрицы вторых моментов (8) при оптимальных $K_1(t)$ и $K_2(t)$ имеют вид

$$\dot{P}_{11}(t) = [2r(t) - \rho(t)]P_{11}(t) + \rho(t)P_{22}(t)e^{\int_t^{\mathrm{T}} 2\{\mu_0(\tau) - r(\tau)\} d\tau};$$
$$\dot{P}_{12}(t) = [r(t) + \mu_0(t) - \rho(t)]P_{12}(t) + \rho(t)P_{22}(t)e^{\int_t^{\mathrm{T}} \{\mu_0(\tau) - r(\tau)\} d\tau};$$
$$\dot{P}_{22}(t) = 2\mu_0(t)P_{22}(t).$$

Так как $P_{22}(t) = \left(V^0(T)\right)^2 e^{-\int_t^{\mathrm{T}} 2\mu_0(\tau) d\tau}$, то

$$\dot{P}_{11}(t) = [2r(t) - \rho(t)]P_{11}(t) + \rho(t)\left(V^0(T)\right)^2 e^{-\int_t^{\mathrm{T}} 2r(\tau) d\tau};$$
$$\dot{P}_{12}(t) = [r(t) + \mu_0(t) - \rho(t)]P_{12}(t) + \rho(t)\left(V^0(T)\right)^2 e^{-\int_t^{\mathrm{T}} \{\mu_0(\tau) + r(\tau)\} d\tau}.$$

Решение данных уравнений в конечный момент времени имеет вид

$$P_{11}(T) = (V^0(0))^2 e^{\int_0^{\mathrm{T}} \{2r(t) - \rho(t)\} dt} + (V^0(T))^2 \left[1 - e^{-\int_0^{\mathrm{T}} \rho(t) dt}\right]; \qquad (11)$$

$$P_{12}(T) = (V^0(0))^2 e^{\int_0^{\mathrm{T}} \{r(t) + \mu_0(t) - \rho(t)\} dt} + (V^0(T))^2 [1 - e^{-\int_0^{\mathrm{T}} \rho(t) dt}]. \qquad (12)$$

Подставляя (11), (12) в критерий (9), запишем выражение для оптимального значения критерия

$$J = e^{-\int_0^{\mathrm{T}} \rho(t) dt} \left[V^0(0) e^{\int_0^{\mathrm{T}} r(t) dt} - V^0(T)\right]^2. \qquad (13)$$

Получим уравнение для дисперсии оптимального портфеля в момент времени T. Рассмотрим уравнение динамики портфеля при оптималь-

ном управлении

$$dV(t) = \left([r(t) - \rho(t)]dt - \sum_{i=1}^{n} \sigma_i(t)\gamma(t)[\mu(t) - \overline{r}(t)]^{\mathrm{T}} dw_i(t)\right) V(t) +$$

$$+ \left(\rho(t)dt + \sum_{i=1}^{n} \sigma_i(t)\gamma(t)[\mu(t) - \overline{r}(t)]^{\mathrm{T}} dw_i(t)\right) e^{\int_0^t \{\mu_0(\tau) - r(\tau)\}d\tau} V^0(t).$$

Очевидно, что

$$\mathbf{M}\{dV(t)\} = [r(t) - \rho(t)]\mathbf{M}\{V(t)\}\, dt + \rho(t) e^{\int_0^t \{\mu_0(\tau) - r(\tau)\}d\tau} V^0(t)dt.$$

Учитывая начальное условие $\mathbf{M}\{V(0)\} = V^0(0)$, получим решение данного уравнения:

$$\mathbf{M}\{V(T)\} = e^{\int_0^t \{r(t) - \rho(t)\}dt} V^0(0) + \left[1 - e^{-\int_0^t \rho(t)dt}\right] V^0(T). \qquad (14)$$

С учетом (11), (12) и (14) дисперсия оптимального управляемого портфеля

$$\mathbf{D}\{V(T)\} = \beta(T) e^{-\int_0^t \rho(t)dt} \left[V^0(0) e^{\int_0^t r(t)dt} - V^0(T)\right]^2, \qquad (15)$$

где $\beta(T) = 1 - e^{-\int_0^t \rho(t)dt}$. Сравнивая уравнение (15) с (13), получаем, что

$$D\{V(T)\} = \beta(T) J(T).$$

В результате в рамках предложенного подхода как частный случай получено решение задачи динамической оптимизации ИП по критерию среднее-вариация (mean-variance portfolio selection problem) [6].

СПИСОК ЛИТЕРАТУРЫ

1. *Шарп У., Александер Г., Бейли Дж.* Инвестиции. М.: ИНФРА-М, 1997.
2. *Лукашин Ю. П.* Оптимизация структуры портфеля ценных бумаг // Экономика и математические методы. 1995. Т. 31. № 1. С. 138-150.
3. *Merton R. C.* Continuous-time finance. Cambr. Ma.: Balckwell, 1990.
4. *Ширяев А. Н.* Вероятностно-статистические модели эволюции финансовых индексов // Обозрение прикладной и промышленной математики. 1995. Т. 2. Вып. 4. С. 527-555.

5. *Athans M.* The matrix minimum principle // Information and Control. 1968. Vol. 11. P. 592-606.
6. *Bajeuex-Besnainou J., Portait R.* Dynamic asset allocation in a mean-variance framework // Management Science. 1998. Vol. 44. No. 11. P. 579-595.

Т.О. Перемитина, Ю.М. Полищук

Комплексный анализ многомерных данных на основе метода главных компонент с применением ГИС

Для анализа многомерных массивов данных широкое применение получил метод главных компонент (МГК), позволяющий сократить размерность исходного признакового пространства путем перехода к новым переменным, называемым главными компонентами [1]. Существуют области научного исследования, в которых изучаемые объекты имеют выраженные пространственные свойства, учет которых позволяет получать в условиях недостаточности количественных данных более полное представление об исследуемых объектах. Однако традиционные процедуры обработки данных методом главных компонент не позволяют учитывать пространственные свойства объектов исследования. В связи с этим целью настоящей работы является изложение методических вопросов проведения комплексного анализа многомерных данных методом главных компонент (МГК) с применением средств геоинформационных систем (ГИС).

Суть предлагаемого подхода, называемого далее геоинформационным подходом, к анализу многомерных данных заключается в сочетании двух методов – пространственного анализа данных с применением средств геоинформатики и статистического анализа данных с применением МГК. Такой комплексный анализ многомерных данных в рамках геоинформационного подхода включает последовательность следующих этапов:

1. Пространственный анализ данных о состоянии объекта исследования средствами ГИС.

2. Обработка массивов данных МГК.

3. Анализ и интерпретация полученных результатов.

Как отмечено выше, применение МГК для анализа многомерных данных позволяет сокращать размерность исходного массива данных. При этом исходное множество данных заменяется факторным множеством с меньшим числом независимых переменных (главных компонент), являющихся линейными комбинациями исходных переменных. В

новой системе координат, определяемой главными компонентами, анализируемые объекты группируются в классы. На основании анализа этих классов можно выявлять закономерности в сходстве или различии группировок объектов. Координаты собственных векторов, определяемые как проекции наблюдаемых переменных, называются нагрузками, которые позволяют определить вклад каждой переменной в картину распределения (группирования) объектов в пространстве главных компонент.

Геоинформационный подход к анализу многомерных данных реализуется в виде программного комплекса, функционирующего в среде ГИС и выполняющего следующие процедуры:

1. Формирование рабочих файлов из базы данных для проведения анализа.

2. Пространственный анализ данных с применением средств ГИС, классификация объектов исследования и формирование массивов данных на основе классификации данных.

3. Обработка массивов данных по алгоритму МГК и выдача результатов обработки в табличном виде.

4. Графическое отображение результатов обработки данных МГК в виде распределения классов в пространстве главных компонент и распределения нагрузок на две первые главные компоненты.

Первый этап предполагает подготовку массива данных для обработки, включая восстановление пропущенных значений, удаление столбцов и др. Пространственный анализ данных на втором этапе предполагает отображение объектов исследования на тематических цифровых картах с применением средств ГИС и последующий анализ их распределения на несколько групп (классов). На третьем этапе проводится традиционный анализ МГК по алгоритму [2]. Как показано в [3], для природных объектов исследования, обладающих выраженными пространственными свойствами, предпочтительно использовать метод Якоби для нахождения собственных значений и векторов корреляционной матрицы. Этот метод не требует приведения исходного массива данных к симметрическому виду, что позволяет уменьшить погрешность и сократить время обработки.

Четвертый этап проведения комплексного анализа позволяет наглядно и оперативно отобразить результаты обработки данных МГК с целью проведения интерпретации результатов. Для лучшей наглядности результатов анализа МГК была разработана дополнительная процедура нанесения на график средних значений и доверительных интервалов (с различными доверительными вероятностями) для каждой груп-

пы исследуемых образцов в пространстве двух главных компонент. Это позволяет показать статистически значимое различие между группами объектов.

В структуру программного комплекса включены следующие компоненты: база данных, общее программное обеспечение, специальное программное обеспечение и СЦК – система цифровых карт. В составе общего программного обеспечения (ОПО) используются средства геоинформационной системы ArcView 3.x, позволяющие проводить пространственный анализ данных и отображать интересующую пользователя информацию или результат ее обработки на цифровых картах. В качестве СУБД (система управления базой данных) используются встроенные средства ГИС, которые поддерживают формат данных, совместимый с dBASE IV. Специальное программное обеспечение (СПО) предназначено для задач формирования рабочих файлов данных, пространственного анализа, обработки многомерных данных МГК, их графического отображения и реализует описанные выше процедуры обработки и анализа. При разработке СПО использовалась среда программирования Delphi 5. Система цифровых карт включает топооснову и тематические слои. В качестве топографической основы используется цифровая карта масштаба 1:1 000 000 (производства Роскартографии), включающая следующие базовые покрытия: административное деление, гидрографию, населчнные пункты и др.

Рассмотрим далее практическое применение предложенного геоинформационного подхода на примере комплексного анализа многомерных данных из базы данных о физико-химических свойствах нефти. Для выполнения пространственного анализа данных была разработана цифровая карта фациального районирования нижней и средней юры Западно-Сибирской нефтегазоносной провинции на основе схемы фациального районирования, приведенной в [4]. В обработку были включены 200 образцов нижне-средне-юрских нефтей Западно-Сибирской нефтегазоносной провинции, каждый из которых описывается следующим перечнем характеристик:

1) температура пласта;
2) давление пласта;
3) содержание газа в нефти;
4) вязкость (при 20 град.);
5) вязкость (при 50 град.);
6) плотность;
7) содержание твердого парафина;
8) содержание общей серы;

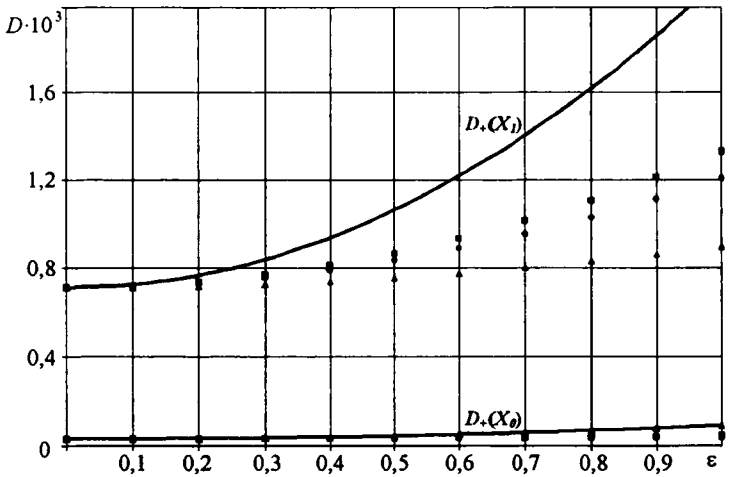


Рис. 1 Результаты обработки данных о физико-химических свойствах нефти

9) содержание общего азота;
10) содержание силикагелеловых смол;
11) содержание асфальтенов;
12) содержание асфальто-смолистых веществ;
13) содержание кокса;
14) содержание бензиновой фракции (мас.%);
15) содержание светлой фракции (мас.%);
16) содержание фракции (об. %);
17) содержание фракции (об. %).

Данные были разбиты на три основных группы по принципу принадлежности анализируемых образцов к одной из фациальных зон: морской, переходной или континентальной. В результате обработки сформированных трех массивов данных получено, что первая главная компонента (ГК1) составляет 75%, вторая ГК2 – 13%, а ГК3 – 5%, ГК4 – 1%. Результаты обработки далее анализировались в пространстве первых двух главных компонент ввиду малых значений компонент ГК3 и ГК4.

На рис. 1,*а* представлено исходное распределение объектов в пространстве двух главных компонент, а на рис. 1,*б* даны те же результаты, представленные в виде средних значений и доверительных интервалов главных компонент для трех групп образцов. Сравнение графиков

показывает, что отображение результатов на правом графике более наглядно и удобно для анализа и интерпретации результатов обработки многомерных данных. Из анализа графиков распределения объектов исследования в пространстве ГК1 и ГК2 (рис. 1,*б*) видно, что различие между тремя группами образцов из разных фациальных зон статистически значимо. Значения нагрузок перечисленных выше характеристик нефти на две главные компоненты сильно различаются для образцов различных фациальных зон. Так как первая главная компонента описывает в среднем 75% исходной информации, далее приведены выводы, полученные на основе рассмотрения нагрузок только на первую главную компоненту. Для нефти из морской фациальной зоны наибольшее значение нагрузки имеет характеристика 16 – содержание светлой фракции, для образцов из континентальной зоны – характеристика 1 – температура пласта, а для образцов переходной зоны первая главная компонента зависит в наибольшей степени от характеристики 3 – содержания газа в нефти. Следовательно, каждая группа образцов имеет свою информационно значимую характеристику, изменение которой и будет наиболее сильно отражаться на распределении образцов в пространстве главных компонент.

Таким образом, совместное применение пространственного анализа и метода главных компонент позволяет выявлять закономерности в описании природных объектов со сложной пространственной структурой на основе обработки многомерных данных, получаемых в результате комплексных исследований таких объектов.

СПИСОК ЛИТЕРАТУРЫ

1. *Леонов В. П.* Обработка экспериментальных данных на программируемых микрокалькуляторах. Томск, 1990. 372 с.
2. *Дубров А. М.* Обработка статистических данных методом главных компонент. М.: Статистика, 1978. 135 с.
3. *Девис Дж.* Статистика и анализ геологических данных. М.: Мир, 1977. 571 с.
4. *Сурков В. С., Серебренникова О. В., Казаков А. М. и др.* Седиментогенез и геохимия нижне-среднеюрских отложений юго-востока Западной Сибири. Новосибирск: Наука, 1999. 213 с.

В. В. Поддубный

Выделение из таблиц сводных данных таблиц сопряженности признаков для их трехмерного графического отображения и обработки в пакете STATISTICA [1]

Рассматривается задача выделения таблиц сопряженности признаков из сводных данных, представленных в виде файлов электронных таблиц в форматах, допускающих импорт в среду пакета STATISTICA. Описываются алгоритмы выделения таблиц сопряженности и их обработки по критерию Пирсона. Приводятся листинги полностью работоспособных программ, реализующих алгоритмы на языке Statistica Basic. Приводится методика работы с этими программами, позволяющая пользователю, даже не владеющему языком программирования Statistica Basic, получать из сводной таблицы данных необходимые ему таблицы сопряженности признаков, обрабатывать их и отображать в виде трехмерных диаграмм графическими средствами пакета STATISTICA.

1. Введение

Во многих случаях результаты первичной статистической обработки данных в медицине, биологии, социологии, экономике и в других областях представляются сводными таблицами вида табл. 1, строки которых соответствуют значениям некоторой упорядоченной переменной **CASE** (например, дате, году, номеру эксперимента и т.п.). Имена столбцов обозначают либо группирующие признаки (такие, как **A** – признаки с числовыми кодами, или **C** – с символьными кодами), либо уровни $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n$ признаков (факторов) **B**, представляющих интерес для исследования. Ячейки столбцов, соответствующие тем или иным уровням признаков, содержат сводные статистические данные (количество событий, число больных, число сторонников чего-либо и т.д.).

При работе с такими таблицами часто возникает необходимость исследовать статистические зависимости или связи признаков типа **B** с группирующими признаками типа **A**, **C** или **CASE** (последний является естественным группирующим признаком). Однако все существующие в настоящее время профессиональные пакеты статистической обработки данных позволяют непосредственно исследовать зависимости признаков типа **B**, заданных столбцами уровней $\{\mathbf{B_j}\}$ таблицы вида 1, только от естественной упорядоченной переменной типа **CASE**, но не от

[1] Работа поддержана грантом РФФИ 98–07–03194.

Таблица 1

Data: TABLE.STA 12v*10c											
Сводные статистические данные											
CASE	A	B1	B2	B3	B4	B5	B6	B7	B8	B9	C
1	1	174	199	69	64	42	34	18	107		C1
2	1	90	5	143	43	43	21	28	25	4	C1
3	2	57	129	62	32	38	3	15	80		C2
4	1	122	151	89	38	33	12	29	59	3	C1
5	3	105	126	22	22	5	6	15	5		C3
6	3	89	102	5	39	22	10	13	33		C3
7	2	3	3	9	17	16	2		12		C2
8	1	39	32	5	2		7	6	2		C1
9	2	23	9	19	11	4	5	2	3		C2
10	3	8	20	17	3	1	3	3	5		C3

A или **C**. Например, нет способов отображать графически непосредственно (т.е. без пересчета, суммирования, сортировки данных) содержимое ячеек как функцию признаков **A** и **B** или **C** и **B** в виде $3D$-диаграммы или проверять гипотезу о независимости этих признаков по критерию χ^2 непосредственно по табл. 1).

В связи с этим в данной работе предлагается инструментарий, который может оказаться полезным для многих пользователей пакета STATISTICA фирмы StatSoft, Inc. Это программы, написанные на языке пакета Statistica Basic, позволяющие любому пользователю, даже не владеющему языком программирования Statistica Basic, провести соответствующую обработку сводных статистических данных, представленных в виде таблицы, имеющей структуру табл. 1, получить таблицы сопряженности выбранных для исследования признаков, обработать их по критерию χ^2 Пирсона и графически представить в виде трехмерных диаграмм графическими средствами пакета STATISTICA.

2. Алгоритм выделения таблиц сопряженности признаков и его программная реализация

Пусть таблица сводных статистических данных, являющаяся исходной для нашего исследования, имеет структуру табл. 1, рассматриваемой в качестве примера. Описание этой структуры приведено выше. Пусть мы хотим установить, имеет ли место значимая статистическая связь между фактором **A**, имеющим числовые уровни **A1**, **A2**, ..., **Ak** (в нашем примере 1, 2, 3), или фактором **C**, имеющим символьные

уровни **C1, C2, ..., Ck** (в нашем примере **C1, C2, C3**), и фактором **B**, уровни которого заданы именами столбцов **B1, B2, ..., Bm** (в нашем примере $m = 9$). Поскольку ячейки столбцов **B1, B2, ..., Bm** содержат количества событий, произошедших на пересечении строк и столбцов уровней факторов **A** (или **C**) и **B** при каждом значении группирующей переменной **CASE**, то такая задача вполне разрешима. Нужно только подсчитать суммарные количества событий, произошедших на пересечении строк и столбцов уровней факторов безотносительно к значениям группирующей переменной типа **CASE**, и получится таблица сопряженности признаков типа табл. 2, которую уже можно обрабатывать методами математической статистики (скажем, по критерию Пирсона).

Таблица 2

Data: NEWTABLE.STA 10v*3c									
Extracted, Summarized and Sorted Data (table.sta)									
A	B1	B2	B3	B4	B5	B6	B7	B8	B9
1	425	387	306	147	117	74	81	193	7
2	83	141	90	60	58	10	16	95	0
3	202	248	44	64	28	19	31	43	0

Однако для рядового пользователя, обычно не владеющего или слабо владеющего встроенными в пакет средствами программирования, и указанное преобразование данных, и использование статистических процедур для обработки преобразовнных данных могут представить значительные трудности (особенно, если табл. 1 имеет большую размерность – сотни или тысячи строк, например). Такую работу проделать вручную достаточно сложно, к тому же при этом трудно избежать ошибок и, следовательно, искажения данных. А профессиональные пакеты статистической обработки данных, в том числе и пакет STATISTICA, не имеют встроенных средств автоматического или меню-управляемого выделения из сводных таблиц типа 1 таблиц сопряженности признаков типа 2, а также (по крайней мере, пакет STATISTICA) встроенных процедур обработки таблиц типа 2 на предмет проверки гипотезы о независимости выделенных признаков.

В приложении приведен полный листинг программы Extract.stb, автоматически выделяющей из таблицы сводных данных типа табл. 1 таблицу сопряженности желаемых признаков типа табл. 2. Программа написана на встроенном в пакет STATISTICA алгоритмическом языке Statistica Basic. Для первого использования программы после запуска

пакета STATISTICA (например, модуля STATISTICA File Server) и открытия файла сводных данных (в нашем примере – файла table.sta) необходимо создать файл программы Extract.stb. Для этого в меню "Analisys" следует выбрать опцию "STATISTICA BASIC" и тем самым открыть окно редактирования новой программы с именем файла программы по умолчанию new.stb. Сохраним эту еще не написанную программу под именем Extract.stb в каталоге, где находится файл данных table.sta (используя меню "File / Save As ..." или клавишу "F12" и диалоговое окно "Save File As"). Перепишем листинг программы Extract.stb, начиная с первой строки, содержащей служебное слово RandomAccess, задающее режим работы программы, до последней строки, соблюдая точно написание всех операторов. После этого сохраним программу, используя пиктограмму с изображением дискеты (или меню "File / Save", или клавиши "Ctrl + S"). Программа готова к использованию. При последующих применениях программы следует открыть сначала файл сводных данных, а затем файл программы, используя опцию меню "File / Open Other / STATISTICA BASIC Program ...".

Запуск программы Extract.stb при активном окне программы осуществляется кнопкой "Execute" на панели инструментов. Если окно программы было закрыто, следует загрузить программу, используя меню "File / Open Other / STATISTICA BASIC Program ..." и диалоговое окно "Open STATISTICA BASIC File", после чего запустить ее кнопкой "Execute".

После запуска программы открывается окно "Select Variables" выбора переменных для исследования. В левой половине окна следует выбрать только одну переменную (фактор) "Grouping Var" типа **A** или **C**. В правой половине окна следует выбрать все переменные (уровни фактора) "Joining Vars (with Integer Values)" типа **B**. После нажатия кнопки "OK" откроется окно "Enter Full Path for Saving Converted Files" выбора пути сохранения нового файла, содержимым которого будет таблица типа 2. В открывшемся окне в строку редактирования "Path:" выводится путь "e : \my\my – sta\table_3d\" (без кавычек), указанный в программе по умолчанию. Можно его оставить или отредактировать.

После нажатия кнопки "OK" на экран выводится электронная таблица (Scrollsheet) "Extracted Data (table.sta)", содержащая данные, извлеченные (скопированные) в соответствии с нашим выбором из исходной таблицы типа 1 (т.е. из файла table.sta в нашем примере). Выводится также окно "Message Box" сообщения, указывающего, какие переменные выбраны в качестве уровней признака (фактора) типа **B**. Нажатием кнопки "OK" закрывается окно сообщения и открывается окно

"Enter New File Name (<=8 symbols)" ввода в строку "New File Name:" имени файла, содержащего пересчитанную и отсортированную таблицу извлеченных данных, приведенную к виду табл. 2. Для нашего примера введем имя файла "newtable" (без кавычек).

После ввода имени файла и нажатия кнопки "OK" на экран выводится итоговая электронная таблица (Scrollsheet) "Extracted, Summarized and Sorted Data (table.sta)" вида табл. 2, содержащая пересчитанные и отсортированные данные, извлеченные из сводной таблицы типа 1 (в нашем примере из файла table.sta) в соответствии с нашим выбором. Эта таблица автоматически сохраняется под указанным нами именем в указанном нами каталоге в формате файла данных пакета STATISTICA (в нашем примере – в файле e : \my\my – sta\table_3d\newtable.sta). Диалоговое окно "Delete Scrollsheets?" позволяет оставить полученные электронные таблицы на экране или закрыть их. После любого выбора появляется окно "Message Box" сообщения, извещающее о завершении работы программы ("Done!!!").

Пересчет и сортировка извлеченных данных производится программой "Extract.stb" по следующему алгоритму. Сначала производится сортировка извлеченных данных в порядке возрастания значения числового кода признака типа **A** или значения числового эквивалента символьного кода признака типа **C**. Затем определяется, сколько разных значений кода имеет переменная типа **A** или **C** и сколько раз в таблице извлеченных данных появляется каждое значение кода. Затем производится суммирование данных в каждом столбце уровня признака типа **B** по строкам, соответствующим одному и тому же значению кода группирующей переменной типа **A** или **C**, то есть суммируется содержимое ячеек, соответствующих одному и тому же значению кода переменной типа **A** или **C** для каждой переменной уровня фактора типа **B**. Если в ячейках встречаются пропущенные значения, они заменяются при суммировании нулями. Результаты суммирования записываются в новый массив данных, выводятся на экран в форме электронной таблицы (Scrollsheet) вида табл. 2 и сохраняются в виде этой же таблицы в файле с расширением .sta в формате данных пакета STATISTICA.

Для наглядного графического представления данных табл. 2 можно теперь воспользоваться графическими возможностями пакета STATISTICA. Опция меню "Graphs / Stats 3D Sequential Graphs / Raw Data Plots ..." открывает окно "3D Raw Data Plots", в котором можно выбрать тип графика (например "Box Plot" в разделе "Graph Type:"). Кнопкой "Variables:" можно открыть окно "Select Variables for 3D Sequential Plot" и выбрать уровни признака **B** для отображения соответствующих им

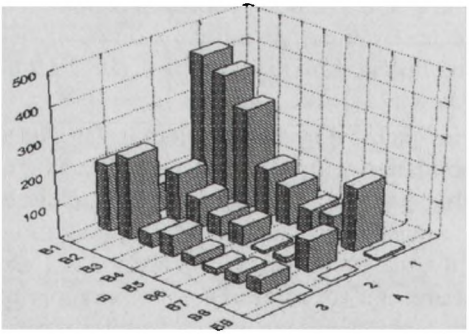


Рис. 1. 3D-диаграмма сопряженности признаков

данных на графике. Кнопкой "Options ..." можно открыть окно "Stats Graphs: Options" и в разделе "DISPLAY / Case Labels" включить радиокнопку "Var", указав в строке ввода переменную "A" (ту, которая имеется в табл. 2) в качестве второго признака. После нажатия кнопки "OK" на экране появляется трехмерный график-диаграмма, по вертикальной оси которого отложено содержимое ячеек столбцов уровней признака **B** табл. 2, а по горизонтальным осям – значения уровней признаков **B** и **A**. Остается ввести названия осей, отредактировать (или удалить) заголовок, цвета и типы раскрасок графика, развернуть график в соответствии с желаемой точкой зрения и т.д. Полученный таким образом 3D-график, отображающий данные таблицы 2, приведен на рис. 1.

3. Алгоритм статистической обработки таблиц сопряженности признаков и его программная реализация

Задача статистической обработки табл. 2 сопряженности признаков сводится обычно к проверке гипотезы H_0 о независимости признаков **A** и **B**. Имеющаяся в пакете STATISTICA процедура построения и обработки таблиц сопряженности признаков может работать только с исходными выборочными данными, но не с таблицами вида 2, а уж тем более не со сводными данными вида табл. 1. Поэтому приведем здесь алгоритм обработки таблицы вида 2 по критерию Пирсона и программную реализацию этого алгоритма.

Как известно [1, с. 112], для проверки гипотезы о независимости признаков по таблице вида 2 можно применить критерий χ^2, основанный

на статистике x^2 К. Пирсона:

$$x^2 = \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{(n_{ij}/n - n_{i.}/n \cdot n_{.j}/n)^2}{n_{i.}/n \cdot n_{.j}/n}, \qquad (1)$$

где n_{ij} – содержимое (i,j)-й ячейки табл. 2, т.е. число случаев, обладающих признаками A_i и B_j, $n_{i.}$ – число случаев, обладающих признаком A_i, т.е. сумма содержимого ячеек i-й строки, $n_{.j}$ – число случаев, обладающих признаком B_j, т.е. сумма содержимого ячеек j-го столбца, n – общее число случаев:

$$n_{.j} = \sum_{i=1}^{k} n_{ij},\ n_{i.} = \sum_{j=1}^{m} n_{ij},\ n = \sum_{i=1}^{k} \sum_{j=1}^{m} n_{ij} = \sum_{i=1}^{k} n_{i.} = \sum_{j=1}^{m} n_{.j}. \qquad (2)$$

Асимптотически (при $n \gg 1$) статистика x^2 при верной нулевой гипотезе имеет распределение χ^2 с $\nu = (k-1)(m-1)$ степенями свободы [1, с. 112]. Гипотеза H_0 отвергается на уровне значимости α, если вычисленное значение статистики x^2 превышает квантиль $\chi^2_{\nu,1-\alpha}$ уровня $1-\alpha$ этого распределения или, что то же самое, если достигнутый уровень значимости критерия p оказывается меньше величины α. В этом случае принимается решение о наличии значимой зависимости между признаками **A** и **B**. В противном случае нет оснований отвергать гипотезу H_0 о независимости признаков.

Таблица 3

2-Way Summary Table: Observed Frequencies										
	B1	B2	B3	B4	B5	B6	B7	B8	B9	**Row** Total
A1	0,14	0,13	0,10	0,05	0,04	0,02	0,03	0,07	0,00	0,58
A2	0,03	0,05	0,03	0,02	0,02	0,00	0,01	0,03	0,00	0,19
A3	0,07	0,08	0,02	0,02	0,01	0,01	0,01	0,01	0,00	0,23
Totals	0,24	0,26	0,15	0,09	0,07	0,03	0,05	0,11	0,00	1,00

Программа Pearson.stb, реализующая этот алгоритм на языке Statistica Basic, приведена в приложении. Файл программы создается или открывается точно по тем же правилам, что и файл программы Extract.stb. Программа работает с таблицей вида 2, т.е. с файлом данных, созданным с помощью программы Extract.stb (в нашем примере – с файлом newtable.sta). После запуска программы Pearson.stb кнопкой "Execute" на экран выводятся электронные таблицы (Scrollsheets) "2-Way Summary

Table: Observed Frequencies" и "2-Way Summary Table: Expected Frequencies", содержащие наблюдаемые $\{n_{ij}/n\}$ и ожидаемые $\{n_{i\cdot}/n \cdot n_{\cdot j}/n\}$ по гипотезе H_0 частоты появления пар уровней признаков $\{A_i, B_j\}$. Это таблицы 3 и 4. В строке "Totals" приведены $\{n_{\cdot j}/n\}$, в столбце "Row Total" — $\{n_{i\cdot}/n\}$, на их пересечении — n/n, т.е. 1.

Таблица 4

2-Way Summary Table: Expected Frequencies										
	B1	B2	B3	B4	B5	B6	B7	B8	B9	**Row** Total
A1	0,15	0,15	0,09	0,05	0,04	0,01	0,03	0,06	0,00	0,58
A2	0,04	0,05	0,03	0,02	0,01	0,01	0,01	0,02	0,00	0,19
A3	0,05	0,06	0,03	0,02	0,02	0,01	0,01	0,03	0,00	0,23
Totals	0,24	0,26	0,15	0,09	0,07	0,03	0,05	0,11	0,00	1,00

Кроме того, на экран выводится таблица "2-Way Summary Table: Pearson Chi-square" (табл. 5), содержащая результаты проверки гипотезы о независимости признаков по критерию χ^2 Пирсона. В этой таблице приводится общий объем выборки n, число степеней свободы "d.f." (то же, что ν), значение критерия "Chi-square" (т.е. значение статистики x^2), коэффициент "fi-coeff." (коэффициент $\varphi = \sqrt{x^2/n}$), коэффициент "contingency coeff." (коэффициент контингенции, сопряженности $c = \sqrt{x^2/(x^2+n)}$) и достигнутый уровень значимости "p-level" (то же, что p, для которого $x^2 = \chi^2_{\nu,1-p}$, т.е. квантиль уровня $1-p$ распределения χ^2 с ν степенями свободы равна значению статистики x^2). Если в нашем примере взять допустимый уровень значимости $\alpha = 0{,}05$, то, как видим, $p \gg \alpha$, так что на 5%-м уровне значимости нет никаких оснований отвергнуть нулевую гипотезу – гипотезу о независимости признаков **A** и **B**.

Таблица 5

2-Way Summary Table: Pearson Chi-square					
n	**d.f.**	**Chi-square**	**fi-coeff.**	**contingency coeff.**	**p-level**
2971	16	0,0595	0,0045	0,0045	0,9999

В заключение подчеркнем, что приведенные ниже листинги соответствуют полностью работоспособным программам. Их достаточно просто скопировать и запустить, если сводная таблица исходных данных имеет структуру табл. 1.

4. Приложение (листинги программ для версии STATISTICA 5.0)

Программа Extract.stb

```
RandomAccess;
NoDataFileVariableNames;
{Путь по умолчанию для сохранения нового файла данных:}
path$ :=' e : \my\my - sta\table_3d\';
ReDim sellist2(NVars);
SelectVariables2('Select Variables',
  1, 1, sellist1, numlist1, 'Grouping Var',
  1, NVars, sellist2, numlist2, 'Joining Vars (with Integer Values)');
NewCases := NCases;
for i := 1 to NCases do begin
  k := 0;
  if not Valid(Data(i, sellist1)) then NewCases := NewCases - 1
  else begin
    for j := 1 to numlist2 do
      if not Valid(Data(i, sellist2(j))) then k := k + 1;
    if k = numlist2 then NewCases := NewCases - 1;
  end;
end;
ReDim Arr1(NewCases, numlist1 + numlist2);
kk := 0;
for i := 1 to NCases do begin
  k := 0;
  if not Valid(Data(i, sellist1)) then kk := kk + 1
  else begin
    for j := 1 to numlist2 do
      if not Valid(Data(i, sellist2(j))) then k := k + 1;
    if k = numlist2 then kk := kk + 1;
  end;
  if i - kk > 0 then begin
    Arr1(i - kk, numlist1) := Round(Data(i, sellist1));
    for j := 1 to numlist2 do
      if Valid(Data(i, sellist2(j))) then
        Arr1(i - kk, numlist1 + j) := Round(Data(i, sellist2(j)))
    else Arr1(i - kk, numlist1 + j) := 0;
  end;
end;
DisplayInputBox('Enter Full Path for Saving Extracted Files',
  'Path :', path$);
VNames$ := VarName(sellist1);
VNames$ := VNames$ +' |' + VarName(sellist2(1));
```

```
if numlist2 > 1 then
  for j := 2 to numlist2 do
    VNames$ := VNames$ +' |' + VarName(sellist2(j));
Scroll1 := NewScrollsheet(NewCases, numlist1 + numlist2,
  Arr1,"Extracted Data",,VNames$);
message$ :=' Selected Vars for Joining : ';
for j := 1 to numlist2 do
  message$ := message$ + VarName(sellist2(j)) +'  ';
DisplayMessageBox(MB_OK,' Message Box',message$);
ReDim Gr(NewCases);
for i := 1 to NewCases do
  Gr(i) := Arr1(i,numlist1);
VectorSort(Gr,Sort_Ascending);
minGr := Gr(1); maxGr := Gr(NewCases);
NCas := maxGr - minGr + 1;
if Trunc(maxGr/10) = 0 then Symbs := 1
  else if Trunc(maxGr/100) = 0 then Symbs := 2
  else if Trunc(maxGr/1000) = 0 then Symbs := 3
  else Symbs := 4;
ReDim Cod(NewCases,NCas);
ReDim SCod(NCas);
NC := NCas;
for k := 1 to NCas do begin
  SCod(k) := 0;
  for i := 1 to NewCases do begin
    if Gr(i) = k + minGr - 1 then Cod(i,k) := 1
      else Cod(i,k) := 0;
      SCod(k) := SCod(k) + Cod(i,k);
  end;
  if SCod(k) = 0 thenNC := NC - 1;
end;
ReDim Codes(NC);
Codes(1) := Gr(1); j := 1;
for i := 2 to NewCases do
  if Gr(i) > Gr(i - 1) then begin
    j := j + 1;
    Codes(j) := Gr(i);
  end;
VNames2$ := Str(minGr,Symbs,-1);
for k := minGr to maxGr - 1 do
  VNames2$ := VNames2$ +' |' + Str(k + 1,Symbs,-1);
ReDim Arr2(NC,numlist1 + numlist2);
for j := 1 to NC do begin
  Arr2(j,numlist1) := Codes(j);
```

```
  for k := numlist1 + 1 to numlist1 + numlist2 do begin
    Arr2(j,k) := 0;
    for i := 1 to NewCases do
      if Arr1(i,numlist1) = Codes(j) then
        Arr2(j,k) := Arr2(j,k) + Arr1(i,k);
  end;
end;
DisplayInputBox('Enter New File Name (<= 8 symbols)',
  'New File Name : ',name$);
VNames1$ := VarName(sellist1);
VNames1$ := VNames1$ +' |' + VarName(sellist2(1));
if numlist2 > 1 then
  for j := 2 to numlist2 do
     VNames1$ := VNames1$ +' |' + VarName(sellist2(j));
Scroll2 := NewScrollsheet(NC,numlist1 + numlist2, Arr2,
  "Extracted, Summarized and Sorted Data",,VNames1$);
SaveScrollsheetAsDataFile(Scroll6,path$ + name$ + ".sta");
button := DisplayButtonBox('Delete Scrollsheets?',
  'Delete All Scrollsheets');
if button = 1 then begin
  DeleteScrollsheet(Scroll1);
  DeleteScrollsheet(Scroll2);
end;
DisplayMessageBox(MB_OK,'Message Box','Done!!!');
```

Программа Pearson.stb

```
RandomAccess;
NoDataFileVariableNames;
ReDim n1(NCases);
ReDim n2(NVars - 1);
for i := 1 to NCases do begin
  n1(i) := 0;
  for j := 2 to NVars do
    n1(i) := n1(i) + Data(i,j);
end;
for j := 1 to NVars - 1 do begin
  n2(j) := 0;
  for i := 1 to NCases do
    n2(j) := n2(j) + Data(i,j + 1);
end;
n := 0;
for i := 1 to NCases do
```

```
  n := n + n1(i);
ReDim ObsFreq(NCases + 1, NVars);
ReDim ExpFreq(NCases + 1, NVars);
for i := 1 to NCases + 1 do
  for j := 1 to NVars do begin
    ObsFreq(i, j) := 0; ExpFreq(i, j) := 0;
  end;
for i := 1 to NCases do begin
  for j := 1 to NVars - 1 do begin
    ObsFreq(i, j) := Data(i, j + 1)/n;
    ExpFreq(i, j) := (n1(i)/n) * (n2(j)/n);
    ObsFreq(NCases + 1, j) := n2(j)/n;
    ExpFreq(NCases + 1, j) := n2(j)/n;
  end;
  ObsFreq(i, NVars) := n1(i)/n;
  ExpFreq(i, NVars) := n1(i)/n;
end;
ObsFreq(NCases + 1, NVars) := 1;
ExpFreq(NCases + 1, NVars) := 1;
VarNames1$ := VarName(1) + "1";
for i := 2 to NCases do begin
  if Trunc(i/10) = 0 then Symbs := 1
  else if Trunc(i/100) = 0 then Symbs := 2
  else if Trunc(i/1000) = 0 then Symbs := 3
  else Symbs := 4;
  VarNames1$ := VarNames1$ + " " + VarName(1) + Str(i, Symbs, -1);
end;
VarNames1$ := VarNames1$ + " " + "Totals";
VarNames2$ := VarName(2);
for j := 3 to NVars do
  VarNames2$ := VarNames2$ + " " + VarName(j);
VarNames2$ := VarNames2$ + " " + "RowTotals";
Scroll1 := NewScrollsheet(NCases + 1, NVars, ObsFreq,
        "2 - Way Summary Table : Observed Frequencies",
        VarNames1$, VarNames2$);
Scroll2 := NewScrollsheet(NCases + 1, NVars, ExpFreq,
        "2 - Way Summary Table : Expected Frequencies",
        VarNames1$, VarNames2$);
x2 := 0;
for i := 1 to NCases do
```

```
  for j := 1 to NVars − 1 do
    x2 := x2 + (ObsFreq(i, j) − ExpFreq(i, j))^2/ExpFreq(i, j);
df := (NCases − 1) * (NVars − 2); p := 1 − ichi2(x2, df);
fi := sqrt(x2/n); c := sqrt(x2/(x2 + n));
Dim Out(6);
Out(1) := n; Out(2) := df; Out(3) := x2;
Out(4) := fi; Out(5) := c; Out(6) := p;
Scroll3 := NewScrollsheet(1, 6, Out,
        "2 − Way Summary Table : Pearson Chi − square",
        "PearsonChi − square",
        "n|d.f.|Chi − square|fi − coeff.|contingancy coeff.|p − level");
DisplayMessageBox(MB_OK,' Message Box',' Done!!!');
```

СПИСОК ЛИТЕРАТУРЫ

1. *Афифи А., Эйзен С.* Статистический анализ: Подход с использованием ЭВМ / Пер. с англ. М.: Мир, 1982. 488 с.

С.П. Сущенко, М.С. Сущенко

О быстродействии многоуровневой памяти неблокирующего типа

Предложена модель влияния параметра глубины неблокируемости кэша на операционные характеристики двухуровневой памяти. Функционирование иерархической памяти описывается работой двухфазного конвейера, на первой фазе которого выполняется обращение к кэш-памяти, а на второй в случае промаха – доступ к оперативной памяти. Работа второй фазы конвейера моделируется марковской системой массового обслуживания с многоэтапным обслуживанием и дискретным временем.

1. Введение

Важнейшим элементом вычислительных систем является подсистема многоуровневой памяти. Разрыв скорости обработки данных в центральном процессоре и скорости доступа к адресуемым операндам в оперативной памяти современных компьютеров составляет 2-3 порядка. Очевидно, что высокое быстродействие вычислительной системы в значительной мере определяется организацией ее многоуровневой памяти.

Главными требованиями, которым должна удовлетворять подсистема памяти, являются достаточно большая емкость, высокое быстродействие и экономическая эффективность с точки зрения технической реализации [1, 2]. Желательно также, чтобы при минимальных физических размерах память обладала как можно большей информационной емкостью. Удовлетворить все эти требования одновременно в одном устройстве невозможно, поэтому обычно комбинируют несколько запоминающих устройств с различными параметрами, добиваясь создания комплексного решения с требуемыми характеристиками. Обычно виртуальная память является средством обеспечения большой информационной емкости, а буферная память (кэш) – средством повышения быстродействия подсистемы памяти.

Как правило, различные уровни иерархической памяти работают асинхронно, поэтому между ними предусмотрены буферные устройства, которые позволяют параллельно выполнять операции доступа к различным уровням. Одним из параметров подсистемы памяти, определяющих ее операционные характеристики, является глубина неблокируемости кэша (емкость межуровнего буферного устройства). Данный параметр задает потенциально достижимую степень параллелизма выполнения совокупности транзакций доступа к различным уровням иерархической памяти.

2. Принципы функционирования и модель кэша неблокирующего типа

Рассмотрим двухуровневую память, имеющую кэш неблокирующего типа с обратной записью [1, 2]. Будем полагать, что процессором порождается неограниченный поток обращений к подсистеме памяти. Процесс функционирования многоуровневой памяти с кэшем неблокирующего типа может быть описан работой двухстадийного конвейера. На первой фазе конвейера выполняется обращение к кэш-памяти. Полагаем, что обращения происходят к различным блокам кэш-памяти. Длительность этой фазы равна времени доступа к кэшу t. При попадании адресуемого объекта в кэш выполняется следующий запрос к иерархической памяти. В случае промаха одновременно происходит обработка текущего запроса на второй фазе – фазе доступа к оперативной памяти и следующей транзакции – на первой фазе. Таким образом, факт обработки транзакции доступа к иерархической памяти на второй фазе является случайным событием. Время обработки транзакции на второй фазе T в $K \geq 2$ раз превышает время обращения к кэш-памяти t. При глубине неблокируемости $N > 1$ подсистема многоуровневой памяти

имеет буфер емкости N для хранения запросов к оперативной памяти, которые последовательно обрабатываются элементами управления памятью на второй фазе конвейера. Если количество таранзакций, обрабатываемых в фазе обращения к оперативной памяти, совпадает с N, то кэш-память оказывается заблокированной. В заблокированном состоянии первая стадия конвейера не работает. Разблокирование первой фазы наступает при завершении обработки одной транзакции на второй фазе конвейера.

Для описания процесса обработки запросов к двухуровневой памяти с глубиной блокировки, равной N, будем моделировать стадию доступа к оперативной памяти марковской системой масового обслуживания (СМО) с дискретным временем и K-этапным обслуживающим прибором [3] с детерминированным временем обслуживания. Длительность цикла дискретной СМО составляет время t. Время между поступлениями заявок (транзакций обращения к оперативной памяти) кратно t и имеет геометрическое распределение с параметром, равным вероятности промаха R.

Считаем, что доступ к многоуровневой памяти выполняется только на чтение или только на запись. Будем полагать, что в каждом такте выполняется обращение к кэш-памяти (при незаблокированной первой фазе конвейера) и с вероятностью промаха R в СМО поступает заявка с временем обработки T, равным трудоемкости выполнения K этапов СМО длительности t. Поскольку обработка различных запросов к оперативной и кэш-памяти выполняется параллельно, то цепь Маркова, описывающая динамику количества этапов обработки совокупности заявок на доступ к оперативной памяти, будет содержать KN состояний. Номер состояния цепи Маркова соответствует количеству этапов обслуживания транзакций доступа к оперативной памяти, накопленных в СМО. Вероятность того, что в стационарном режиме система находится в состоянии k, обозначим через P_k, $k = \overline{0, NK-1}$. Заблокированным соответствует множество состояний $i = \overline{K(N-1)+1, KN-1}$, в которых СМО содержит N транзакций. Переходные вероятности цепи Маркова имеют вид

$$\Pi_{ij} = \begin{cases} R, & i = 0,\ j = K; \\ R, & i = \overline{1, K(N-1)},\ j = i + K - 1; \\ 1 - R, & i = \overline{1, K(N-1)},\ j = i - 1; \\ 1, & i = \overline{K(N-1)+1, KN-1},\ j = i - 1. \end{cases}$$

Система уравнений равновесия для вероятностей состояний P_i, $i = \overline{0, NK-1}$, стационарного режима функционирования дискретной

СМО, описывающей процесс доступа к иерархической памяти, для произвольной конечной глубины неблокируемости $N \geq 2$ и целочисленного отношения времен обращения к оперативной памяти и кэш-памяти $K \geq 2$ записывается следующим образом:

$$\begin{cases} P_0 R = P_1(1-R); \\ P_i = P_{i+1}(1-R),\ i = \overline{1, K-1}; \\ P_K = P_{K+1} + (P_0 + P_1)R,\ N = 2; \\ P_K = P_{K+1}(1-R) + (P_0 + P_1)R,\ N > 2; \\ P_i = P_{i+1}(1-R) + P_{i-K+1}R,\ i = \overline{K+1, (N-1)K-1}; \\ P_i = P_{i+1} + P_{i-K+1}R,\ i = \overline{(N-1)K, NK-2}; \\ P_{NK-1} = P_{(N-1)K+1}R. \end{cases} \tag{1}$$

Важнейшими операционными характеристиками иерархической памяти являются вероятность блокировки кэш-памяти, среднее время доступа и пропускная способность. Вероятность блокировки определяется суммой вероятностей состояний, в которых на фазе доступа к оперативной памяти (в СМО) выполняется обработка N транзакций:

$$Q(N,K) = \sum_{i=(N-1)K+1}^{NK-1} P_i. \tag{2}$$

Среднее время доступа к адресуемым объектам складывается из времени разблокирования кэш-памяти, времени обращения к кэшу и времени обращения к оперативной памяти (при промахе в кэш):

$$\overline{T}(N,K) = t\left\{ \sum_{i=(N-1)K+1}^{NK-1} (i-(N-1)K)P_i + 1 + R\sum_{i=1}^{NK-1} iP_i \right\}. \tag{3}$$

Пропускная способность определяется количеством транзакций доступа к многоуровневой памяти, обработанных за время $\overline{T}(N,K)$. С учетом (3) данное определение можно записать в виде

$$C(N,K) = \frac{1}{Kt} + \frac{K(1-R)-1}{K\overline{T}(N,K)}. \tag{4}$$

3. Вероятностно-временные характеристики

При $N = 2$ и конечном $K \geq 2$ решение системы уравнений равновесия (1) с учетом условия нормировки вероятностей P_i, $i = \overline{0, 2K-1}$, принимает вид

$$P_i = P_0 \frac{R}{(1-R)^i},\quad i = \overline{1, K};$$

$$P_{K+i} = P_0 \frac{R^2}{(1-R)^K} \sum_{j=0}^{K-i-1} (1-R)^j, \quad i = \overline{1, K-1};$$

$$P_0 = \frac{(1-R)^K}{1 + R^2 \sum_{j=0}^{K-2} (K-j-1)(1-R)^j}.$$

Операционные характеристики (2) – (4) при этом определяются зависимостями

$$Q(2,K) = \frac{R^2 \sum_{j=1}^{K-1} (K-j)(1-R)^{j-1}}{1 + R^2 \sum_{j=0}^{K-2} (K-1-j)(1-R)^j}, \quad \overline{T}(2,K) = t +$$

$$+t\left(R^2 \sum_{j=0}^{K-1} \frac{1+K-j}{2} (K-j)(1-R)^j + R^3 \sum_{j=1}^{K-1} \frac{3K+1-j}{2} \times \right.$$

$$\left. \times (K-j)(1-R)^{j-1} \right) \Big/ \left(1 + R^2 \sum_{j=0}^{K-2} (K-1-j)(1-R)^j \right).$$

Для $K = 2$ и конечной глубины неблокируемости $N \geq 2$ решением (1) являются соотношения

$$P_1 = P_0 \frac{R}{1-R}; \; P_i = P_0 \frac{R^{i-1}}{(1-R)^i}, \; i = \overline{2, 2N-2};$$

$$P_{2N-1} = P_0 \left(\frac{R}{1-R} \right)^{2N-2}; \; P_0 = \frac{(1-R)^{2N-2}}{R^{2N-2} + \sum_{j=0}^{N-2} R^{2j} (1-R)^{2(N-j-2)}}.$$

В случае неограниченного N стационарное состояние существует при $R < 1/2$, и решение системы уравнений равновесия принимает вид

$$P_0 = 1 - 2R, \; P_1 = (1-2R) \frac{R}{1-R}, \; P_i = (1-2R) \frac{R^{i-1}}{(1-R)^i}, \; i \geq 2.$$

Операционные показатели для $N = \infty$ описываются соотношениями

$$Q(\infty, 2) = 0, \; \overline{T}(\infty, 2) = t \left(1 + R^2 \frac{3-4R}{1-2R} \right),$$

$$C(\infty,2)=\frac{1}{2t}+\frac{[2(1-R)-1](1-2R)}{2t[1-2R+R^2(3-4R)]}.$$

При $N=3$ и $K\geq 2$ решение (1) имеет вид

$$P_i=P_0\frac{R}{(1-R)^i},\ i=\overline{1,K};$$

$$P_i=P_0\frac{R}{(1-R)^i}[1-(1-R)^{K-1}(1+R(i-K-1))],\ i=\overline{K+1,2K};$$

$$P_{3K-i}=\frac{P_0R^2}{(1-R)^{2K}}\sum_{j=0}^{i-1}(1-R)^j[1-(1-R)^{K-1}(1+R(K-1-j))],$$

$$i=\overline{1,K-1};$$

$$P_0=(1-R)^{2K}\Big/\Big\{(1-R)^K+R\sum_{i=0}^{K-1}(1-R)^{K-1-i}[1-(1-R)^{K-1}\times$$

$$\times(1+iR)]+R^2\sum_{i=1}^{K-1}\frac{K-i}{(1-R)^{1-i}}[1-(1-R)^{K-1}(1+(K-i)R)]\Big\}.$$

Для глубины неблокируемости $N=4$ и произвольного $K\geq 2$ решение (1) записывается следующим образом:

$$P_i=P_0\frac{R}{(1-R)^i},\ i=\overline{1,K};$$

$$P_i=P_0\frac{R}{(1-R)^i}[1-(1-R)^{K-1}(1+R(i-K-1))],\ i=\overline{K+1,2K};$$

$$P_i=P_0\frac{R}{(1-R)^i}\Big\{1-(1-R)^{K-1}[1+R(i-K-1)]+$$

$$+(i-2K)R(1-R)^{2(K-1)}\Big[1+R\frac{i-2K-1}{2}\Big]\Big\},\ i=\overline{2K+1,3K};$$

$$P_{4K-i}=\frac{P_0R^2}{(1-R)^{3K}}\sum_{j=0}^{i-1}(1-R)^j\Big\{1-(1-R)^{K-1}(1+R(2K-j-1))+$$

$$(K-j)R(1-R)^{2(K-1)}\Big[1+R\frac{K-j-1}{2}\Big]\Big\},\ i=\overline{1,K-1};$$

Вероятность P_0 находится из условия нормировки.

Численные исследования операционных характеристик показывают, что при заданной вероятности промаха в кэш-памяти пропущенный по системной шине поток транзакций доступа к адресуемым объектам оперативной памяти растет с увеличением глубины неблокируемости, а среднее время выборки операндов и команд из иерархической памяти падает.

СПИСОК ЛИТЕРАТУРЫ

1. *Кохонен Т.* Ассоциативные запоминающие устройства. М.: Мир, 1982.
2. *Мотоока Т., Томита С., Танака Х. и др.* Компьютеры на СБИС: В 2-х кн. / Пер. с япон. М.: Мир, 1988. Кн. 1. 392 с.
3. *Клейнрок Л.* Теория массового обслуживания. М.: Машиностроение, 1979.

О.В. Шестернева

Об использовании уравнения Винера–Хопфа для идентификации линейных динамических систем в условиях пассивного эксперимента [1]

Работа посвящена вопросам идентификации линейной динамической системы в условиях нормальной эксплуатации объекта.

1. Введение

В работе рассматривается задача идентификации в "широком" смысле в условиях непараметрической неопределенности, то есть в случае, когда структуру объекта с точностью до набора параметров определить невозможно. Актуальность разработки непараметрических методов и алгоритмов идентификации определяется тем фактом, что постановка задач идентификации в "широком" смысле преобладает во множестве практических приложений. Зачастую исследователю приходится сталкиваться с малоизученными процессами и объектами, структура моделей для которых неизвестна. Вследствие этого на современном этапе активно разрабатываются подходы к идентификации динамических систем в условиях неопределенности. Одним из таких подходов является использование непараметрических методов теории идентификации. Непараметрический подход к идентификации ЛДС подразумевает представление линейной системы в виде интеграла Дюамеля с последующим непараметрическим оцениванием весовой функции системы.

[1] Работа выполнена при поддержке Красноярского краевого фонда наук.

В настоящее время непараметрическая идентификация в значительной степени основана на снятии переходных характеристик, т.е. на активном эксперименте. Активные методы идентификации характерны тем, что на вход исследуемого объекта подают стандартные входные воздействия и исследуют выходной сигнал. На практике во многих случаях нарушение нормального функционирования объекта искусственными пробными воздействиями совершенно недопустимо. В этих случаях применяются пассивные методы идентификации, в которых используются случайные естественные колебания входных сигналов.

2. Математическая постановка задачи

Проблема идентификации поясняется на рис. 1.

Здесь объект – это линейный динамический стационарный объект, на который действует $\xi(t)$ – ненаблюдаемое случайное воздействие, $u(t)$ – входные переменные объекта, $x(t)$ – выходные переменные объекта, $u_t^\varepsilon, x_t^\varepsilon$ – соответствующие наблюдения переменных процесса, которые далее будем для простоты обозначать $\{u_i, x_i, i = \overline{1,s}\}$, $\varepsilon(t)$ – случайные помехи в каналах измерения, $\widehat{x}(t)$ – выход модели объекта. Доступная априорная информация содержит только выборки измерений входных и выходных переменных объекта. Данные о структуре объекта отсутствуют, необходимо построить математическую модель стохастического объекта, адекватно описывающую его поведение при произвольном входном воздействии.

Суть метода построения непараметрической модели ЛДС заключается в следующем: известно, что реакция ЛДС $x(t)$ на входное воздействие $u(t)$ описывается интегралом Дюамеля

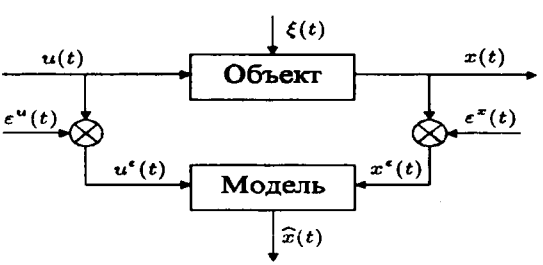


Рис. 1. Общая схема идентификации линейного динамического объекта

$$x(t) = k(0)u(t) + \int_0^t h(t-\tau)u(\tau)d\tau. \qquad (1)$$

Вычисление значения выхода объекта $x(t)$ при этом возможно, если известна его весовая функция $h(t)$. Поэтому основная идея идентификации ЛДС в условиях непараметрической неопределенности состоит в непараметрическом оценивании весовой функции линейной динамической системы. Предлагается применять метод определения импульсной переходной характеристики при произвольном сигнале, используя уравнение Винера–Хопфа [1].

$$K_{ux}(t) = \int_0^t h(t-\tau)K_{uu}(\tau)d\tau. \qquad (2)$$

где $Kuu(\tau)$ и $Kux(t)$ – соответственно автокорреляционная и взаимная корреляционная функции.

Предлагаемый подход к идентификации в условиях нормального функционирования сводится к нахождению весовой функции из уравнения Винера–Хопфа на основе наблюдений "вход-выход" объекта, где в качестве взаимной и автокорреляционной функции используются их оценки. Существует класс случайных процессов, называемых эргодическими и обладающих весьма важным свойством, среднее по множеству от величины, характеризующей процесс, равно среднему по времени с вероятностью, равной единицы. В работе рассматриваются эргодические процессы. Пусть на вход ЛДС поступает произвольный сигнал, представляющий собой смесь полезного сигнала и помехи (помеха аддитивная, статистически независима с нулевым средним и ограниченной дисперсией). Автокорреляционная функция будет иметь вид [2]

$$K_{uu}(\tau) = M\{u(t)u(t+\tau)\} = \int_{\omega(u,u)} u(t)u(t+\tau)p(u_t, u_\tau)\,d\omega =$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_0^T u(t)u(t+\tau)\,dt,$$

где T – интервал времени, на котором производятся измерения.

Взаимная корреляционная функция выглядит следующим образом:

$$K_{ux}(\tau) = M\{x(t)u(t+\tau)\} = \int\limits_{\omega(u,x)} x(t)u(t+\tau)p(u,x)\,d\omega = \\ = \lim_{T\to\infty}\frac{1}{T}\int\limits_0^T u(t+\tau)x(t)\,dt.$$

Оценки автокорреляционной и взаимной корреляционной функции представлены ниже:

$$K_{uu}[\tau] = \frac{1}{s}\sum_{i=0}^{s} u[t_i]u[t_i+\tau]; \tag{3}$$

$$K_{ux}[\tau] = \tfrac{1}{s}\sum_{i=0}^{s} x[t_i]u[t_i+\tau]. \tag{4}$$

Определение импульсной характеристики заключается в численном решении уравнения Винера–Хопфа. Подставляя оценки корреляционных функций (3), (4) в уравнение Винера–Хопфа, получим следующее выражение:

$$\sum_{i=0}^{s} x[t_i]u[t_i+\tau] = \sum_{\nu=0}^{t/\Delta t}\sum_{i=0}^{s} h[t-\nu]u[t_i]u[t_i+\nu]\Delta\nu, \tag{5}$$

где ν – переменная интегрирования.

Численно решая интегральное уравнение Винера–Хопфа (5), генерируем выборку, необходимую для восстановления весовой функции ЛДС:

$$h[t] = \frac{\sum\limits_{i=0}^{s} x[t_i]u[t_i+\tau] - \sum\limits_{\nu=0}^{t/\Delta t}\sum\limits_{i=0}^{s} h[t-\nu]u[t_i]u[t_i+\nu]\Delta\nu}{\sum\limits_{i=0}^{s} u[t_i]u[0]\Delta\nu}. \tag{6}$$

Полученные значения весовой функции можно рассматривать как регрессию, поскольку шаг дискретизации постоянен, оценка весовой функции выглядит следующим образом [3]:

$$h_s[t] = \frac{1}{sC_s}\sum_{i=0}^{s} h[t_i]H\Big(\frac{t-t_i}{C_s}\Big), \tag{7}$$

где H – колоколообразная функция, C_s – параметр размытости, которые удовлетворяют следующим условиям сходимости [3]:

$$\frac{1}{C_s}\int\limits_{\Omega(t)} H\Big(\frac{t-t_i}{C_s}\Big)dt = 1,\ \lim_{s\to\infty}\frac{1}{C_s}\int\limits_{\Omega(t)} H\Big(\frac{t-t_i}{C_s}\Big) = \delta(t-t_i), \tag{8}$$

$$C_s > 0,\ \lim_{s\to\infty} sC_s = \infty,\ \lim_{s\to\infty} C_s = 0. \tag{9}$$

Весовая функция линейной динамической системы является, как известно, ее определяющей характеристикой, то есть, подставив оценку весовой функции (7) в интеграл Дюамеля (1), получим модель ЛДС:

$$x_s(t) = \frac{1}{sC_s}\sum_{\nu=0}^{t/\Delta t}\sum_{i=0}^{s} h[t-\nu]H\Big(\frac{t-t_i}{C_s}\Big)u(\nu)\Delta\nu. \tag{10}$$

Настройка непараметрической модели производится при помощи параметра размытости C_s, который выбирается из минимума квадратичного критерия:

$$W(C_s) = \sum_{i=1}^{s}\left(x(t_i) - x_s(t_i, C_s)\right)^2 \to \min_{C_s}. \tag{11}$$

Данная непараметрическая модель строится при произвольном входном сигнале, структура объекта и дифференциальное уравнение, описывающее исследуемый объект при синтезе модели были неизвестны.

3. Численное исследование

Было проведено численное исследование предлагаемого алгоритма идентификации, построенного в условиях непараметрической неопределенности. В процессе статистического моделирования в качестве объекта использовалась линейная динамическая модель в виде обыкновенного дифференциального уравнения второго порядка. При синтезе алгоритмов построения моделей параметризованная структура и порядок уравнения, описывающего процесс, были неизвестны.

Допустим, что в момент накопления выборки на объект действует сигнал $u(t) = e^{sin(t)+cos(0{,}5t)}$ с аддитивной помехой 5%. Сгенерируем выборку из 200 наблюдений входных-выходных переменных объекта. Получаем оценки корреляционных функций по формулам (3), (4), которые затем подставляем в уравнение Винера–Хопфа (2), находим оценку весовой функции по алгоритму (6) и сглаживаем полученное решение, используя непараметрическую оценку кривой регрессии, которое имеет вид представленный на рис. 2.

Для того чтобы получить модель линейной динамической системы, подставим полученную оценку весовой функции в интеграл Дюамеля.

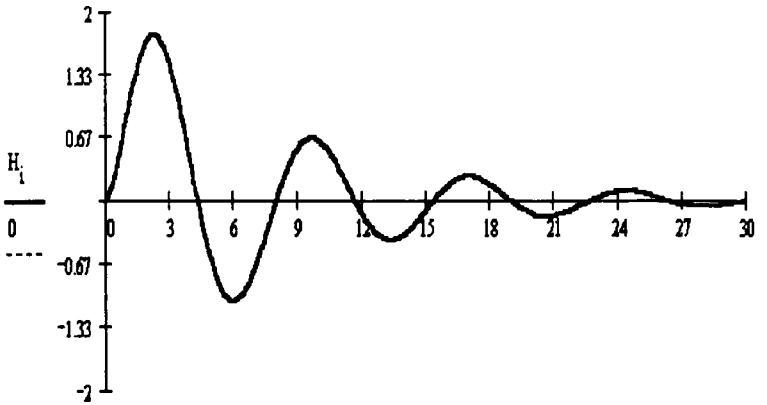


Рис. 2. Оценка весовой функции

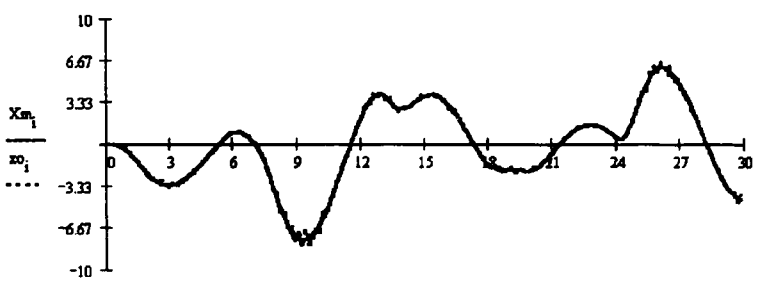


Рис. 3. Реакция объекта X_0 и непараметрической модели X_m на входное воздействие $u(t)$

Проведем статистическое исследование работы полученной модели, подавая в качестве входного воздействия функцию

$$u(t) = \frac{2cos(t)}{3 - 5cos(t)},$$

полученный результат представлен на рис. 3.

Среднеквадратичная ошибка моделирования в этом случае составила 0,03808, что свидетельствует о высоком качестве идентификации объекта.

В заключение отметим, что проведенные численные исследования непараметрической модели и модели, построенной на основе интегрального уравнения Винера–Хопфа, показали достаточно высокое качество идентификации.

СПИСОК ЛИТЕРАТУРЫ

1. *Пугачев В. С.* Теория случайных функций и ее применение к задачам автоматического управления. М.: ГИФМЛ, 1960.
2. *Фельдбаум А. А., Дудыкин А. Д. и др.* Теоретические основы связи и управления. М.: ГИФМЛ, 1963.
3. *Медведев А. В.* Непараметрические системы адаптации. Новосибирск: Наука, 1983.

КОНТАКТНЫЕ ТЕЛЕФОНЫ И E-MAIL АВТОРОВ

Акинфин А.В. БГУ, тел: 375(0)172265704, akinfinav@bsu.by;
Воробейчиков С.Э. ТГУ, тел.: (3822)426099, sev@vmm.tsu.ru
Гальперин В.А. ТГУ, тел.: (3822)426162, gva07@ef.tsu.ru
Демин Н.С. ТГУ, ТНЦ СО РАН ОПИ, dyomns@fpmk.tsu.ru;
Домбровский В.В. ТГУ, тел.: (3822)426162, dombrovs@ef.tsu.ru;
Дмитриев Ю.Г. ТГУ, ТНЦ СО РАН ОПИ, тел.: (3822) 529828, dmit@fpmk.tsu.ru;
Зенкова Ж.Н. ТГУ, ТНЦ СО РАН ОПИ, тел.: (3822)529828, jean@mail2000.ru;
Иванова Е.В. ТГУ, тел.: (3822)426099, vsm@fpmk.tsu.ru;
Кабанова Т.В. ТГУ, тел.: (3822)426099, sev@vmm.tsu.ru;
Кошкин Г.М. ТГУ, ТНЦ СО РАН ОПИ, тел.: (3822)529828, kgm@fpmk.tsu.ru;
Красноштанов А.П. САА, тел.: (3912)640068, krasno-irina@yandex.ru;
Лемешко Б.Ю. НГТУ, тел.: (3832)463754, headrd@fpm.ami.nstu.ru;
Ляшенко Е.А. ТГУ, тел.: (3822)426162, dombrovs@ef.tsu.ru;
Медведев А.Г. БГУ, тел: (37517)2282333, alexey_g_medvedev@hotmail.com;
Медведев Г.А. БГУ, тел.: (37517)2393227, MedvedevGA@bsu.by;
Неделько В.М. ИМ СО РАН, тел.: (3832)333291, nedelko@math.nsc.ru;
Пестунов И.А. ИВМ СО РАН, тел.: (3912)432656, pestunov@icm.krasn.ru;
Поддубный В.В. ТГУ, тел.: (3822)426096, poddubny@inf.tsu.ru;
Помадин С.С. НГТУ, тел.: (3832)463754, headrd@fpm.ami.nstu.ru;
Постовалов С.Н. НГТУ, тел.: (3832)463754, headrd@fpm.ami.nstu.ru;
Рожкова О.В. ТПУ, svr@hm.tpu.ru;
Рожкова С.В. ТПУ, svr@hm.tpu.ru;
Смагин В.И. ТГУ, тел.: (3822)426099, vsm@fpmk.tsu.ru;
Сущенко М.С. ТГУ, тел.: (3822)426099, vas@vmm.tsu.ru;
Сущенко С.П. ТГУ, тел.: (3822)426096, ssp@inf.tsu.ru;
Тарасенко П.Ф. ТГУ, ТНЦ СО РАН ОПИ, тел.: (3822)529828, ptara@ich.tsu.tomsk.su;
Тарасенко Ф.П. ТГУ, ТНЦ СО РАН ОПИ, тел.: (3822) 529828, ftara@ich.tsu.tomsk.su;
Устинов Ю.К. ТГУ, тел.: (3822) 410105, ujk@ic.tsu.ru;

Харин Ю.С. БГУ, тел: 375(0)172265704, kharin@bsu.by;
Шестернева О.В. НИИ СУ ВПТ, тел.: (3912)432863, Kuznetcova_o@mail.ru;
Шуленин В.П. ТГУ, ТНЦ СО РАН ОПИ, тел.: (3822)529828.

Список сокращений организаций

БГУ — Белорусский государственный университет, Республика Беларусь, 220050, Минск, пр. Ф. Скорины, 4.

ИВМ СО РАН — Институт вычислительного моделирования СО РАН, Красноярск, Академгородок;

ИМ СО РАН — Институт математики СО РАН, Россия 630090, Новосибирск, пр. Коптюга, 4;

КГТУ — Красноярский государственный технический университет, Россия, 660074, Красноярск, ул. Киренского, 26;

НГТУ — Новосибирский государственный технический университет, Новосибирск, Россия, г.Новосибирск, пр. К. Маркса, 21;

НИИ СУ ВПТ — НИИ систем управления, волновых процессов и технологий, Россия, 660038, Красноярск, ул. Баумана, 20-в;

САА — Сибирская аэрокосмическая академия им. ак. М. Ф. Решетнева, Красноярск, ул. Красноярский рабочий, 31;

ТГУ — Томский государственный университет, Россия, 634050, г. Томск, пр. Ленина, 36;

ТНЦ СО РАН ОПИ — Томский научный центр Сибирского отделения академии наук, отдел проблем информатизации, Россия, 634021, г. Томск, пр. Академический, 2;

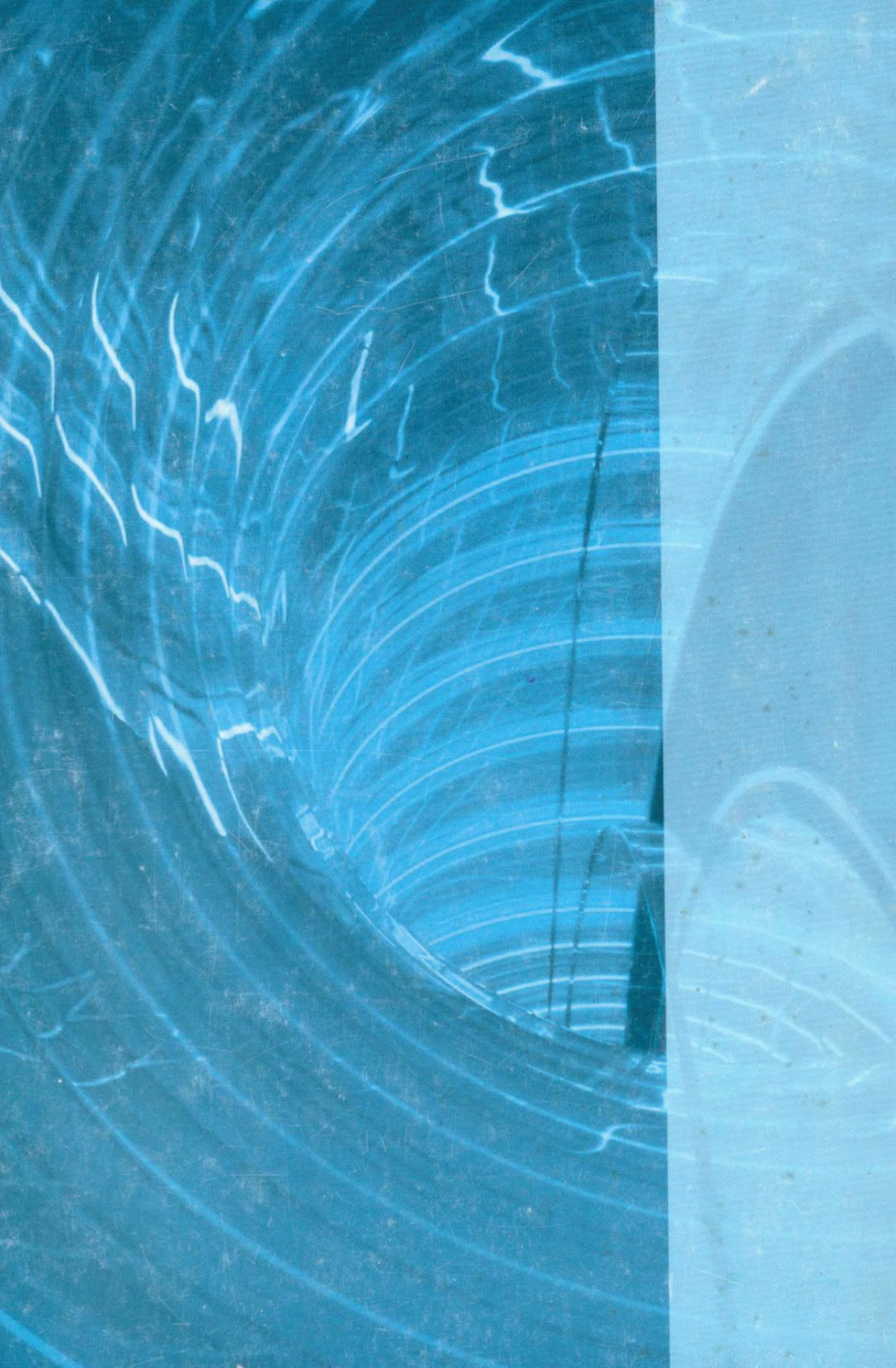
ТПУ — Томский политехнический университет, Россия, 634050, г. Томск, пр. Ленина, 30.

ТРУДЫ
X ЮБИЛЕЙНОГО СИМПОЗИУМА
ПО НЕПАРАМЕТРИЧЕСКИМ И РОБАСТНЫМ СТАТИСТИЧЕСКИМ МЕТОДАМ В КИБЕРНЕТИКЕ

Редактор *Т.С. Портнова*
Верстка *В.Н. Романенко*

К-ОКП ОК-005-93, код продукции 953340

Изд. лиц. ИД № 04000 от 12.02.01. Подписано к печати 12.01.04.
Формат 60×84 1/16. Бумага офсетная. Печать офсетная. Гарнитура «Таймс».
Усл. печ. л. 13,02. Уч.-изд. л. 14,58. Тираж 500 экз.

ЗАО «Издательство научно-технической литературы»
634050, Томск, пр. Ленина, 34а, тел. (382-2) 53-33-35

Отпечатано в типографии ЗАО «М-Принт», г. Томск, ул. Пролетарская, 38/1