Таврический Вестник Информатики и Математики

№1 ' 2009

МЕЖДУНАРОДНОЕ НАУЧНО-ТЕОРЕТИЧЕСКОЕ ИЗДАНИЕ
КРЫМСКИЙ НАУЧНЫЙ ЦЕНТР НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК
И МИНИСТЕРСТВА ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
ТАВРИЧЕСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ им. В.И. ВЕРНАДСКОГО

ОСНОВАН В 2002 ГОДУ

Свідоцтво про державну реєстрацію друкованого засобу масової інформації

КВ №7826 від 04.09.2003

Згідно до постанови ВАК України від 30 червня 2004 р. 3—05/7, перелік №4, журнал "Таврійський вісник информатики та математики" внесено до переліку журналів ВАК України, у яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів кандидата й доктора наук за спеціальностями "Теоретичні основи інформатики та кібернетики", "Математичне моделювання та обчислювальні методи", "Математичне і програмне забезпечення обчислювальних машин і систем", "Системний анализ і теорія оптимальних рішень".

КРЫМСКИЙ НАУЧНЫЙ ЦЕНТР НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК
И МИНИСТЕРСТВА ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
ТАВРИЧЕСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ им. В.И. ВЕРНАДСКОГО

АДРЕС РЕДАКЦИИ:
Крымский научный центр Национальной Академии наук
и Министерства образования и науки Украины
пр-т Вернадского, 2, г. Симферополь, Крым, 95007, Украина

ДЛЯ ПЕРЕПИСКИ:
Факультет математики и информатики ТНУ
пр-т Вернадского, 4, г. Симферополь, Крым, 95007, Украина

Тел. гл.редактора:	(0652) 63-75-42
Тел. редакции:	(0652) 602-466
e-mail (гл.редактор):	donskoy@ccssu.crimea.ua
e-mail (для переписки):	twim_taurida@mail.ru
сайт журнала:	www.twim.crimea.edu

Журнал публикует оригинальные и обзорные статьи
по вопросам теоретической и прикладной информатики и математики

Ведущие тематические разделы:

Функциональный анализ и его приложения

Интегральные, дифференциальные уравнения и динамические системы

Нелинейный анализ и его применение

Спектральные и эволюционные задачи

Математические проблемы гидродинамики

Дискретная оптимизация

Математическая логика, теория алгоритмов и теория сложности вычислений

Математические модели и методы прогнозирования

Машинное обучение и извлечение закономерностей

Дедуктивные системы и базы знаний

Знаниеориентированные и гибридные математические модели принятия решений

Синтез моделей принятия решений при неполной начальной информации

Вычислительная математика

Математическая теория, алгоритмы и системы распознавания образов

Печатается по решению научно технического Совета
КНЦ НАН и Министерства образования и науки Украины
Протокол №1 от 15 января 2009 г.

СОДЕРЖАНИЕ

НАЦІОНАЛЬНОЇ АКАДЕМІЇ НАУК УКРАЇНИ – 90 РОКІВ

НАНУ заснована 27 листопада 1918 року у м. Києві. Її першим президентом був видатний вчений із світовим ім‘ям В.І. Вернадський.

Національна академія наук України згідно з чинним законодавством є вищою державною науковою організацією України, що заснована на державній власності та користується правами самоврядності, яка полягає у самостійному визначенні тематики досліджень, своєї структури, вирішенні науково-організаційних, господарських, кадрових питань, здійсненні міжнародних наукових зв'язків. Академія об‘єднує дійсних членів, членів-кореспондентів та іноземних членів, усіх науковців її установ, організовує і здійснює фундаментальні та прикладні дослідження з найважливіших проблем природничих, технічних та соціогуманітарних наук.

Вищим органом НАН України є Загальні збори її членів. У період між сесіями Загальних зборів керівництво роботою Академії здійснює Президія НАН України, яка обирається Загальними зборами строком на 5 років. До складу Президії НАН України, вибори якої відбулися у квітні 2004 року, входять 32 особи, в тому числі президент, перший віце-президент – головний учений секретар, три віце-президенти, 14 академіків-секретарів відділень, 13 членів Президії. У засіданнях також беруть участь з правом дорадчого голосу 6 в.о. членів Президії та 10 радників Президії НАН України.

В НАН України функціонують 3 секції (фізико-технічних і математичних наук, хімічних і біологічних наук, суспільних і гуманітарних наук), що об'єднують 14 відділень наук: математики, інформатики, механіки, фізики та астрономії, наук про Землю; фізико-технічних проблем матеріалознавства, фізико-технічних проблем енергетики; ядерної фізики та енергетики, хімії, біохімії, фізіології і молекулярної біології, загальної біології, економіки, історії, філософії та права, літератури, мови та мистецтвознавства. В Академії діють 6 регіональних наукових центрів подвійного з Міністерством освіти і науки України підпорядкування: Донецький (м. Донецьк), Західний (м. Львів), Південний (м. Одеса), Північно-східний (м. Харків), Придніпровський (м. Дніпропетровськ), Кримський (м. Сімферополь) та Інноваційний центр по м. Києву.

Основною ланкою структури НАН України є науково-дослідні інститути та прирівняні до них наукові установи. В структурі НАН України діють національні заклади – Національна бібліотека України ім. В. І. Вернадського, Національний науковий центр «Харківський фізико-технічний інститут», Національний історико-археологічний заповідник «Ольвія», Національний ботанічний сад ім. М. М. Гришка, Національний дендрологічний парк «Софіївка», Національний науково-природничий музей. До структури НАН України входять також організації дослідно-виробничої бази (дослідні підприємства, конструкторсько-технологічні організації, обчислювальні центри). Всього на цей час в НАН України діють 176 наукових установ та 49 організацій і підприємств дослідно-виробничої бази.

Між науковими установами Академії та вищими навчальними закладами в 2007 році укладено близько 180 договорів про співробітництво. Щорічно розробляється 200–300 спільних наукових проектів. Функціонує майже 140 спільних з освітянами науково-навчальних структур (комплексів, центрів, лабораторій, філій кафедр тощо), які широко використовують потенціал НАН України для підготовки фахівців високої кваліфікації для потреб вищої школи та НАН України. Протягом останніх років 1600–1800 висококваліфікованих науковців НАН України викладають у вузах, з них кожний десятий академік та член-кореспондент НАН України. У 2007 році майже 80 науково-педагогічних працівників захистили у спеціалізованих радах наукових установ НАН України дисертаційні роботи на здобуття вченого ступеня доктора наук та понад 300 – кандидата наук. Понад 1400 студентів виконували дипломні роботи під керівництвом провідних вчених НАН України; побачили світ понад 90 монографій, підготовлених у творчій співпраці з освітянами.

Науковими установами НАН України щороку впроваджується в різні галузі економіки України понад 2000 новітніх розробок, серед яких передові технології, у тому числі інформаційні, машини, устаткування, матеріали, автоматизовані комплекси і системи, програмні продукти, бази даних і бази знань, сорти рослин, методичні рекомендації та методики, стандарти. В 2007 році значна кількість впроваджених розробок була спрямована на підвищення рівня охорони здоров'я.

Велика увага приділяється забезпеченню більш ефективної діяльності технологічних парків, що створені в Україні, в тому числі за участю провідних інститутів Академії, розгляду та затвердженню напрямів інноваційної діяльності існуючих технопарків, науково-технологічній експертизі їх інноваційних проектів.

В 2007 році укладено 49 ліцензійних угод і контрактів, здобуто 653 патенти на винаходи і корисні моделі.

Академія має власні видавництва «Наукова думка» та Видавничий дім «Академперіодика». Протягом 2007 року установами НАН України видано майже 950 наукових книг, з них 685 – монографії та понад 250 – збірники наукових праць; майже 400 книг навчальної, довідкової та науково-популярної літератури. Поточні праці науковців публікувалися в 81 науковому журналі та понад 50 серійних виданнях. 26 журналів НАН України перекладаються англійською мовою, серед яких 17 перевидаються за кордоном, редакції 9 журналів власними силами перекладають та видають англійською мовою всі свої номери. Практично всі журнали НАН України мають свої сторінки в мережі Інтернет, де розміщують змісти, анотації кількома мовами, близько 40% – повнотекстові версії видань.

Станом на 01.01.08 в НАН України діє близько 100 договорів, проектів, меморандумів про співробітництво понад 150 наукових організацій Академії з науковими організаціями 45 країн світу та міжнародними організаціями, серед яких, зокрема, Міжнародний інститут прикладного системного аналізу IIASA, Європейська організація ядерних досліджень CERN, Європейська наукова асоціація геофізичних досліджень EISCAT. За міжакадемічними угодами виконується близько 100 спільних проектів з організаціями Польщі, Угорщини, Болгарії, Чехії, Словаччини та Румунії. В середньому вчені Академії отримують щорічно близько 600 грантів для виконання наукових проектів, проведення та участі в конференціях, стажування в зарубіжних наукових центрах.

На умовах паритетного фінансування проводяться спільні конкурси науково-дослідних проектів з Сибірським відділенням РАН, Російським фондом гуманітарних досліджень, діє спільна «Програма цільових досліджень та розвиваючих ініціатив» НАН України з Науково-технологічним центром в Україні. Частково фінансуються також спільні проекти Академії з Національним центром наукових досліджень Франції (CNRS) та Радою з науки і технологій Туреччини (TÜBITAK).

Щорічно установами НАН України виконуються роботи в рамках більш як 300 контрактів на загальну суму 50 млн. грн. Триває співпраця з такими корпораціями, компаніями, концернами як "INTEL", "MOTOROLA", "BOEING", "GENERAL ELECTRIC", (США); "Folgat AG" (Німеччина); "Sodern" (Франція); "Sigma Aldrich" (Швейцарія); "Global Metal Technology" (Корея); "NORINKO" (Китай) тощо.

Загальна кількість працюючих в НАН України за станом на 01.01.2008 складала 43349 чол., в тому числі 19024 наукових працівників. Серед них 2568 докторів наук та 8076 кандидатів наук. Середній вік наукових працівників становив близько 50,2 року, докторів наук – 62,2 року, кандидатів наук – 51,4 року.

Персональний склад. За станом на 01.01.2008 до складу НАН України входять 182 дійсних члена (академіка), 343 члена-кореспондента та 115 іноземних членів.

КРЫМСКИЙ НАУЧНЫЙ ЦЕНТР АКАДЕМИИ НАУК УКРАИНЫ: ДЕЯТЕЛЬНОСТЬ В ОБЛАСТИ ИНФОРМАТИКИ И МАТЕМАТИКИ К 90-ЛЕТИЮ НАНУ

Национальная академия наук Украины (НАНУ) была создана 27 ноября 1918 года в г. Киеве. Её первым президентом был видный учёный с мировым именем Владимир Иванович Вернадский.

Сегодня НАНУ – высшая научная организация Украины, которая имеет в своей структуре исследовательские институты и шесть региональных научных центров: Южный (г. Одесса), Западный (г. Львов), Донецкий (г. Донецк), Приднепровский (г. Днепропетровск), Крымский (г. Симферополь) и Северо-восточный (г. Харьков).

Крымский научный цент НАНУ создан 29 января 1997 года, но связи информатиков и математиков Крыма с Академией наук Украины имеют более давнюю историю. Прежде всего, нужно упомянуть совместную работу по становлению и развитию Малой академии наук Крыма «Искатель». Энтузиастом подготовки школьников к будущей научной деятельности был профессор Симферопольского госуниверситета Валентин Николаевич Касаткин, который активно начал работать с юными кибернетиками. Это важное направление педагогической деятельности поддержал в конце 60-х годов вице-президент Академии наук Украины, выдающийся учёный, один из основоположников компьютерной математики и признанный во всём мире лидер в области кибернетики и разработки электронно-вычислительных машин академик Виктор Михайлович Глушков. В одной из статей в «Бюллетене ВАК СССР» он определил трёхступенчатую схему подготовки кибернетиков для академической науки: Малая академия наук – ВУЗ – аспирантура. Виктор Михайлович читал лекции крымским мановцам – юным кибернетикам.

С этого времени началось сотрудничество крымских математиков и кибернетиков с Академией наук, с Институтом кибернетики (ИК) АН УССР (г. Киев). В Симферопольский университет приезжали с лекциями специалисты из ИК АН УССР, крымские учёные защищали диссертации в этом всемирно известном институте.

В 1989 году по инициативе профессора Донского В. И. и чл.-корр. РАН Рудакова К. В. была организована постоянно действующая международная научная конференция «Интеллектуализация обработки информации» (ИОИ). Она была проведена в г. Севастополе под председательством академика АН Украины В. С. Михалевича. Эта конференция стала местом постоянных встреч крымских информатиков и математиков с учёными НАНУ. В разные годы проведение этой конференции возглавляли: академик НАНУ и РАН Ю. И. Журавлёв, чл.-корреспондент НАНУ и РАН А. А. Стогний, академик АН Беларуси В. С. Танаев и другие известные учёные.

Издание трудов ИОИ осуществлялось при поддержке Института искусственного интеллекта НАНУ в г. Донецке и его директора чл.-корр. НАНУ А. И. Шевченко.

До открытия КНЦ в Симферополе существовало Крымское отделение Академии наук Украины, возглавляемое академиком Беляевым В. И. Отделение проводило региональные научные конференции, в которых активно участвовали крымские математики и информатики. Осуществлялось сотрудничество и с Южным научным центром Академии наук, труды крымских учёных издавались в Одессе в научных сборниках ЮЦ АН Украины. Основные исследования были связаны с математическим моделированием рекреационных систем.

В 2000–2008 гг. КНЦ под руководством чл.-корреспондента НАНУ Багрова Н. В. ведёт большую координационную работу и становится действительно центром науки Крыма. В работе КНЦ активно участвуют математики: профессора Донской В. И. и Копачевский Н. Д. В этот период четыре ученика проф. Донского защитили кандидатские диссертации в области математических проблем информатики в Институте кибернетики им. академика В. М. Глушкова. В Институте математики НАНУ защитили докторские диссертации крымские математики И. В. Орлов и М. А. Муратов. Важной вехой сотрудничества стало издание в Институте математики НАНУ монографии М. А. Муратова «Алгебры измеримых и локально измеримых операторов» (2007 г.).

В 2002 году КНЦ НАНУ основывает новый научный журнал «Таврический вестник информатики и математики» (ТВИМ), который выходит в свет дважды в год по настоящее время. Журнал возглавляет главный научный сотрудник КНЦ проф. Донской В. И. В журнале публикуются работы украинских и зарубежных учёных, крымских математиков и информатиков в области математической теории управления, оптимизации при неполной информации, теории машинного обучения и распознавания.

Сегодня, в юбилейные дни Академии, можно подвести очень важный итог. Многолетняя и целенаправленная деятельность в области информатики и математики дала весомые результаты. Команда симферопольских студентов программистов под руководством доцента Козлова А. И. – ученика академика Глушкова – заняла второе место в Европе на олимпиаде по информатике в 2008 году. Журнал ТВИМ получил статус Высшей аттестационной комиссии для публикации материалов кандидатских и докторских исследований. Международная научная конференция ИОИ, проводимая Крымским научным центром, регулярно собирает ведущих специалистов для обмена опытом и продвижения научных исследований в области математики и информатики.

К 90-ЛЕТИЮ СО ДНЯ РОЖДЕНИЯ ПРЕЗИДЕНТА НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК УКРАИНЫ АКАДЕМИКА Б. Е. ПАТОНА

27 ноября 2008 г. исполнилось 90 лет со дня рождения
президента Национальной академии наук Украины академика Б. Е. Патона

Борис Евгеньевич Патон – выдающийся ученый в области сварки, металлургии и технологии металлов, имеющий мировую известность и признание. Он автор многих основополагающих исследований и созданных на их основе высоких технологий, талантливейший организатор науки, видный государственный и общественный деятель. Ему присуща исключительная способность предвидеть перспективные тенденции развития науки, играющие решающую роль на определенном этапе научно-технического прогресса.

Б. Е. Патон творчески развил идеи, продолжил дело, начатое его отцом академиком Е. О. Патоном – крупнейшим ученым, основателем Института электросварки, – и достиг новых вершин в науке.

На протяжении 50 лет он возглавляет всемирно известный академический Институт электросварки им. Е. О. Патона, выросший в мощный научно-технический комплекс, в структуру которого входят научно-исследовательский институт, конструкторско-технологические и экспериментальные подразделения, три опытных завода, а также инновационные организации, научно-инженерные и аттестационный центры. Под руководством Б. Е. Патона и при его непосредственном участии в Институте проведены глубокие исследования и получены крупные результаты в разработке прогрессивных технологий неразъемного соединения и обработки металлов и неметаллов в различных условиях и средах. К ним относятся сварка и наплавка под флюсом, сварка в защитных газах сплошной и порошковой проволокой, электрошлаковая сварка, стыковая сварка оплавлением, газотермическое напыление, лучевые технологии и другие процессы. Фундаментальные исследования Б. Е. Патона и его учеников в области взаимодействия сварочных источников нагрева с расплавленным металлом заложили основу для создания новой отрасли металлургии – специальной электрометаллургии. Благодаря ей стало возможным литье особо чистых специальных сталей и сплавов, цветных металлов, получение уникальных композиционных материалов. Открылись перспективы для создания новейших конструкционных и функциональных материалов XXI века. Борис Евгеньевич внес большой вклад в создание новых типов сварных конструкций, индустриальных способов сварки магистральных трубопроводов, крупногабаритных резервуаров для хранения нефти, кожухов доменных печей, высотных башенных конструкций.

Академик Б. Е. Патон первым начал и развил исследования по использованию сварочных процессов в космической технологии, выполненных космонавтами при орбитальных полетах кораблей и в условиях открытого космоса. Он признанный лидер в этой области.

В последнее десятилетие в круг научных интересов Патона-ученого вошла проблема изыскания оригинальных медицинских технологий и разработка уникальных образцов медицинской техники и инструментов. Под его руководством сотрудники Института и ученые-медики создали новый способ соединения (сварки) мягких тканей человека и животных, широко используемый ныне в хирургической практике.

Свыше 40 лет Б. Е. Патон является президентом Национальной академии наук Украины, которая под его руководством превратилась в один из крупнейших научных центров Восточной Европы, широко известный во всем мире. Она играла и играет важную роль в жизни общества и государства, в прогрессе науки и образования, в укреплении обороноспособности и развитии народного хозяйства Украины.

С самого начала важнейшим направлением организаторской деятельности Патона-президента стало всемерное развитие фундаментальных исследований и создание на их основе новейших технологий для широкого промышленного применения, ориентирование академических институтов на этот путь. Постоянное внимание Борис Евгеньевич уделял комплексности и приоритетности научных изысканий по важнейшим проблемам естественных, технических и гуманитарных наук. Инициатива Б. Е. Патона максимально привлечь научные учреждения к решению производственных и экологических проблем на местах проявилась в организации 7 академических региональных научных центров, охватывающих все области Украины. Его стремление поставить достижения ученых на службу экономике, отраслям промышленности и сельскому хозяйству отразилось в развитии целенаправленных фундаментальных исследований, активном участии академических институтов в научно-технических программах различного уровня.

Успешной реализации научных результатов, их использованию на промышленном уровне способствовали созданная в Академии при руководящей роли Б. Е. Патона опытно-производственная и конструкторская база, инженерные центры, а также образованные затем научно-технические комплексы, в том числе межотраслевые. Новая страница в многогранной деятельности Б. Е. Патона открылась в годы независимости Украины. В качестве члена Совета по вопросам науки и научно-технической политики при Президенте Украины и Совета национальной безопасности и обороны Украины Борис Евгеньевич внес большой личный вклад в адаптацию Национальной академии наук и всей науки Украины к условиям рыночной экономики. Как председатель Комитета по Государственным премиям Украины в области науки и техники он неустанно заботится об авторитете и престиже труда ученых. Большое значение Б.Е.Патон придает инновационной деятельности, формированию и совершенствованию первых в Украине технопарков.

Б. Е. Патон играет важную роль в координации деятельности государственных академий наук в нашей стране, сотрудничестве с вузами, расширении их взаимодействия в интересах развития науки и государства в целом. С большим вдохновением академик Б.Е.Патон заботится о научной молодежи, которой принадлежит будущее, о привлечении молодых талантов к научной работе в институтах и аспирантуре. Он всегда связывает подготовку молодых кадров с пропагандой и улучшением условий сложного, но важного для общества труда ученого. Борис Евгеньевич прилагает большие усилия для сохранения и развития международного научного сотрудничества Академии, внешнеэкономических связей ее институтов с деловыми партнерами зарубежных стран. Это сотрудничество проявляется в широком участии ученых Академии в реализации международных научных программ, организации совместных лабораторий и производств, широком обмене информацией, заключении многочисленных лицензионных соглашений и контрактов. Академик Б. Е. Патон является одним из основателей и на протяжении 10 лет бессменным президентом Международной ассоциации академий наук, объединяющей национальные академии наук, а также ряд ведущих научных центров стран СНГ.

Б. Е. Патон – почетный президент Международной инженерной академии, член Академии Европы, почетный член Римского клуба. Он является действительным членом Российской академии наук, иностранным членом Шведской королевской академии инженерных наук, Национальной академии наук Индии, академий наук и научно-технических обществ многих других стран.

Длительное время Борис Евгеньевич сочетал напряженную научную деятельность с государственной на высоких постах заместителя Председателя Верховного Совета СССР и члена Президиума Верховного Совета Украины. Как человек, ученый и гражданин Б. Е. Патон обладает непревзойденными качествами увлекать во имя высокой цели большие коллективы ученых и организаторов науки, заражать их неиссякаемым энтузиазмом, создавать благоприятную творческую обстановку. Он всегда быстро и своевременно откликается и на нужды своих коллег, и на потребности экономики, государства. Его самоотверженный подвижнический труд отмечен многочисленными научными и государственными наградами и премиями. Он награжден Золотыми медалями им. М. В. Ломоносова и С. И. Вавилова, Золотой медалью им. Л. Лозанна Ассоциации металлургов Италии, Золотой медалью им. В. Г. Шухова Союза инженеров и научно-технических обществ России, Золотой медалью Всемирной организации интеллектуальной собственности, Серебряной медалью им. Эйнштейна ЮНЕСКО. За огромные заслуги Б. Е. Патон удостоен Ленинской и Государственной премий, высоких званий дважды Героя Социалистического Труда и Героя Украины, четырех орденов Ленина, орденов Трудового Красного Знамени, Дружбы народов, князя Ярослава Мудрого V степени.

К 85-ЛЕТИЮ СО ДНЯ РОЖДЕНИЯ АКАДЕМИКА В. М. ГЛУШКОВА

85 лет назад родился пионер кибернетики академик Виктор Михайлович Глушков. *Роль этого выдающегося учёного, значение выдвигаемых им идей не смогли оценить руководители Советского государства. Для информатиков, кибернетиков и математиков всего мира Виктор Михайлович – образец учёного, организатора, безгранично честного и преданного любимой науке человека. Редакция журнала ТВИМ, отмечая юбилей своего кумира, перепечатывает интересный материал Ирины Лисниченко («Факты»).*

Основатель и первый директор Института кибернетики АН УССР академик Виктор Глушков очень любил футбол. И в Институте кибернетики можно было часто наблюдать, как в обеденный перерыв маститые ученые азартно гоняли мяч. По просьбе Валерия Лобановского сотрудники института участвовали в разработке уникальной системы управления тренировками киевских динамовцев. Даже предлагалось укрепить микрофон (подобие нынешнего мобильного телефона) в ухе спортсмена, чтобы по нему сообщать, где какой игрок находится в это время на поле и где находится мяч. Академик был горд и доволен как мальчишка, что может помочь любимой команде. Несмотря на постоянную занятость (кроме должности директора института, у Глушкова было еще 56(!) общественных должностей), Виктор Михайлович побывал на знаменитом матче за Суперкубок УЕФА между киевским «Динамо» и мюнхенской «Баварией», который проходил в 1975 году в Киеве. И вместе со своими дочерьми Олей и Верой радовался победному голу Олега Блохина.

О своем отце корреспонденту «Фактов» рассказала младшая дочь академика Вера Глушкова.

«Еще в 1967-м нашу ЭВМ «Мир» на выставке в Лондоне купила компания IBM»

– Папа постоянно находился под пристальным вниманием Запада, – рассказывает старший научный сотрудник Института кибернетики НАН Украины, кандидат физико-математических наук Вера Глушкова. – Малейший его отрыв вперед, особенно это касалось общегосударственной автоматизированной системы управления экономикой страны, – как тут же «Вашингтон пост» и «Гардиан» выходили с негативными статьями об академике Глушкове. Мол, молодой кибернетик хочет автоматизировать Кремль, заменить людей роботами и так далее.

– В чем была суть этой разработки Глушкова?

Свой проект Виктор Глушков представил руководителю СССР Никите Хрущеву еще в 1964 году. Это ноу-хау уже тогда опережало сегодняшний интернет, завоевавший мир благодаря скорости обмена информацией. По замыслу отца, объединение вычислительных центров и автоматизированных систем управления предприятиями в одной структуре позволяло бы получать четкую картину происходящего в народном хозяйстве и выбирать самый оптимальный вариант управления каждым предприятием, каждой отраслью. Например, ЭВМ могла определить, что в данный момент на заводе в Тольятти следует выпускать столько-то «Жигулей» с такими-то параметрами. И так по каждой продукции, производимой в СССР. Папа еще в конце 50-х годов понял, что без автоматизации управления экономика рассыплется.

В 60-х годах управление Советским Союзом нуждалось в использовании ЭВМ. Потому что в стране производилось свыше 20 тысяч наименований товаров, за выпуском и распределением которых человек уже не мог уследить. Поскольку в СССР действовала централизованная система управления, можно было поставить ЭВМ на все предприятия и из единого центра следить за их работой. Сегодняшняя база данных банков, работа с карточками клиентов – это фрагмент той системы, которую предлагал внедрить отец. Она включала и банки, и бухгалтерский учет, в том числе и безналичную выдачу зарплаты, и производство, и транспорт, и армию... До сих пор ни одна страна в мире не сделала этого в полном объеме!

Большим достижением команды под руководством Виктора Глушкова стало создание электронно-вычислительной машины (ЭВМ) «Мир». В то время машины напоминали огромные шкафы и занимали целые комнаты. Созданная под руководством отца машина «Мир» помещалась на столешнице, в ней использовались основные принципы работы ПК, поэтому ее по праву можно считать прообразом нынешних персональных компьютеров.

Английская компания IBM купила ЭВМ «Мир» на выставке в Лондоне в 1967 году, когда еще не существовало ни единого персонального компьютера. Компания сделала это для того, чтобы в международном суде доказать конкурирующей фирме, что принцип программирования, который запатентовали конкуренты, советские ученые давно изобрели и использовали в машине «Мир». Представители IBM доказали это, о чем имеется соответствующий международный сертификат.

Машина для инженерных расчетов
МИР1 1966 год

Виктор Михайлович Глушков
за пультом ЭВМ «Промінь»

Во время посещения
фирмы IBM

Президент и вице-президент АН Украины:
Б. Е. Патон и В. М. Глушков.
Двадцать лет общего пути.
60-е - 70-е годы XX века

– *Давайте вернемся к идее академика Глушкова автоматизировать всю советскую экономику. Как отнесся Кремль к этому предложению?*

– Проект зарубили. Во-первых, разработчики просили на его внедрение большие деньги, что совершенно объяснимо. Ведь поставить на каждое предприятие дорогие ЭВМ (тогда это были большие ламповые машины) довольно сложно. Во-вторых, с внедрением такой системы сразу стала бы видна ненужность громоздкого управленческого аппарата. То ли дело для управленцев привычно писать наверх отчеты, которые по всей огромной стране не перепроверишь, липовые они или нет. Машинный контроль, который предлагал внедрить Глушков, уменьшал бы значимость управляющей пирамиды. Это они прекрасно понимали...

«О работах молодого кибернетика Глушкова докладывали лично президенту США Джону Кеннеди»

– Теперь остается только гадать, что было бы, если бы проект Глушкова внедрили. Наверное, СССР оказался бы впереди планеты всей.

– Возможно, холодную войну не проиграл бы. И история тогда бы пошла по другому пути. Так что американские специалисты знали цену работам отца и даже лично докладывали о Викторе Глушкове президенту США Джону Кеннеди. Они предлагали папе бросить страну и переехать в США. Сначала пытались купить: американцы готовы были заплатить ему почти миллион долларов в год за чтение двух лекций в неделю. Папа лишь улыбнулся на такое предложение: «Что же вы платите такие большие деньги за такую маленькую работу?» «Не скажите, – ответили ему. – Мы прекрасно понимаем, что вы настоящий ученый. И если к вам будут подходить студенты и задавать вопросы, вы будете на них отвечать. А это для нас самое главное». Но отец отказался.

После этого на него было совершено два покушения. Отец был прагматиком, четко настроенным на дело. Поэтому заподозрить Виктора Глушкова в чрезмерной мнительности невозможно. Папа рассказывал, что впервые на него покушались в Канаде. Самолет, в котором он летел, поднялся, долетел до океана и вдруг повернул обратно.

Папа отлично знал английский, свободно читал лекции и вел переписку. После приземления он услышал, как механики в аэропорту говорили друг другу, что кто-то насыпал песок в двигатель. Слава Богу, летчики обратили внимание на нехарактерный гул мотора и вернулись.

Второй подобный случай произошел в Югославии, когда наперерез автомобилю, в котором ехал академик Глушков, вдруг выехал грузовик. Если бы не профессионализм водителя легковушки, катастрофы было не миновать.

Кроме этих ЧП, возникало и немало других случайностей, но на них папа не акцентировал внимания. Офицеры КГБ, курировавшие работу Института кибернетики, тоже признавали, что происшествия в Канаде и Югославии очень похожи на покушение.

– Негативное отношение власти к предложенному плану автоматизации страны, критика Запада не охладили пыл Виктора Михайловича?

– Неудача никогда не останавливала отца. Во-первых, разработку начало внедрять Министерство обороны СССР, на что дал добро министр обороны Дмитрий Устинов. Во-вторых, автоматизированные системы управления (АСУ) стали появляться на передовых заводах. Одним из первых, кто согласился на установку такой АСУ на своем предприятии, был директор львовского завода телевизоров Степан Петровский.

– Вот почему львовский «Электрон» считался лучшим телевизором!

– Папа обладал невероятным талантом увлекать за собой. Вот и Степана Остаповича уговорил. И продолжал ездить по стране с лекциями, просвещать народ. Иногда за один день мог побывать в трех городах: начало его лекций подгадывали под время прилета самолета. Чтобы послушать академика Глушкова, студенты набивались в аудиторию под завязку, стояли в проходах. Вы до сих пор можете встретить людей, которые после его выступлений круто изменили свою судьбу: переводились на факультеты кибернетики, оставляли привычную работу и шли строить АСУ. Папа умел зажечь людей, показать перспективу!

– Вы тоже закончили факультет кибернетики КГУ?

– Было бы неестественно, если бы я выбрала другой факультет. Этот факультет организовал отец. Сейчас там установлена мемориальная доска. Папа сразу мне сказал: «Ни за кого я просить не буду». К поступлению меня готовил преподаватель по математике.

– Папа утверждал, что сила духа определяет все. Он считал потерянным каждый день, в который не преодолел себя. Зарядку делал всегда до седьмого пота. Если плавал, то плыл несколько километров с такой нагрузкой, что на берегу падал от изнеможения. Несколько раз я вместе с родителями отдыхала в Болгарии. Папа любил заплывать далеко в море, пока однажды его не схватила судорога. После этого он сказал маме: «Я буду плыть вдоль берега, а ты иди и смотри, все ли со мной в порядке». Неудивительно, что, кроме должности директора Института кибернетики АН УССР, у папы было еще... 56 общественных должностей (Вера Викторовна достает из папки четыре страницы машинописного текста с перечнем – вице-президент АН УССР, член ЦК Компартии Украины, депутат Верховного Совета СССР, член Комитета по Ленинским и Государственным премиям, заведующий кафедрой Московского физико-технического института, член редколлегии научно-популярных журналов «Наука и жизнь», «Техника–молодежи» и нескольких международных научных журналов и т. д.). Причем на этих должностях папа не просто числился! Если уж он начинал чем-то заниматься, то достигал в новом деле успехов. В Институте кибернетики занимались вычислительной техникой, начиная от железа и заканчивая химией и биологией. В этом же институте работал академик Николай Амосов. В структуре института был даже отдел прогнозирования политики...

«Цитируя Маркса и Энгельса, папа спас своего учителя – известного ученого, которого объявили врагом народа»

– Отец обожал читать – продолжает моя собеседница. – Вечерами мы с ним вместе читали журналы «Техника–молодежи», «Наука и жизнь», разгадывали кроссворды. Это было его любимым занятием. В журнале «Техника–молодежи» папа был не только членом редколлегии, но и членом комиссии по изучению НЛО. Как он рассказывал, большинство случаев подтвердилось и все они имели под собой реальную основу. Допустим, возникшее зарево над Петрозаводском объяснялось взрывом на военных учениях, проводившихся в области...

– *В мистику известный кибернетик верил?*

– В мистику – нет, а в Бога верил. Когда мы с отцом ездили в Москву, посещали Троице-Сергиеву лавру, на территории которой похоронен преподобный Сергий Радонежский. Хорошо запомнила ту поездку. А вот куличи у нас дома не пекли, пасхальные яйца не красили, в церковь мы не ходили. В советские времена это не приветствовалось. Мы же были под наблюдением. Лишнего слова не могли сказать по телефону!

Расскажу о поступке отца, который я считаю одним из главных в его жизни. Об этом еще никто не писал. После окончания Новочеркасского индустриального института папа вместе с мамой, Валентиной Михайловной, по распределению приехал в Свердловск и стал работать в Лесотехническом институте. Руководителем его диссертации был профессор Сергей Николаевич Черников, автор известных учебников по алгебре и высшей математике. И вот однажды профессора объявили врагом народа. Созвали собрание, на котором все должны были клеймить его позором. Не прийти на такое собрание или промолчать – уже это было гражданским подвигом!

– Папа собирался на собрание, – рассказывает Вера Глушкова. – Мать плакала и умоляла: «Витя, не ходи!» Он посмотрел на жену и сказал: «Ты знаешь, если не пойду, я с этим жить не смогу». Встал и пошел. У отца была блестящая память, он знал наизусть Маркса, Ленина, Энгельса, Гегеля, Фейербаха. Когда заканчивал институт, ему даже предлагали работу на кафедре марксизма-ленинизма. На собрании в Свердловске отец встал и начал сыпать цитатами: «По этому вопросу Маркс, Энгельс сказали то-то. Этот классик говорил так, другой писал следующее»... Словом, засыпал президиум цитатами, с помощью которых объяснил, что в случае с профессором Черниковым они поступают неправильно. И профессора отпустили. Даже с работы не выгнали!

«18-летний Виктор Глушков трижды просился на фронт, но его не взяли из-за плохого зрения – минус 12»

– У каждого человека есть своя миссия, – продолжает Вера Глушкова. – Папе Бог дал таланты и знания, к тому же он сумел развить их до абсолюта. В детстве усовершенствовал систему скорочтения, оттачивал ее в вузе и в результате мог запоминать сразу до 20 страниц математического текста. Отец был невероятно любознательным и интересовался всеми науками.

В детстве с ним много занимался отец, мой дед. Дедушка родился в семье потомственных донских казаков в станице Луганская, что на границе Луганской и Ростовской областей, еще до революции получил высшее образование и работал главным инженером в городе Шахты Ростовской области. Здесь и родился его единственный сын Виктор. Дед и мой маленький папа разработали электронную пушку, стрелявшую на заданное расстояние, собрали приемник-телевизор. Так что папа многое умел делать своими руками, с техникой он был на ты.

Будучи школьником, Витя Глушков прошел со своим отцом всю вузовскую программу. Поэтому, когда поступил в Новочеркасский индустриальный институт, там ему особо делать было нечего. Экстерном закончил папа и Ростовский университет.

– *Почему не престижный МГУ?*

– Папа подавал документы в Москву, но из-за того, что в годы войны он находился на оккупированной территории, получил отказ. Только в 70-е годы стала доступной информация, что наши войска на южном направлении по линии Ростов-Таганрог попали в котел (этот факт был отражен в фильме «Они сражались за Родину»), так что все пути бегства мирных жителей оказались отрезанными. Виктор Глушков вместе с матерью оказался в оккупации.

В начале войны 18-летний Виктор Глушков трижды ходил в военкомат и просился на фронт, но ему отвечали: «Без вас обойдемся!» Ведь у папы было очень плохое зрение – минус 12. Без очков он вообще ничего не видел! Поэтому папа остался в Шахтах. Его мать Вера Львовна, в честь которой меня и назвали, работала в подполье, распространяла листовки, написанные по сводкам Информбюро, которые она слушала по собранному сыном радиоприемнику. Однажды по дороге домой моего отца остановили соседи и предупредили, что его мать арестовали гестаповцы и расстреляли. Почти полтора года папа жил вместе со своим другом в разрушенной школе, питаясь дохлой кониной, и почти каждый день ходил по пересыльным лагерям в надежде увидеть маму... Сейчас в Шахтах стоит памятник погибшим в Великой Отечественной войне, среди которых есть и фамилия моей бабушки.

Отец хорошо помнил свою родословную, утерянную в годы войны, и очень гордился, что он из донских казаков. К сожалению, восстановить ее не удалось, так как она была в единственном рукописном варианте. Один из предков Глушковых был адъютантом атамана Платова и брал Париж в 1812 году. Второй в 1905 году демонстрировал на выставке в Париже собственноручно выращенную то ли тыкву, то ли дыню. Мой дед был первым из всей династии Глушковых, который получил высшее образование, закончил политехнический институт в Екатеринославе (сейчас Днепропетровск), в годы гражданской войны как незаинтересованная сторона присутствовал на переговорах представителей Красной Армии с батькой Махно. Дед рассказывал, как Махно проводил собрания. Батько ставил на голосование какой-то вопрос и, направляя дуло нагана на людей, говорил: «Кто против? Единогласно».

«Отец знал, что умирает. Сам поставил себе диагноз»

– *Как обычно отдыхал ваш отец?*

– У папы в домашнем кабинете до трех часов ночи горел свет – он работал. Единственный месяц, когда с папой можно было спокойно поговорить, был наш отдых в Болгарии. Сколько себя помню, он или работал, или шутил. Папа знал массу анекдотов, забавных случаев. Любил Высоцкого и последние годы жизни часто слушал его песни. Все юмористические стихи барда папа знал наизусть и меня научил. Он прекрасно пел, в каком-то конкурсе перепел знаменитого исполнителя. Даже Анатолий Соловьяненко, отдавая должное пению отца, оставил ему свой автограф. Особенно папа любил «Дивлюсь я на небо», «Два кольори», «Чорнобривці».

– *Когда вы узнали о болезни Виктора Михайловича?*

– Последний год папа очень тяжело ходил. Сила воли у него была невероятная. Сейчас я понимаю, насколько ему было плохо и что он тогда преодолевал, продолжая жить в привычном для себя ритме. Преодолевал болезнь и никогда никому не жаловался. Лишь когда папе сделали пункцию спинного мозга, признался маме: «Валя, мне было так плохо, что я чуть не застонал».

Помню, когда я приехала в больницу в Москву (после больницы в Феофании Виктор Михайлович два месяца пролежал в «Кремлевке», ему долго не могли поставить диагноз. – Авт.), то я не смогла находиться в палате и, чтобы не расплакаться, вышла. Тогда многие плакали. Но папа своим духом заражал всех, смеялся, шутил и поддерживал здоровых. Хотя сам был в трубках, проводах и аппаратах, поддерживающих в нем жизнь. Проведать папу пришел и министр обороны СССР Дмитрий Устинов. Дмитрий Федорович поинтересовался, чем может помочь. Папа пошутил: «Пришлите танк!»

Папа знал, что умирает... Он сам поставил себе диагноз. Обложился специальной литературой и понял, что у него опухоль продолговатого мозга. Когда конец стал для него очевиден, посадил рядом с собой мою старшую сестру Олю и надиктовал ей свою исповедь.

У нас, по-моему, было всего два томографа, в том числе в 4-м Управлении Минздрава УССР, где не смогли расшифровать томограмму Глушкова. Пригласили специалиста из Германии, который посмотрел томограмму и подтвердил диагноз, поставленный отцом. Продолговатый мозг отвечает за рефлексы и оперировать его очень опасно...

30 января 1982 года папа умер. Ему было 58 лет. Похоронили отца на центральной аллее Байкового кладбища в Киеве.

УДК 519.962.22

МІНІМАКСНІ СЕРЕДНЬОКВАДРАТИЧНІ ОЦІНКИ ТРЕНДУ В ЗАДАЧАХ РЕГРЕСІЇ

© Демиденко С.В., Наконечний О.Г.

Київський національний університет ім. Т.Г.Шевченка,
факультет кібернетики
пр-т академіка Глушкова-2, корпус 6, м. Київ 03680, Україна
e-mail: *s.demidenko@gmail.com*

Abstract. This paper describes an approach to the linear minimax estimation of the generalized polinom with unknown partially unbounded parameters. Linear minimax estimations are constructed on the basis of measurements with random noise. We introduce notions of the upper and lower linear minimax estimations. The sufficient conditions for the minimax estimation existence are formulated. Numerical simulation that illustrates these results is presented.

Вступ

Проблеми апроксимації функції тісно зв'язані із задачами оптимізації [1, 2] з одного боку, а також із питаннями оцінки функцій за даними спостережень, що досліджувалися в математичній статистиці [3, 4].

У даній роботі розглядається випадок, коли спостерігається узагальнений поліном із шумом. При цьому кореляційна функція випадкового процесу що моделює шум невідома і належить певній області. При обмеженнях на коефіцієнти розроблено алгоритм побудови гарантованої лінійної оцінки такого полінома.

1. Постановка задачі

Нехай на відрізку $[0, T]$ спостерігається реалізація випадкового процесу $y(t)$, що має вигляд

$$y(t) = \sum_{i=1}^{m} \alpha_i \varphi_i(t) + \xi(t) \tag{1}$$

де $\varphi_i(t)$ $i = \overline{1..m}$ – неперервні на відрізку $[0, T]$ функції, $\xi(t)$ – реалізація неперервного в середньому квадратичному випадкового процесу з нульовим середнім та невідомою кореляційною функцією $R(t, s) \in K$

$$K = \{R : \int_0^T \int_0^T (R(t, s) - R_0(t, s))^2 dtds \leq q^2\} \tag{2}$$

$R_0(t, s)$ – відома кореляційна функція неперервна на $[0, T] \times [0, T]$. Нехай також відомо, що коефіцієнти $\alpha_i, i = \overline{1..r}, r \leq m$ належать множині G, де

$$G = \{\alpha : |\alpha_i - \widehat{\alpha}_i| \leq \beta_i, i = \overline{1..r}\} \tag{3}$$

Позначимо через $\widehat{P(T)}$ лінійну оцінку тренду $P(T) = \sum\limits_{i=1}^{m} \alpha_i \varphi_i(T)$ вигляду $\widehat{P(T)} = \int\limits_0^T u(t)y(t)dt + c$, де $u(t)$ належить простору вимірних за Лебегом інтегрованих з квадратом функцій, c – довільна константа. Позначимо через $\sigma(R, \alpha, u, c)$ середньоквадратичну похибку такої оцінки, тобто

$$\sigma^2(R, \alpha, u, c) = M(P(T) - \widehat{P(T)})^2 \tag{4}$$

Означення 1. Оцінка для якої $\widehat{u}$ та $\widehat{c}$ належать множині $(\widehat{u}, \widehat{c}) \in \arg\min \sigma_1^2(u, c)$, де

$$\sigma_1^2(u, c) = \sup_{G,K} \sigma^2(R, \alpha, u, c)$$

називається мінімаксною середньоквадратичною оцінкою.

2. Умови існування та вигляд мінімаксної оцінки

Твердження 1. Припустимо, що $\varphi_i(t)$, $i = \overline{(r+1)..m}$, лінійно незалежні, тоді існує єдина мінімаксна середньоквадратична оцінка причому

$$\inf_{u,c} \sigma_1^2(u, c) = \inf_{u \in U} \sigma_1^2(u, \widehat{c}) = \sigma_2^2(\widehat{u}) \tag{5}$$

де $\sigma_2^2(u) = \sigma_1^2(u, \sum_{i=1}^{r} \widehat{\alpha}_i z_i(0))$, $\widehat{u} = \arg\min_{\widehat{u} \in U} \sigma_2^2(u)$, U множина, що визначається із умови

$$U = \{u : z_i(0) = 0, i = \overline{(1+r)..m}\} \tag{6}$$

а $\widehat{c} = \sum\limits_{i=1}^{r} \widehat{\alpha}_i \widehat{z}_i(0)$, $\widehat{z}_i(0) = z_i(0)|_{u=\widehat{u}}$, де $z_i(0) = \varphi_i(T) - \int\limits_0^T u(t)\varphi_i(t)dt$.

Доведення. Зауважимо, що має місце рівність

$$\sigma^2(R, \alpha, u, c) = (\sum_{i=1}^{m} \alpha_i z_i(0) - c)^2 + \int_0^T \int_0^T R(t, s)u(t)u(s)dtds \tag{7}$$

$$\sigma_1^2(u, c) = \sup_G (\sum_{i=1}^{m} \alpha_i z_i(0) - c)^2 + \sup_R \int_0^T \int_0^T R(t, s)u(t)u(s)dtds = I_1(u, c) + I_2(u) \tag{8}$$

Крім того $I_1 = \infty$ при $u \notin U$ і

$$I_1(u, c) = \left(\sum_{i=1}^{r} \beta_i |z_i(0)| + \left| c - \sum_{i=1}^{r} \widehat{\alpha}_i z_i(0) \right| \right)^2 \tag{9}$$

при $u \in U$.

$$I_2(u) = \int_0^T \int_0^T R_0(t, s)u(t)u(s)dtds + q^2 \int_0^T u^2(t)dt \tag{10}$$

Далі зауважимо, що так як функції $\varphi_i(t)$ лінійно незалежні, то множина U є непорожньою опуклою замкненою в $L_2(0,T)$ множиною. Функціонал

$$\sigma_2^2(u) = \left(\sum_{i=1}^{r} \beta_i |z_i(0)|\right)^2 + I_2(u) \tag{11}$$

є слабконапівнеперервним, сильно опуклим функціоналом і значить існує єдина функція $\widehat{u} \in \arg\min \sigma_2^2$, в силу нерівності

$$\sigma_1^2(u,c) \geq \sigma_2^2(u) \geq \min_{u\in U} \sigma_2^2(u) = \sigma_2^2(\widehat{u}) \tag{12}$$

яка перетворюється в рівність при $u = \widehat{u}$, $c = \widehat{c}$, одержимо необхідне. □

Позначимо далі через V множину матриць B, що мають вигляд

$$\begin{aligned} &B = (b_{ij})_{i,j=\overline{1,r}} \\ &b_{ij} = \int_{-1}^{1} .. \int_{-1}^{1} x_i x_j \nu(dx) \end{aligned} \tag{13}$$

$\nu(\cdot)$ – пробігає множину ймовірносних мір, що зосереджені на гіперкубі $[-1..1] \times .. \times [-1..1]$.

Твердження 2. Має місце рівність

$$\min_{u\in U} \sigma_2^2(u) = \max_{B\in V} \min_{u\in U} \sigma_3^2(u,B) \tag{14}$$

де

$$\begin{aligned} &\sigma_3^2(u,B) = \sum_{i,j=1}^{r} \beta_i\beta_j b_{ij} z_i(0) z_j(0) + I_2(u) = (Bz_1(0), z_1(0)) + I_2(u) \\ &z_1(0) = (\beta_1 z_1(0), .., \beta_r z_r(0)) \end{aligned} \tag{15}$$

Доведення. Так як

$$\begin{aligned} &\left(\sum_{i=1}^{r} \beta_i |z_i(0)|\right)^2 = \max_{-1\leq x_i\leq 1} \left(\sum_{i=1}^{r} \beta_i x_i z_i(0)\right)^2 = \\ &= \max_{\nu(\cdot)} \int_{-1}^{1} .. \int_{-1}^{1} \left(\sum_{i=1}^{r} \beta_i x_i z_i(0)\right)^2 \nu(dx) = \max_{B\in V} (Bz_1(0), z_1(0)) \end{aligned} \tag{16}$$

то в силу теореми про мінімакс одержимо необхідну рівність. □

Твердження 3. Має місце рівність

$$\min_{u\in U} \sigma_3^2(u,B) = \sigma_3^2(\widehat{u},B) \tag{17}$$

Доведення. Нехай $z_i(t)$ є розв'язком рівняння

$$\frac{z_i(t)}{dt} = -\varphi_i(t)u(t), z_i(T) = \varphi_i(T), i = \overline{1,..,m}, z_i(0) = 0, i = \overline{(r+1),..,m} \tag{18}$$

Введемо також функції $p_i(t)$, що є розв'язками рівнянь

$$\frac{p_i(t)}{dt} = 0, i = \overline{1,..,m} \tag{19}$$

$$p_i(0) = \sum_{j=1}^{r} b_{ij}\beta_j z_j(0), i = \overline{1,..,r} \tag{20}$$

Тоді похідна Гато від функціонала $\sigma_3^2(u, B)$ буде мати вигляд

$$\frac{1}{2}\left(\sigma_3^2\right)' = -\sum_{i=1}^{m} \varphi_i(t)p_i(t) + \int_0^T R_0(t,s)u(s)ds + q^2 u(t) \tag{21}$$

Таким чином для $\widehat{u}(s)$ одержимо інтегральне рівняння

$$\int_0^T R_0(t,s)\widehat{u}(s)ds + q^2\widehat{u}(t) = \sum_{i=1}^{m} \varphi_i(t)p_i(t) \tag{22}$$

Запишемо $\widehat{u}$ у іншому вигляді, для цього зауважимо, що $p_i(t)$ не залежить від t і тому якщо позначити їх через p_i, то одержимо, що

$$\widehat{u}(t) = \sum_{i=1}^{m} p_i\psi_i(t) \tag{23}$$

де $\psi_i(t)$ є розв'язком інтегрального рівняння

$$\int_0^T R_0(t,s)\psi_i(s)ds + q^2\psi_i(t) = \varphi_i(t) \tag{24}$$

Тоді $z_i(t) = -\sum_{j=1}^{m} p_i z_{ij}(t) + \varphi_i(T)$, де $z_{ij}(t) = \int_t^T \varphi_i(s)\psi_j(s)ds.$

Так як

$$p_i = p_i(0) = \sum_{j=1}^{r} b_{ij}\beta_j z_j(0), \quad i = \overline{1,..,r},$$

то для $z_j(0)$ одержимо систему рівнянь

$$z_i(0) = -\sum_{j=1}^{m}\sum_{k=1}^{r} b_{ik}z_{ij}(0)\beta_k z_k(0) + \varphi_i(T) \tag{25}$$

а константи p_i, $i = \overline{(r+1)..m}$ визначаються із системи рівнянь

$$\sum_{i=1}^{m} \int_{0}^{T} \varphi_j(t)\psi_i(t)dtp_i = \varphi_j(T), j = \overline{r+1,..,m} \tag{26}$$

□

Наслідок 1. Нехай r=0. Тоді мінімаксна оцінка має вигляд

$$\widehat{\widehat{P}} = \sum_{i=1}^{m} p_i \int_{0}^{T} \psi_i(t)y(t)dt \tag{27}$$

Величини p_i визначаються з системи рівнянь

$$\sum_{i=1}^{m} p_i \int_{0}^{T} \varphi_j(t)\psi_i(t)dt = \varphi_j(T), j = \overline{1,..,m} \tag{28}$$

Наслідок 2. Припустимо, що $R_0(t,s) = 0$. Тоді $\psi_i(t) = q^{-2}\varphi_i(t)$.

Наслідок 3. Мінімаксна оцінка має вигляд

$$\widehat{\widehat{P}} = \sum_{i=1}^{m} \widehat{p_i} \int_{0}^{T} \psi_i(t)y(t)dt + \widehat{c} \tag{29}$$

числа $\widehat{p_i}$ знаходяться з розв'язку задачі на мінімум наступної функції

$$F(p_1,..,p_m) = \left(\sum_{k=1}^{r} \beta_k |\varphi_k(T) - \sum_{i=1}^{m} p_i z_{ki}(0)|\right)^2 + \sum_{i,j=1}^{m} p_i p_j (R_0\psi_i, \psi_j) + q^2 \sum_{i,j=1}^{m} p_i p_j (\psi_i, \psi_j) \tag{30}$$

де

$$(R\psi_i, \psi_j) = \int_{0}^{T} .. \int_{0}^{T} R_0(t,s)\psi_i\psi_j dtds, (\psi_i, \psi_j) = \int_{0}^{T} .. \int_{0}^{T} \psi_i\psi_j dtds, \tag{31}$$

при обмеженнях

$$\varphi_i(T) = \sum_{k=1}^{m} p_k \int_{0}^{T} \varphi_i(t)\psi_k(t)dt, \quad i = \overline{r+1,..,m} \tag{32}$$

Доведення. Із твердження 3 одержимо наступний вигляд функції $\widehat{u}(t)$ на якій досягається мінімум середньоквадратичної похибки оцінки $\widehat{P}(T)$, $\widehat{u}(t) = \sum_{i=1}^{m} \widehat{p}_i\psi_i(t)$, враховуючи вигляд функціоналу $\sigma^2(u,c)$ одержимо необхідне. □

Приклад 1. Проілюструємо застосування наслідку 3 на прикладі наступної моделі $\varphi_1(t) = \sin(t), \varphi_2(t) = \sin(t^2), \varphi_3(t) = \cos(t), \beta_1 = 2, R_0 = 0, \widehat{\alpha_1} = 0$. Випадковий процес $\xi(t)$ має вигляд $\xi(t) = \xi_1 * sin(20 * t)$, де ξ_1 – випадкова величина, що має нормальний розподіл з параметрами $(0, 0.2)$. Нехай $\alpha = (1\ 1\ 1)^T$ і реалізація випадкового процесу $\xi(t)$ має вигляд $\xi(t) = -0.226 * \sin(20 * t)$. Процедура оцінювання буде проводитися в умовах коли всі зазначені вище параметри фіксовані, а параметри T та q будуть варіюватися в межах $T \in [1, 3]$ та $q \in [0.2, 2.2]$ з кроком 0.05 та 0.4. Відповідний алгоритм було реалізовано в системі Mathematica 6, результати обчислень зображені на наступних графіках.

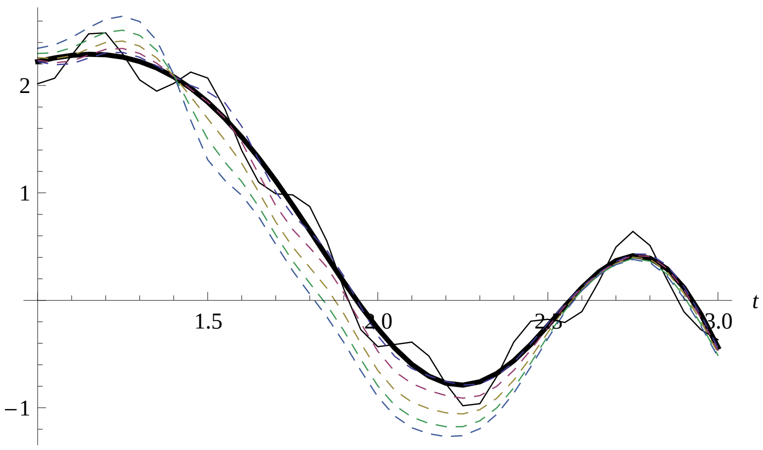

Рис. 1. Реальні значення $P(T)$ (суцільна лінія), реалізація $y(t)$ (суцільна тонка лінія), оцінки $\widehat{P(T)}$ при різних значеннях параметра q (штрихові лінії)

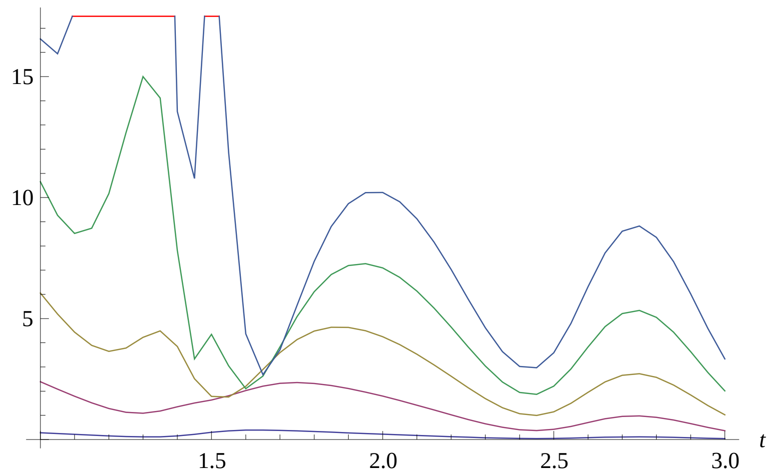

Рис. 2. Мінімаксна середньоквадратична похибка при різних значеннях параметра q

Отже, як видно з малюнку 1, при зменшенні параметра q якість оцінювання покращується, при $q = 0.2$ оцінки розташовані найближче до реальних значень тренду. При цьому найменші значення мінімаксної середньоквадратичної похибки теж досягаються при $q = 0.2$ (мал. 2).

3. Верхня та нижня мінімаксні оцінки

Означення 2. Оцінка $\widehat{\widehat{P}}_+(T)$ та $\widehat{\widehat{P}}_-(T)$ називаються верхньою та нижньою мінімаксною якщо вони мають вигляд

$$\begin{aligned} \widehat{\widehat{P}}_+(T) &= \int\limits_0^T \widehat{u}_+(t)y(t)dt + \widehat{c}_+ \\ \widehat{\widehat{P}}_-(T) &= \int\limits_0^T \widehat{u}_-(t)y(t)dt + \widehat{c}_- \end{aligned} \tag{33}$$

де $(\widehat{u}_+, \widehat{c}_+)$, $(\widehat{u}_-, \widehat{c}_-)$ знаходяться із мінімізації функціоналів $\sigma_+^2(u, c), \sigma_-^2(u, c)$ відповідно і для яких справедливі нерівності

$$\sigma_-^2(u, c) \leq \sigma^2(u, c) \leq \sigma_+^2(u, c) \tag{34}$$

Твердження 4. Оцінки для яких має місце рівність

$$\begin{aligned} \widehat{u}_+(t) &= \sum_{i=1}^m p_i^+ \psi_i(t), \\ \widehat{u}_-(t) &= \sum_{i=1}^m p_i^- \psi_i(t) \end{aligned} \tag{35}$$

для яких числа $p_i^\pm, i = \overline{1, .., m}$ знаходяться із систем рівнянь

$$\psi_i(T) = \sum_{i=1}^m p_k^\pm + z_{ik}(0) + \delta_{ir} p_i^\pm (\beta_i^\pm)^2 \tag{36}$$

де

$$\delta = \begin{cases} 1, & i \leq r \\ 0, & i > r \end{cases}$$

є відповідно нижніми та верхніми мінімаксними оцінками.

Доведення. Введемо множини G_+ та G_-

$$\begin{aligned} G_+ &= \{\alpha : \sum_{i=1}^r \frac{(\alpha_i - \widehat{\alpha}_i)^2}{{\beta_i^+}^2} \leq 1\} \\ G_- &= \{\alpha : \sum_{i=1}^r \frac{(\alpha_i - \widehat{\alpha}_i)^2}{{\beta_i^-}^2} \leq 1\} \end{aligned} \tag{37}$$

числа β_i^+ та β_i^- виберемо так, щоб $G_- \subset G \subset G_+$. Тоді очевидно, що

$$\begin{aligned} \sigma_+^2(u,c) &= \sup_{G_+,R} \sigma^2(R,\alpha,u,c) \\ \sigma_-^2(u,c) &= \sup_{G_-,R} \sigma^2(R,\alpha,u,c) \end{aligned} \tag{38}$$

Легко бачити, що

$$\sigma_+^2(u,c) = ((\sum_{i=1}^{r}(\beta_i^+)^2 z_i^2(0))^{\frac{1}{2}} + |c - \sum_{i=1}^{r}\widehat{\alpha}_i z_i(0)|)^2 + (R_0 u, u) + q^2(u,u) \tag{39}$$

Звідки $\widehat{c}_+ = \sum_{i=1}^{r}\widehat{\alpha}_i \widehat{z}_i^+(0)$, де $z_i^+(t)$ є розв‘язком рівняння

$$\frac{d\widehat{z}_i^+}{dt} = -\varphi_i(t)\widehat{u}^+(t), \widehat{z}_i^+(T) = \varphi_i(T) \tag{40}$$

де $\widehat{u}^+(t) \in \arg\min \sigma_+^2(u)$, $\sigma_+^2 = \sum_{i=1}^{r}(\beta_i^+)^2 z_i^2(0) + (R_0 u, u) + q^2(u,u)$

Звідки як і в твердженні 3 одержимо, що

$$\widehat{u}_+(t) = \sum_{i=1}^{m} p_i^+ \psi_i(t) \tag{41}$$

де p_i^+ знаходиться із системи рівнянь

$$\begin{cases} p_j^+(\beta_j^+)^2 = \varphi_j(T) - \sum_{i=1}^{m} p_i^+ z_{ij}(0),\ j = \overline{1,..,r} \\ \varphi_j(T) = \sum_{i=1}^{m} p_i^+ z_{ij}(0),\ j = \overline{1+r,..,m} \end{cases}$$

Зауваження 1. Нехай множина значень параметрів є G_+, тоді мінімаксна оцінка співпадає із верхньою мінімаксною оцінкою.

$\square$

Висновки

В роботі показано, що при деяких умовах існують єдині мінімаксні середньоквадратичні оцінки узагальнених поліномів. Дані оцінки знайдені у вигляді лінійної комбінації розв'язків інтегральних рівнянь Фредгольма другого роду. Вводяться поняття верхніх і нижніх мінімаксних оцінок та наводиться їх вигляд. При певних умовах на коефіцієнти розроблено алгоритми для обчислення оцінок. В системі Mathematica 6 створено відповідну програмну реалізацію.

Список літератури

1. *Лоран Г.Ж.* Апроксимация и оптимизация. – М.: Мир, 1975. – 496 с.
2. *Тихомиров В.М.* Некоторые вопросы теории приближений. – М.: МГУ, 1980. – 304 с.
3. *Себер Д.* Линейный регрессионный анализ. – М.: Мир, 1980. – 456 с.
4. *Норман Дрейпер, Гарри Смит* Прикладной регрессионный анализ. – 3-е изд. – М.: «Диалектика», 2007. – С. 912.
5. *Полянин А. Д., Манжиров А. В.* Справочник по интегральным уравнениям. – М.: Физматлит, 2003. – 608 с.

Статья поступила в редакцию 13.10.2008

УДК 519.23

ЭФФЕКТИВНЫЙ МЕТОД РАНЖИРОВАНИЯ НЕЗАВИСИМЫХ ПЕРЕМЕННЫХ И ОТБРАСЫВАНИЯ НЕСУЩЕСТВЕННЫХ ПАРАМЕТРОВ ПРИ МНОГОФАКТТОРНОМ СТАТИСТИЧЕСКОМ АНАЛИЗЕ[1]

© Поляков Б.Н.

E-MAIL: *bpoliakov@hotmail.com*

Abstract. The reliable criteria of ranging of independent variables and rejection of insignificant parameters at the multifactorial statistical analysis substantiations are resulted and are offered, which efficiency is illustrated by a concrete example and proves to be true more than 30-years practice of successful carrying out of statistical researches in mechanical engineering, metallurgy and medicine.

Введение

В работе [1] отмечалось, что для нахождения криволинейного уравнения множественной регрессии методом Брандона, независимые переменные необходимо ранжировать, т.е. располагать последние в порядке уменьшения силы их влияния на зависимую переменную. Ранжировать независимые переменные можно на основе линейного регрессионного анализа, как первого этапа многофакторного анализа, несколькими способами: по коэффициенту полной корреляции d_{1i}, по коэффициенту частной корреляции r_{1i} или по стандартизованному коэффициенту регрессии β_i.

Коэффициент полной корреляции d_{1i} характеризует тесноту связи между зависимой переменной x_1 и независимой x_i вне зависимости от того, чем обусловлена эта связь, действительным влиянием x_i, либо влиянием других независимых переменных, корреляционно связанных с x_i и, вследствие этого, искажающих силу влияния рассматриваемой независимой переменной на зависимую. Особенно это отмечается в том случае, когда независимые переменные сильно коррелируют между собой.

Стандартизованный коэффициент регрессии β_i является количественной характеристикой силы влияния независимой переменной, выраженной в единицах среднеквадратического отклонения зависимой переменной при устранении другой линейной связи с остальными независимыми переменными. Но в анализе связи двух переменных обычно принято пользоваться не абсолютными оценками, а относительными, т.е. более общими характеристиками.

Коэффициент частной корреляции r_{1i} является относительной характеристикой силы влияния независимой переменной на зависимую при постоянстве других, участвующих в анализе, т.е. выражает влияние, очищенное от действия других независимых факторов. Из определения коэффициента частной корреляции ясно, что последний является наилучшей оценкой силы связи между независимой и зависимыми переменными на основе линейного приближения к эмпирическим данным. И поэтому принято в последовательном многофакторном анализе [1] независимые переменные располагать по мере убывания $|r_{1i}|$.

---

[1]Разработан совместно с канд.техн.наук Ю.Д. Макаровым и инж.- математиком Ф.М.Карлинской

1. Критерии ранжирования независимых переменных и отбрасывания несущественных параметров

Рассмотрим теперь вопрос отбрасывания несущественных параметров. Обычно исследователь стремится зафиксировать как можно больше независимых переменных, которые, по его мнению, каким-то образом влияют на изучаемое явление. При этом многие «независимые» факторы могут быть на самом деле тесно взаимосвязаны друг с другом. Большое количество независимых переменных иногда искажает физический смысл определяемого уравнения, к тому же коэффициенты регрессии, вычисленные для сильно коррелированных переменных, малонадежны и будут иметь широкие доверительные интервалы, поэтому часть независимых переменных на основе какого-либо критерия необходимо исключить из рассмотрения. Таким образом, возникает вопрос о критерии отбрасывания несущественных параметров.

О существенности влияния независимой переменной на зависимую переменную можно судить уже по величине доверительного интервала коэффициента регрессии. Если доверительный интервал проходит через нуль, то вопрос о существенности влияния соответствующей независимой переменной ставится под сомнение. Поэтому выполнение неравенства (1) будет являться естественным критерием существенности независимого фактора

$$t_{iI} = \left| \frac{a_i}{S_{a_i}} \right| > t_\alpha, \tag{1}$$

где a_i – коэффициент регрессии, S_{a_i} – его среднеквадратическое отклонение, t_α – квантиль нормированного нормального распределения, соответствующий вероятности $(1-\alpha)$. Данный критерий (неравенство (1)) будем называть первым, на что указывает римская цифра I у индекса в обозначении критерия $t_{\alpha I}$. Но очень часто, по многим причинам: неправильно выбраны независимые параметры, недостаточны величина выборки и точность экспериментальных данных, и вследствие этого, небольшая точность результатов анализов и т.д., – почти для всех коэффициентов регрессии несправедливо неравенство (1) или доверительные интервалы этих коэффициентов проходят через нуль. И одновременное отбрасывание несущественных независимых переменных по критерию I может резко уменьшить коэффициент множественной корреляции, увеличить стандартную ошибку оценки и вообще привести к неправильному объяснению изучаемого процесса.

Число же всевозможных вариантов отбрасывания независимых переменных на основе критерия I растет по закону $C_k^1 + C_k^2 + ... + C_k^{k-1} + C_k^k$ где k – количество независимых переменных, которые нужно было бы отбросить по критерию I.

В. Визорке и др. [2] предложили метод одновременного отбрасывания нескольких независимых переменных на основе собственного опыта, полученного при обработке большого количества экспериментальных данных. Суть этого метода в следующем.

Рассмотрим величину $t_i = a_i/S_{a_i}$, эквивалентной формулой которой будет являться следующая:

$$t_i = \frac{\beta_i \sqrt{1 - R_i^2}}{\sqrt{1 - R^2}} \sqrt{n - m},$$

где β_i – стандартизованный коэффициент регрессии, R – коэффициент множественной корреляции, R_i – коэффициент множественной корреляции i – той независимой переменной с остальными независимыми переменными, n – количество значений зависимой переменной, m – число рассматриваемых параметров, включая зависимый. Если x_i коррелирует только с одной независимой переменной, и если последняя является несущественной по критерию I, то после ее отбрасывания t_i возрастает в $1/\sqrt{1 - R_i^2}$ раз, что не учитывает рассмотренный выше критерий.

Поэтому необходимо относиться очень осторожно к независимым переменным, которые тесно взаимосвязаны с другими независимыми переменными, и в первую очередь, рекомедуется отбросить те независимые переменные, несущественные по критерию I , коррелирующие слабо с другими, так как данный критерий не учитывает взаимной корреляции между независимыми переменными. Но независимая переменная может коррелировать с несколькими независимыми переменными и t_i может возрасти [2] приблизительно в $1 + R_i\sqrt{2}/\sqrt{1 - R_i^2}$ раз после отбрасывания несущественных параметров и

$$t_{iII} = |t_{iI}| \frac{1 + R_i\sqrt{2}}{\sqrt{1 - R_i^2}} > t_\alpha, \tag{2}$$

Экспериментальная проверка показала, что критерий II отбрасывания нескольких независимых параметров очень слаб, что будет проиллюстрировано в приведённом ниже примере. Для практического применения этого критерия можно его усилить. В алгоритме многофакторного статистического анализа [1] заложен следующий критерий:

$$t_{iIII} = |t_{iI}| \frac{1}{\sqrt{1 - R_i^2}} > t_\alpha, \tag{3}$$

Критерий отбрасывания II отличается от критерия III сомножителем $1 + R_i\sqrt{2}$ в числителе правой части выражения (2).

На Рис. 1. и Рис. 2. показаны зависимости $1+R_i\sqrt{2}/\sqrt{1 - R_i^2}$ и $1/\sqrt{(1 - R_i^2)}$ от R_i. Из рассмотрения этих графиков и неравенств (2) и (3) можно сделать вывод о том, что в первую очередь будут отброшены те независимые переменные, для которых мало значение R_i, т.е. слабо связанные с другими независимыми переменными.

После работы критерия II могут остаться ещё несколько независимых переменных существенных по критерию III, но не существенных по критерию I. В этом случае предлагается на втором этапе исследования несущественные переменные отбрасывать по следующему критерию:

$$t_{iIV} = |t_{iI}| \frac{1 + |r_{1i}|\sqrt{2}}{\sqrt{1 - r_{1i}^2}} > t_\alpha, \tag{4}$$

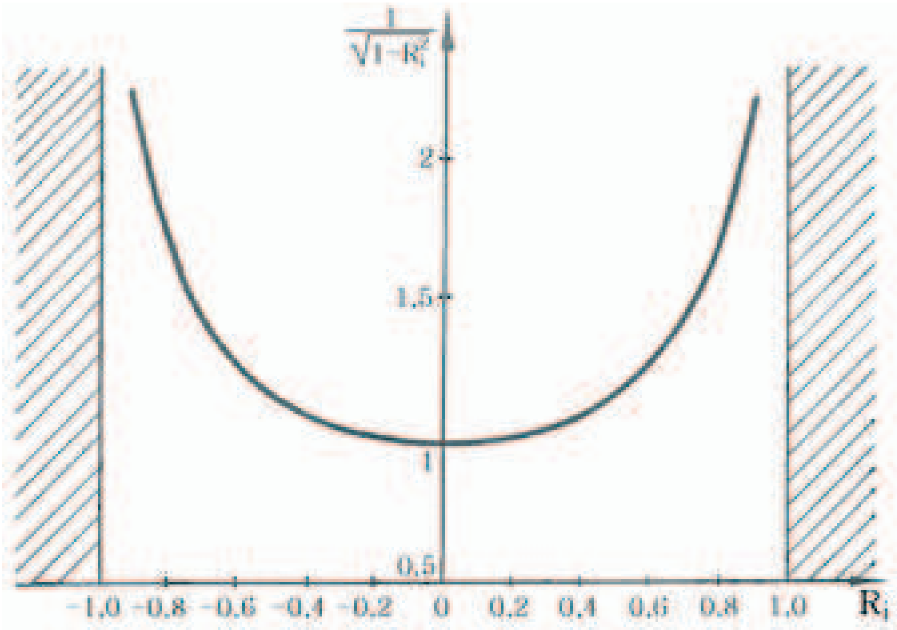

Рис. 1. Зависимость $\frac{1}{\sqrt{1-R_i^2}}$ от R_i

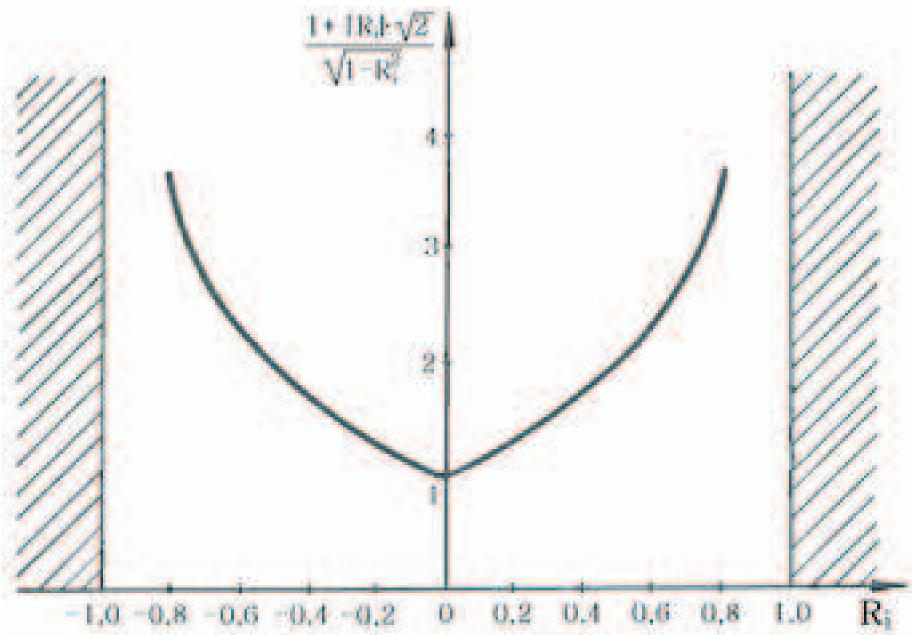

Рис. 2. Зависимость $\frac{1+|R_i|\sqrt{2}}{\sqrt{1-R_i^2}}$ от R_i

Сопоставляя критерии отбрасывания независимых переменных II и IV, можно заметить, что они имеют одинаковую структуру. Поэтому график функции

$1 - R_i\sqrt{2}/\sqrt{1-R_i^2}$ полностью совпадает с графиком функции $1 + |r_{1i}|\sqrt{2}/\sqrt{1-r_{1i}^2}$. Из Рис. 2 и неравенства (4) следует, что в первую очередь будут отбрасываться те несущественные независимые переменные, которые слабо влияют на зависимую переменную. Практическое применение критерия IV дало положительный результат. Критерий IV сильнее критерия III, поэтому он применяется уже на втором этапе отбрасывания независимых переменных.

2. Пример

Действия приведенных выше критериев проиллюстрированы примером. Опытные данные для примера взяты из источника [3], в котором обобщены технологические режимы прокатки на блюмингах ряда отечественных металлургических заводов и дополнены данными по схемам и режимам обжатий, применяющимися на блюминге 1300 меткомбината «Криворожсталь».

Пример. На основе опытных данных определяется регрессионная зависимость количества пропусков (N) от среднего обжатия в пропуске за цикл прокатки ($\Delta_{\text{ср}}$, мм), площади поперечного сечения слитка ($F_{\text{сл}}$, мм^2), площади поперечного сечения конечной заготовки ($F_{\text{пр}}$, мм^2), числа кантовок (K), массы слитка ($Q_{\text{сл}}$, т), числа пропусков до первой кантовки (n_{K1}), числа пропусков между первой и второй кантовкой (n_{K2}), числа пропусков между второй и третьей кантовкой (n_{K3}), числа пропусков между третьей и четвертой кантовкой (n_{K4}).

Таблица 1. Первое приближение

Параметр	a_i	$a_i^{(1)}$	$a_i^{(2)}$	r_{1i}	t_{iI}	t_{iII}	t_{iIII}	t_{iIV}	$\sqrt{1-R_i^2}$
N	13,81	13,99	13,63	-	7,522	-	-	-	-
$\Delta h_{\text{ср}}$, мм	-0,094	-0,077	-0,110	-0,857	-10,95	19,08	12,033	-47,04	0,910
K	0,278	0,770	-0,214	0,166	1,108	6,547	2,843	1,388	0,389
$Q_{\text{сл}}$, т	0,235	0,517	0,047	0,241	1,630	11,47	4,916	2,253	0,332
$F_{\text{сл}}$, мм^2	$0,88\cdot$ $\cdot 10^{-5}$	$0,13\cdot$ $\cdot 10^{-4}$	$0,45\cdot$ $\cdot 10^{-5}$	0,521	3,999	25,36	10,946	8,134	0,365
$F_{\text{пр}}$, мм^2	$-0,3\cdot$ $\cdot 10^{-4}$	$-0,19\cdot$ $\cdot 10^{-4}$	$-0,40\cdot$ $\cdot 10^{-4}$	-0,640	-5,462	17,61	8,463	-13,54	0,645
n_{K1}	0,082	0,231	-0,068	0,161	1,068	2,825	1,462	1,329	0,731
n_{K2}	0,086	0,325	0,154	0,106	0,700	1,735	0,929	0,810	0,753
n_{K3}	0,216	0,434	-0,001	0,285	1,949	6,534	3,110	2,836	0,627
n_{K4}	0,291	0,604	-0,023	0,267	1,818	8,873	3,936	2,600	0,462
S=0,604 R=0,944									

Первое приближение запишется следующим образом (см. также таблицу):

$$N = 13,81 - 0,094\Delta h_{\text{ср}} - 0,3\cdot 10^{-4} F_{\text{пр}} + 0,88\cdot 10^{-5} F_{\text{сл}} + 0,216 n_{K3} +$$
$$+0,291 n_{K4} + 0,235 Q_{\text{сл}} + 0,278K + 0,082 n_{K1} + 0,086 n_{K2}.$$

Таблица 2. Второе приближение

N	14,66	14,84	14,48	-	7,985	-	-	-	-
$\Delta h_{\text{ср}}$, мм	-0,094	-0,077	-0,110	-0,855	-11,05	19,19	12,13	47,03	0,911
K	0,231	0,704	-0,243	-0,141	0,954	5,454	2,376	1,155	0,402
$Q_{\text{сл}}$, т	0,206	0,478	-0,065	0,217	1,493	10,14	4,355	1,999	0,343
$F_{\text{сл}}$, мм^2	$0,94\cdot$ $\cdot 10^{-5}$	$0,135\cdot$ $\cdot 10^{-4}$	$0,53\cdot$ $\cdot 10^{-5}$	0,556	4,498	27,30	11,83	9,665	0,380
$F_{\text{пр}}$, мм^2	$-0,31\cdot$ $\cdot 10^{-4}$	$-0,21\cdot$ $\cdot 10^{-4}$	$-0,41\cdot$ $\cdot 10^{-4}$	-0,667	-5,981	17,84	8,773	15,59	0,682
n_{K3}	0,202	0,386	0,016	0,303	2,132	5,618	2,913	3,195	0,732
n_{K4}	0,249	0,551	-0,053	0,234	1,617	7,622	3,397	2,215	0,476
S=0,615 R=0,942									

Таблица 3. Третье приближение

N	15,38	15,56	15,20	-	8,377
$\Delta h_{\text{ср}}$, мм	-0,95	- 0,079	- 0,112	-0,861	-11,64
$Q_{\text{сл}}$, т	0,229	0,496	-0,037	0,241	1,684
$F_{\text{сл}}$, мм^2	$0,93\cdot$ $\cdot 10^{-5}$	$0,134\cdot$ $\cdot 10^{-5}$	$0,52\cdot$ $\cdot 10^{-5}$	0,547	4,429
$F_{\text{пр}}$, мм^2	$-0,32\cdot$ $\cdot 10^{-5}$	$-0,23\cdot$ $\cdot 10^{-5}$	$-0,42\cdot$ $\cdot 10^{-5}$	-0,698	-6,969
n_{K3}	0,249	0,407	0,091	0,414	3,089
n_{K4}	0,367	0,547	0,188	0,509	3,998
S=0,621 R=0,941					

Доверительные интервалы коэффициентов регрессии следующих параметров проходят через нуль: K, $Q_{\text{сл}}$, n_{K1}, n_{K2}, n_{K3}, n_{K4}. По критерию II отбрасывается только n_{K2} по критерию III должно отбрасываться n_{K1} и n_{K2} а по критерию IV – n_{K1}, n_{K2}, K. Так как в программе многофакторного анализа сначала работает критерий III, то второе приближение, после отбрасывания n_{K1} и n_{K2} будет следующим:

$$N = 14,66 - 0,094\Delta h_{\text{ср}} - 0,31 \cdot 10^{-4} F_{\text{пр}} + 0,94 \cdot 10^{-5} F_{\text{сл}} + 0,202 n_{K3} + \\ +0,249 n_{K4} + 0,206 Q_{\text{сл}} + 0,231 K.$$

При этом коэффициент множественной регрессии уменьшается с $0,944$ до $0,942$, а остаточное среднеквадратическое отклонение увеличивается с $0,604$ до $0,615$.

Во втором приближении доверительные интервалы коэффициентов $K\,Q_{\text{сл}}$, n_{K4} регрессии параметров проходят через нуль, но критерии II и III дают отрицательный ответ на отбрасывание этих независимых параметров. По критерию IV можно отбросить независимый параметр K. Тогда в третьем приближении уравнение регрессии запишется следующим образом:

$$N = 15,38 - 0,095\Delta h_{\text{ср}} - 0,32 \cdot 10^{-4} F_{\text{пр}} + 0,93 \cdot 10^{-5} F_{\text{сл}} + 0,367 n_{K4} + \\ +0,249 n_{K3} + 0,229 Q_{\text{сл}}.$$

Хотя после третьего приближения доверительный интервал у независимого параметра $Q_{\text{сл}}$ проходит через нуль, но ни один из выше рассмотренных критериев (II, III, IV) не отбрасывает его, так как t_{iI}, равный $1,684$, достаточно близок к $t_\alpha = 1,96$ и влияние на зависимую переменную N значительно $(r_{14} = 0,241)$.

После третьего приближения коэффициент множественной корреляции равен $0,941$, а остаточное среднеквадратическое отклонение равно $0,621$.

Отбрасывание неэависимых параметров n_{K1}, n_{K2} согласуется с физическим процессом прокатки на блюминге, потому что среднее обжатие до первого ящичного калибра (где прокатка идет со стесненным уширением) определяется только размерами слитка и его положением при начальном обжатии, а в дальнейшем – размерами поперечного сечения конечного раската, последовательностью расположения и шириной калибров, которые в какой-то степени характеризуются параметрами n_{K1} и n_{K2}. Отсюда очевидно, что и количество кантовок мало влияет на среднее обжатие за цикл прокатки.

Заключение

Таким образом, предлагаемые критерии ранжирования независимых переменных и отбрасывания несущественных параметров, проверенные на многочисленных примерах из более чем 30-летней практики успешного проведения статистических исследований в машиностроении, металлургии и медицине [4, 5], которые подтверждают их надёжность и эффективность для проведения разнообразных многофакторных статистических исследований.

Список литературы

1. *Реноме* Статистические методы в алгоритмах и примерах (из практики прокатного производства). Учебное пособие ". ISBN 978-5-98947-081-5„СПб.: "Реноме", декабрь 2007, - 182с.
2. *Von H. Knüppel , Stumpf A., Wiezorke B.* Mathematische Statistik in Eisenhüttenwerken, Archive fürdas Eisenhüttenwesen, №8, 1958. Перевод № 1492. НИИТЯЖМАШ Уралмашзавода, 1968.
3. *Логоватовский А.А.* Нормирование процессов на блюминге. М.: Металлургия, 1966. 220с.
4. *Коцарь С.Л., Поляков Б.Н., Макаров Ю.Д., Чичигин В.А.* Статистический анализ и математическое моделирование блюминга М: Металлургия, 1974. 280с.
5. *Поляков Б.Н.* Повышение качества технологий, несущей способности конструкций, долговечности оборудования и эффективности автоматических систем прокатных станов. - СПб.: Реноме , 2006, - 528с.

Статья поступила в редакцию 25.10.2008

УДК 519.7

НЕКОТОРЫЕ ПОЛОЖЕНИЯ И ЗАДАЧИ ТЕОРИИ ШАБЛОНОВ

© Анафиев А.С.

Таврический национальный университет им. В.И. Вернадского
факультет математики и информатики
пр-т Вернадского, 4, г. Симферополь, 95007, Украина
e-mail: *aydera@mail.ru*

Abstract. The main definitions, positions and problems of template theory are considered. The new types of templates are introduced. Some properties of polynomial templates are considered. New tasks which play an important role of the construction of high-quality decision rules are emphasized.

Введение

Рассмотрим стандартную постановку задачи обучения по прецедентам. Дано *множество объектов* X, *множество ответов* Y и некоторая неполная информация о целевой зависимости $f^* : X \to Y$ представленная в виде конечного *набора прецедентов* $\{(x_i \in X,\, y_i \in Y)\}_{i=1}^{\ell}$, где $y_i = f^*(x_i)$, $i = 1..\ell$. При этом набор $\{x_1, \ldots, x_\ell\} \subset X$ будем называть *обучающей выборкой* и обозначать X^ℓ.

Задача обучения по прецедентам заключается в том, чтобы восстановить неизвестную целевую зависимость f^* между объектами и ответами, т.е. построить алгоритм $a^* : X \to Y$, который будет удовлетворять следующим требованиям.

- Алгоритм $a^*(x)$ должен воспроизводить на объектах обучающей выборки заданные ответы: $a(x_i) = y_i$, $i = 1..\ell$. Причем, равенство здесь можно понимать и как приближенное, в зависимости от специфики задачи.
- На алгоритм a^* могут накладываться разного рода априорные ограничения, такие как, линейность, монотонность, гладкость и т.д., а также и их сочетания. Более того, может быть задана некоторая модель данного алгоритма [1].
- Алгоритм a^* должен достаточно хорошо восстанавливать неизвестную целевую зависимость f^* не только на объектах обучающей выборки, но и на всем множестве объектов X. Другими словами, алгоритм a^* должен обладать обобщающей способностью.

При этом фраза «достаточно хорошо восстанавливает» означает, что искомый алгоритм a^* должен быть лучшим в некотором семействе алгоритмов относительно некоторого заранее известного функционала качества.

Кроме этого, при решении задачи обучения по прецедентам заранее фиксируется класс решающих правил A, из которого и будет выбираться оптимальный алгоритм, удовлетворяющий вышеописанным требованиям.

При этом, в большинстве случаев, класс решающих правил A выбирается достаточно широким, что приводит к неприменимости многих оценок, например, оценок Вапника-Червоненкиса [3], для оценки качества полученного алгоритма и получения «рецепта» – как выбирать оптимальное (близкое к оптимальному) решающее правило в заданном семействе A.

Для сужения класса решающих правил A с целью построения более качественных решающих правил была предложена теория шаблонов [2]. Рассмотрим основные определения и положения данной теории.

1. Некоторые определения и положения теории шаблонов

Определение 1. Оператор $t : \Theta \to A$ будем называть *шаблоном* множества алгоритмов A относительно параметрического множества Θ. Если при этом $\Theta = P^k$, то шаблон t будем называть *k-параметрическим* и обозначать t^k.

Приведем здесь также определение модели алгоритмов [1].

Определение 2. *Моделью алгоритмов* называется параметрическое семейство отображений $A = \{\varphi(x, \theta) \mid \theta \in \Theta\}$, где $\varphi : X \times \Theta \to Y$ – некоторая фиксированная функция, Θ – множество допустимых значений параметра θ, называемое пространством параметров или пространством поиска.

Данное определение можно переформулировать, используя понятие шаблона.

Определение 3. *Моделью алгоритмов* называется параметрическое семейство отображений $A = \{f(x) \mid f = t(\theta),\, \theta \in \Theta\}$, где $t(\theta)$ – некоторый фиксированный шаблон, Θ – множество допустимых значений параметра θ.

Задавая модель алгоритма с помощью шаблона, мы акцентируем свое внимание на конструктивных особенностях решающих правил и на возможность сужения области неопределенности, задаваемой моделью алгоритмов. При этом параметрическое множество определяет количество параметров и их допустимые значения, а шаблон показывает каким образом связаны между собой параметры и переменные. *Тогда возникают следующие две интересные задачи: первая (прямая) – как можно использовать структуру и вид связей между параметрами и переменными для построения качественных решающих правил адекватных данной выборке; и вторая (обратная) – задача нахождения структуры выборки описываемой в виде шаблона.*

Также можно рассматривать модель отдельной функции.

Определение 4. Тройку $M_f = \langle\, \theta,\, t,\, \Theta \,\rangle$ такую, что $t(\theta) = f$, где t — шаблон множества A относительно параметрического множества Θ и $\theta \in \Theta$, будем называть *моделью функции* $f \in A$.

Определение 5. Множество

$$\Delta_t(\Theta) = \{f \in A \mid M_f = \langle\, \theta,\, t, \Theta \,\rangle\}$$

будем называть *образом шаблона* t на множестве A относительно параметрического множества Θ.

Определение 6. Шаблон t называется *универсальным шаблоном выборки* $X^\ell = (x_i, y_i)_{i=1}^{\ell}$, если $\forall f \in \Delta_t$ и $\forall i = \overline{1, \ell}$ выполняется $f(x_i) = y_i$.

Будем говорить, что шаблон t *удовлетворяет выборке* $X^\ell = (x_i, y_i)_{i=1}^{\ell}$, если $\exists f \in \Delta_t$ такая, что $\forall i = \overline{1, \ell}$ выполняется $f(x_i) = y_i$.

Замечание. В некоторых задачах равенство $f(x_i) = y_i$ можно понимать как приближенное.

Определение 7. Шаблон t_{min} называется *минимальным шаблоном для выборки* X^ℓ, если он имеет наименьшее число параметров среди всех удовлетворяющих выборке X^ℓ шаблонов.

Тогда становится актуальной задача построения минимального или близкого к минимальному шаблонов удовлетворяющих обучающей выборке! И какова сложность отыскания таких шаблонов для различного рода задач обучения по прецедентам.

Рассмотрим несколько различных видов шаблонов.

Обозначим $\tilde{\theta}^{(k)} = (\theta_1, \ldots, \theta_k)$.

Определение 8. Шаблон t^k называется *α-шаблоном*, если его можно представить в виде

$$t^k(\tilde{\theta}^{(k)}) = h\left(h_1\left(\psi_1\left(\tilde{\theta}^{(k)}\right), \varphi_1(x_1, \ldots, x_n)\right), \ldots, h_m\left(\psi_m\left(\tilde{\theta}^{(k)}\right), \varphi_m(x_1, \ldots, x_n)\right)\right), \tag{1}$$

где

1) $x_1, \ldots, x_n$ – символы переменных функций из A;
2) h – m-местный функциональный символ;
3) h_j – 2-местный функциональный символ;
4) φ_j – s-местные функциональные символы, $j = \overline{1, m}$, $s \in \{0, \ldots, n\}$, причем $\varphi_i \neq \varphi_j$, $i \neq j$, функции $\varphi_j(x_1, \ldots, x_n)$ будем называть *свойствами множества X выделяемые шаблоном* t^k;
5) ψ_j – p-местные функциональные символы, $j = \overline{1, m}$, $p \in \{0, \ldots, k\}$.

При этом, если $m = k$ и $\psi_j\left(\tilde{\theta}^{(k)}\right) = \theta_j$, $j = \overline{1, k}$, то шаблон t^k называется *простым.*

Пусть $\mathfrak{F}$ алгебраическое поле с операциями сложения $+$ и умножения $\cdot$. Простой шаблон с $h(z_1, \ldots, z_k) = z_1 + \ldots + z_k$, $h_j(z_1, z_2) = z_1 \cdot z_2$ и функциями-свойствами $\varphi_j(x_1, \ldots, x_n) = x_{i_1} \cdot \ldots \cdot x_{i_r}$, $j = \overline{1, k}$, называется *полиномиальным шаблоном* множества функций $A = \{f : \mathfrak{F}^n \to \mathfrak{F}\}$ над полем $\mathfrak{F}$ относительно параметрического множества $\Theta = \mathfrak{F}^k$.

Согласно определению, каждый полиномиальный шаблон t^k можно представить в виде

$$t(\theta_1, \ldots, \theta_k) = \theta_1 \cdot x_{i_1^1} \cdot \ldots \cdot x_{i_{r_1}^1} + \ldots + \theta_k \cdot x_{i_1^k} \cdot \ldots \cdot x_{i_{r_k}^k} = \sum_{j=1}^{k} \theta_j \cdot x_{i_1^j} \cdot \ldots \cdot x_{i_{r_j}^j}. \tag{2}$$

Причем произведения $s_j = x_{i_1^j} \cdot \ldots \cdot x_{i_{r_j}^j}$, $j = \overline{1, k}$, будут свойствами множества X, выделяемые шаблоном t. Число r_j будем называть *рангом* свойства s_j, а относительно шаблона t будем говорить, что он порожден множеством свойств $\mathrm{S} = \{s_1, s_2, \ldots, s_k\}$ и обозначать $t \leftleftarrows \mathrm{S}$.

Полиномиальный шаблон над полем $\mathfrak{F}$ будем обозначать $\mathfrak{t}_{\mathfrak{F}}$, а через $\mathfrak{t}^{k,r}$ – k-параметрический полиномиальный шаблон, который способен выделять из множества X только свойства ранга не больше, чем r.

Получены следующие оценки ёмкости класса решающих правил порожденных полиномиальными шаблонами [2].

Определение 9. Поля, элементы которых можно представить в виде $\log \mathrm{M}$-разрядного двоичного кода, $\mathrm{M} < \infty$, будем называть M-*полями*. M-поле, содержащееся во множестве рациональных (целых) чисел, будем называть M-*полем рациональных (целых) чисел.*

Теорема 1. *Пусть $\mathfrak{F}$ некоторое M-поле, тогда ёмкость класса $h(\Delta_{\mathfrak{T}_{\mathfrak{F}}^{k,r}})$ образа множества всех k-параметрических полиномиальных шаблонов, порожденных свойствами ранга не выше r, удовлетворяет неравенству*

$$\begin{aligned} h(\Delta_{\mathfrak{T}_{\mathfrak{F}}^{k,r}}) \leqslant \min(\quad & k(r\log(n+1) + \log \mathrm{M}), \\ & k(n\log(r+1) + \log \mathrm{M}), \\ & P(r) \log \mathrm{M}\,), \end{aligned}$$

где $P(r)$ — число свойств множества X, выделяемые всеми шаблонами из $\mathfrak{T}_{\mathfrak{F}}^{k,r}$, ранга не больше, чем r.

Следствие 1. *Ёмкость $h(P_{k,r}(n))$ класса $P_{k,r}(n)$ полиномов над* M-*полем рациональных чисел $\mathbb{Q}$ от n переменных степени не выше r с k отличными от нуля коэффициентами удовлетворяет неравенству*

$$\begin{aligned} h(P_{k,r}(n)) \leqslant \min(\quad & k(r\log(n+1) + \log \mathrm{M}), \\ & k(n\log(r+1) + \log \mathrm{M}), \\ & \frac{(n+r)!}{n!r!} \log \mathrm{M}\,). \end{aligned}$$

Следствие 2. *Ёмкость $h(P^2_{k,r}(n))$ класса $P^2_{k,r}(n)$ полиномов Жегалкина от n переменных степени не выше r с k отличными от нуля коэффициентами удовлетворяет неравенству*

$$h(P^2_{k,r}(n)) \leqslant \min\left(k(r\log(n+1)+1),\ k(n\log(r+1)+1), \sum_{i=0}^{r} C_n^i \right).$$

Как показано в [2], учет числа параметров при оценивании ёмкости класса решающих правил позволяет намного улучшить известные оценки. Кроме этого, данная теорема подчеркивает целесообразность построения минимальных (с минимальным числом параметров) шаблонов.

Приведем численную оценку зависимости вероятности неслучайной настройки на выборку от числа параметров k-параметрического полиномиального шаблона.

Теорема 2. *Для того, чтобы вероятность $\mathsf{P}(\Delta_{\mathfrak{t}^k},\, \ell,\, \delta\ell)$ настройки на какие-нибудь $\ell - \delta\ell$ элементов выборки X^ℓ длины ℓ алгоритмами семейства $\Delta_{\mathfrak{t}^k}$, – образа k-параметрического полиномиального шаблона, – была меньше некоторого заданного $\eta > 0$*

необходимо, чтобы число параметров k *удовлетворяло условию*

$$k < \frac{1}{\log \ell} \left(\ell - \delta\ell + \log \eta - \log \left(C_\ell^{\delta\ell} \cdot \gamma(\ell - \delta\ell) \right) \right),$$

где $\gamma(\ell - \delta\ell) = \dfrac{1}{1 - 2^{-(\ell-\delta\ell)+1}} \to 1$ *при* $\ell - \delta\ell \to \infty$, $\delta\ell$ *— число ошибок, допускаемых на обучающей выборке алгоритмом из семейства* $\Delta_{\mathfrak{t}^k}$ *построенным в результате обучения.*

Доказательство. По теореме 2.7 из [2]

$$\mathsf{P}(\Delta_{\mathfrak{t}^k},\, \ell,\, \delta\ell) < \gamma(\ell - \delta\ell) \cdot C_\ell^{\delta\ell} \cdot 2^{k(\log \ell - 1)} \cdot 2^{\,-\mathsf{R}'(X^\ell, \Delta_{\mathfrak{t}^k})}$$

или, согласно определению степени частичной закономерности R' [2],

$$\mathsf{P}(\Delta_{\mathfrak{t}^k},\, \ell,\, \delta\ell) < \gamma(\ell - \delta\ell) \cdot C_\ell^{\delta\ell} \cdot 2^{k(\log \ell - 1)} \cdot 2^{\,-(\ell-k-\delta\ell)}.$$

Тогда, если

$$\gamma(\ell - \delta\ell) \cdot C_\ell^{\delta\ell} \cdot 2^{k(\log \ell - 1)} \cdot 2^{\,-(\ell-k-\delta\ell)} < \eta, \tag{$*$}$$

то и $\mathsf{P}(\Delta_{\mathfrak{t}^k},\, \ell,\, \delta\ell) < \eta$. Из $(*)$ следует, что

$$\log \left(\gamma(\ell - \delta\ell) \cdot C_\ell^{\delta\ell} \cdot 2^{k(\log \ell - 1)} \cdot 2^{\,-(\ell-k-\delta\ell)} \right) < \log \eta,$$

$$\log \left(C_\ell^{\delta\ell} \cdot \gamma(\ell - \delta\ell) \right) + \log \left(2^{k(\log \ell - 1)} \right) + \log \left(2^{\,-(\ell-k-\delta\ell)} \right) < \log \eta,$$

$$\log \left(C_\ell^{\delta\ell} \cdot \gamma(\ell - \delta\ell) \right) + k(\log \ell - 1) - (\ell - k - \delta\ell) < \log \eta,$$

$$k(\log \ell - 1) - \ell + k + \delta\ell < \log \eta - \log \left(C_\ell^{\delta\ell} \cdot \gamma(\ell - \delta\ell) \right),$$

$$k \log \ell < \ell - \delta\ell + \log \eta - \log \left(C_\ell^{\delta\ell} \cdot \gamma(\ell - \delta\ell) \right).$$

Откуда следует утверждение теоремы

$$k < \frac{\ell - \delta\ell + \log \eta - \log \left(C_\ell^{\delta\ell} \cdot \gamma(\ell - \delta\ell) \right)}{\log \ell}.$$

□

Следствие 3. *Для того, чтобы вероятность* $\mathsf{P}(\Delta_{\mathfrak{t}^k},\, \ell,\, 0)$ *безошибочной настройки на выборку* X^ℓ *длины* $\ell \geqslant 10$ *алгоритмами семейства* $\Delta_{\mathfrak{t}^k}$ *была меньше некоторого заданного* $\eta > 0$ *необходимо, чтобы число параметров* k *удовлетворяло условию*

$$k < \frac{\ell + \log \eta}{\log \ell}.$$

Доказательство. Следует из теоремы 2, учитывая, что $\delta\ell = 0$ и $\gamma(\ell) \approx 1$ при $\ell \geqslant 10$ [1]. □

Кроме этого в [2] была получена численная оценка зависимости вероятности неслучайной настройки на выборку от числа параметров k-параметрического полиномиального шаблона для булевого случая (рассматривается булева выборка $X^{\ell}_{\{0,1\}}$).

Следствие 4. *Для того, чтобы вероятность* $\mathsf{P}(\Delta_{\mathfrak{t}^k}, \ell, \delta\ell)$ *безошибочной настройки на выборку* $X^{\ell}_{\{0,1\}}$ *длины* $\ell \geqslant 10$ *алгоритмами семейства* $\Delta_{\mathfrak{t}^k}$, *– образа k-параметрического шаблона над полем* $\{0,1\}$, *– была меньше некоторого заданного* $\eta > 0$, *необходимо, чтобы число параметров k удовлетворяло условию*

$$k < \ell - \log \ell + \log \eta.$$

Приведем таблицу, которая показывает каким должно быть число параметров k, чтобы вероятность неслучайной безошибочной настройки на выборку $X^{\ell}_{\{0,1\}}$ длины ℓ k-параметрическим полиномиальным шаблоном была выше 90%.

ℓ	$k \leqslant$
10	3
20	12
30	22
40	31
50	41
60	51

ℓ	$k \leqslant$
70	61
80	70
90	80
100	90
1000	987
10000	9983

2. Теорема существования полиномиального шаблона

Теорема 3 (существования полиномиального шаблона). *Для произвольной корректной выборки X^{ℓ} длины ℓ существует k-параметрический полиномиальный шаблон* $\mathfrak{t}^k_{\mathfrak{F}}$, $k \leqslant \ell$, *удовлетворяющий выборке X^{ℓ}.*

Здесь возникают следующие *интересные задачи*:

- *задача поиска условий существования и вида обучающей выборки X^{ℓ}, для которой не существует k-параметрического полиномиального шаблона с $k < \ell$ параметрами удовлетворяющего выборке X^{ℓ};*
- *задача нахождения минимального k, при котором для выборки X^{ℓ} существует удовлетворяющий ей k-параметрический полиномиальный шаблон.*

Заключение

В работе приведены основные определения, положения и проблемы теории шаблонов. Введены новые типы шаблонов, такие как, универсальные и α-шаблоны. Получено условие на количество параметров k-параметрического полиномиального шаблона t^k, при котором вероятность настройки на выборку алгоритмами семейства Δ_{t^k} будет меньше некоторого $\eta > 0$.

Выделены новые задачи теории шаблонов, которые играют важную роль при построении качественных решающих правил и сужении области неопределенности при решении задач обучения по прецедентам.

Список литературы

1. *Donskoy V. I.* The Estimations Based on the Kolmogorov Complexity and Machine Learning from Examples // Proceedings of the Fifth International Conference «Neural Networks and Artificial Intelligence» (ICNNAI'2008), Minsk, 2008, - p.p. 292-297.
2. *Анафиев А.С.* Теория шаблонов в задачах обучения по прецедентам и выбора моделей. Диссертация на соискание ученой степени кандидата физико-математических наук. – 2007. – 135 с.
3. *Вапник В.Н.* Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979. – 448 с.
4. *Воронцов К.В.* Вычислительные методы обучения по прецедентам. Курс лекций по машинному обучению. – 2009 г. – 42 с.
http://www.machinelearning.ru/wiki/images/8/8d/Voron-ML-Intro.pdf

Статья поступила в редакцию 21.12.2008

УДК 519.83

УСЛОВНО КОНТРОЛИРУЮЩИЕ СТРАТЕГИИ И ПОСЛЕДОВАТЕЛЬНЫЙ ВЫБОР РЕШЕНИЙ В МНОГОШАГОВОЙ ИГРЕ С БУЛЕВЫМИ СТРАТЕГИЯМИ

© Блыщик В.Ф.

Таврический национальный университет им. В.И. Вернадского
факультет математики и информатики
пр-т Вернадского, 4, г. Симферополь, 95007, Украина
e-mail: *veb@land.ru*

Abstract. The multimonic games with Boolean strategies and sequential selection of actions by players are considered. The notion of conditionally controlling strategy and suggest the decision algorithm based on this notion is introduced.

Введение

Принятие решений в реальных сложных системах связано с конфликтными ситуациями и неопределенностью, которая вызвана не только отсутствием информации о выборе стратегий противоборствующей стороной, но и невозможностью априорного вычисления выигрыша для некоторых ситуаций. Будем рассматривать класс антагонистических игр, в котором стратегии двух игроков состоят в указании совокупности некоторых действий из заданного множества, выполнение которых зависит от игрока. Такие стратегии являются булевыми векторами (наборами), состоящими из единиц и нулей. Вещественная функция выигрыша предполагается априорно заданной только для некоторого подмножества ситуаций. Такие модели будем называть играми с булевыми стратегиями [1, 2, 3, 5].

Изучение такого класса игр позволяет моделировать и оптимизировать поведение противоборствующих сторон логическими средствами, использовать методы, согласованные со структурой и математическим описанием вычислительных систем, баз знаний, моделей и алгоритмов интеллектуализированной обработки информации, основанных на эмпирической индукции.

1. Принятие решения при последовательном получении информации о выборах игроков

Управляющие переменные игроков $x_1, \ldots, x_m$ и $y_1, \ldots, y_n$ в игре $\Gamma(\Gamma_\Delta)$ [1] могут рассматриваться в приложениях как элементарные составляющие (ЭС) выбираемых стратегий в следующем смысле. Будем считать, что выбор значений элементарных составляющих $\widetilde{x}$ и $\widetilde{y}$ связан с выполнением некоторых действий $\widetilde{A} = A_1, \ldots, A_m$ и $\widetilde{B} = B_1, \ldots, B_n$. Если первый (второй) игрок устанавливает переменную $x_i(y_j)$ в единицу, то действие $A_i(B_j)$ обязательно выполняется, если устанавливает $x_i(y_j)$ в ноль, – действие $A_i(B_j)$ запрещается, если $x_i(y_j)$ не устанавливается ни в ноль, ни в единицу, никакой информации о действии $A_i(B_j)$ нет.

Определим многошаговую игру Γ^1_Δ со стратегиями последовательного выбора игроками по одной ЭС поочередно, полагая, что для Γ^1_Δ известна платежная LQ -матрица [4, 5].

Пусть функции логического описания классов (ЛОК) $j = \overline{1, l}$ имеют вид

$$f_j(\widetilde{x}, \widetilde{y}) = \bigvee_{1 \le s \le \mu_j} L_s^{(j)} \& Q_s^{(j)}$$

и «ход» делает первый игрок. Он устанавливает переменную x_i в единицу или ноль или пропускает ход, отказываясь от выбора. Если $x_i := \sigma$, $\sigma \in \{0, 1\}$, то подстановка вместо x_i значения σ в ЛОК $f_j(\widetilde{x}, \widetilde{y})$ приведет к упрощению соответствующей ДНФ и получению выражения $f_j(x_1, \ldots, x_{i-1}, \sigma, x_{i+1}, \ldots, x_m, y_1, \ldots, y_n)$ от на единицу меньшего числа переменных.

Рассмотрим пример, где ЛОК имеют следующий вид:

$$f_1(\widetilde{x}, \widetilde{y}) = \overline{x}_1 \overline{y}_1, f_2(\widetilde{x}, \widetilde{y}) = x_1 \overline{y}_1, f_3(\widetilde{x}, \widetilde{y}) = y_1;$$

Предположим, $x_1 := 1$; после подстановки этого значения в выражения $f_1(\cdot) = 0; f_2(\cdot) = \overline{y}_1; f_3(\cdot) = y_1$ определяем новый LQ -набор:

$$L_1^{(1)} = \varnothing; Q_1^{(1)} = \varnothing; L_1^{(2)} = \varnothing; Q_1^{(2)} = \overline{y}_1; Q_1^{(3)} = y_1$$

LQ -матрица редуцируется до строки вида показанного на рис. 1.

$$\begin{array}{cc} \overline{y}_1 & y_1 \\ \hline \multicolumn{1}{|c|}{\overline{h}_2} & \multicolumn{1}{c|}{\overline{h}_3} \\ \hline \end{array}$$

Рис. 1. Редуцированная матрица

Что соответсвует вычеркиванию из исходной LQ -матрицы строки $\overline{x}_1$ (рис. 2).

$$\begin{array}{c|c|c|} & \overline{y}_1 & y_1 \\ \cline{2-3} \cancel{\overline{x}_1} & \cancel{\overline{h}_1} & \cancel{\overline{h}_3} \\ \cline{2-3} x_1 & \overline{h}_2 & \overline{h}_3 \\ \cline{2-3} \end{array}$$

Рис. 2. Вычеркивание из исходной LQ -матрицы строки $\overline{x}_1$.

Следующим ходом второй игрок, выбрав $y_1 := 0$, зафиксирует решение игры в седловой точке.

Определение 1. Операцией удаления элементарной составляющей $O^\sigma_{\text{ЭС}}$ называется подстановка значения какой-либо одной переменной во все ЛОК $f_j(\widetilde{x}, \widetilde{y}), j = \overline{1, l}$.

При выполнении операции $O^g_{\text{ЭС}}$ некоторые из LQ конъюнкций могут обратиться в ноль, что соответствует их вычеркиванию из исходного LQ -списка, а некоторые, при обращении соответствующего литерала в единицу, уменьшатся по рангу на единицу.

Поскольку операция $O^g_{\text{ЭС}}$ соответствует подстановке значения только одной переменной, то эта переменная будет принадлежать либо множеству $\{x_1, \ldots, x_m\}$, либо множеству $\{y_1, \ldots, y_n\}$. Учитывая, что в ЛОК входят коньюнкции, вообще говоря, из обоих этих множеств, применение операции может привести к следующим исходам, представленным в таблице 1.

Таблица 1. Выполнение операции $O^g_{\text{ЭС}}$

№пп	Уточнение операции	Результат выполнения операции $O^g_{\text{ЭС}}$
1	$x_i := 0$ ($x_i := 1$)	Конъюнкции множества L, в которых переменная x_i содержится в положительном (отрицательном) литерале, обращаются в ноль. Конъюнкции множества Q, входящие в ЛОК в виде логического сомножителя с обратившейся в ноль конъюнкцией множества L также обращаются в ноль. Из конъюнкций множества L, в которых переменная x_i содержится в отрицательном(положительном) литерале указанная переменная вычеркивается. Ранг этой конъюнкции уменьшается на единицу. Если ранг становится нулевым, то указанная конъюнкция удаляется из L -списка, а конъюнкции Q -списка остаются без изменений.
2	$y_j := 0$ ($y_j := 1$)	Результат аналогичен п.1 с переменой местами в описании списков L и Q исходных конъюнкций LQ -игры.

Рассмотрение возможных результатов выполнения операции $O^g_{\text{ЭС}}$ делает очевидным следующее утверждение

Теорема 1. *Выполнение операции $O^g_{\text{ЭС}}$ равносильно вычеркиванию некоторых строк или столбцов исходной матрицы LQ -игры и, возможно, сокращению на единицу ранга некоторых конъюнкций.*

После выполнения операции $O^g_{\text{ЭС}}$ список LQ конъюнкций модифицируется, как указано в таблице; будем называть новый список $LQ(O^g_{\text{ЭС}})$ -списком, и соответствующую ему матричную игру – $LQ(O^g_{\text{ЭС}})$ -игрой.

Определение 2. Оптимальным шагом первого(второго)игрока называется такое выполнение операции $O^{g}_{\text{ЭС}}$, что в полученной $LQ(O^{g}_{\text{ЭС}})$ матричной игре размерности $p \times q$ значение максимина $\max\limits_{1\le i\le p}\min\limits_{1\le j\le q}\overline{h}_{ij}$ было наибольшим (значение минимакса $\min\limits_{1\le j\le q}\max\limits_{1\le i\le p}\overline{h}_{ij}$ было наименьшим).

Очевидно, что оптимальный шаг обеспечивает игроку выбор наилучшего гарантированного результата в $LQ(O^{g}_{\text{ЭС}})-$ игре, формируемой после преобразования платежной матрицы в соответствии с теоремой 1.

2. Условно контролирующие стратегии и последовательный выбор решений

Определение 3. Конъюнкция K первого(второго) игрока в $LQ(O^{g}_{\text{ЭС}})$ игре называется условно-контролирующей, если в матрице этой $LQ(O^{g}_{\text{ЭС}})$ игры все значения в соответствующей этой конъюнкции K строке(столбцу) одинаковы и равны $h(K)$; величина $h(K)$ называется ценой условно-контролирующей конъюнкции.

В LQ-игре на рис.2 второй игрок имеет условно контролирующую конъюнкцию y_1.

Если на очередном шаге (выполнение операции) $O^{g}_{\text{ЭС}}$ в игре $LQ(O^{g}_{\text{ЭС}})$ для имеющего ход первого (второго) игрока существуют условно контролирующие конъюнкции, то они легко определяются путем просмотра строк и столбцов платежной матрицы.

Определение 4. Условно-контролирующая конъюнкция первого(второго) игрока с наибольшей(наименьшей) ценой называется *наилучшей.*

Перед выполнением очередного шага игры Γ^1_Δ игрок определяет свою наилучшую условно-контролирующую стратегию.

Теорема 2.

1. *Если в исходной LQ -игре условно контролирующую конъюнкцию имеет только первый (второй) игрок, то в игре Γ^1_Δ он может обеспечить себе выигрыш, равный цене его наилучшей условно контролирующей конъюнкции.*
2. *Если в исходной LQ -игре оба игрока имеют условно контролирующие конъюнкции и L^* и Q^* – наилучшие из них, то в игре Γ^1_Δ игроки могут обеспечить себе выигрыш, равный $h(L^*, Q^*)$; при этом пара (L^*, Q^*) является седловой точкой исходной $LQ-$ игры.*

Замечание 1. Если элементарные составляющие из наилучшей условно контролирующей конъюнкции исчерпаны, то игрок может пропускать ходы, и при этом результат игры Γ^1_Δ не изменится.

Следствие 1. *Если исходная LQ -игра имеет седловую точку и цену h^* и первый (второй) игрок имеет наилучшую условно контролирующую конъюнкцию с ценой $h_{\text{ук}} \ge h^*(h_{\text{ук}} \le h^*)$, то он обеспечивает себе выигрыш не меньший (не больший) значения h^* цены исходной LQ -игры.*

Алгоритм A^1_s решения многошаговой игры Γ^1_Δ.

1. Если исходная LQ -игра имеет седловую точку, найти ее и определить цену игры h^*.
2. Если существуют наилучшие условно контролирующие конъюнкции первого (второго) игрока с ценой $h_{\text{УК}} \geq h^*(h_{\text{УК}} \leq h^*)$, то выполнять шаги, описанные в доказательстве теоремы 2.
3. Если седловой точки в исходной LQ -игре нет, но условно контролирующие конъюнкции есть у одного игрока с наилучшей ценой $h_{\text{УК}}$ принять решение о достаточности выигрыша $h_{\text{УК}}$ и реализовать стратегию, соответствующую наилучшей условно контролирующей конъюнкции.
4. Если условно контролирующих конъюнкций нет у обоих игроков, то выполнять оптимальные шаги (определение 2) следующим образом. Поскольку каждый шаг, согласно теореме 1, приводит к вычеркиванию некоторых строк или столбцов платежной матрицы, то рассмотреть все возможные шаги (их число будет не более чем удвоенная сумма рангов L или Q конъюнкций). При выборе каждого шага оценивается гарантированный выигрыш и реализуется оптимальный шаг.

Заключение

Практическое значение полученных результатов состоит в возможности применения разработанных в ней алгоритмов решения многошаговой игры с булевыми стратегиями и принятия решений в рамках теоретико-игрового подхода для построения интеллектуализированных информационных систем, выявления скрытых структурных закономерностей в данных, построения логических описаний классов объектов в виде дизъюнктивных нормальных форм.

Список литературы

1. *Блыщик В.Ф.* Решение игр с булевыми стратегиями и неполной информацией на основе синтеза ДНФ// Искусственный интеллект. – 2000. – № 2. – С.9-12.
2. *Блыщик В.Ф.* Принятие и визуализация решений при последовательном выборе действий в матричной игре с булевыми стратегиями// International Conference. Problems of Decision Making under Uncertainties (PDMU-2003).Abstracts. – Alushta (Ukraine). 2003. – С.74-76.
3. *Блыщик В.Ф.* Игры двух лиц с булевыми стратегиями и множеством переменных перехвата// Таврический вестник информатики и математики. – 2004. – № 1. – С.40-47.
4. *Блыщик В.Ф.* Алгоритм построения классов значений платежной функции по прецедентной начальной информации// Искусственный интеллект. – 2006. – № 2. – С.10-13.
5. *Донской В.И., Башта А.И.* Дискретные модели принятия решений при неполной информации. – Симферополь: Таврия, 1992. – 166 с.

Статья поступила в редакцию 24.12.2008

УДК 519.962.22

МІНІМАКСНІ ОЦІНКИ РОЗВ'ЯЗКІВ РІВНЯНЬ З ЛІНІЙНИМ НЕОБМЕЖЕНИМ ОПЕРАТОРОМ У ГІЛЬБЕРТОВОМУ ПРОСТОРІ

© **Жук С.М.**

Київський національний університет ім. Т.Г.Шевченка,
факультет кібернетики
пр-т академіка Глушкова-2, корпус 6, м.Київ 03680, Україна
e-mail: *Serhiy.Zhuk@gmail.com*

Abstract. This paper describes an approach to minimax estimation of the solution of a linear equation with closed dense defined mapping in Hilbert space. A class of the linear minimax mean-square estimations is considered. A necessary an sufficient condition for the minimax mean-square error to be finite is introduced. A representation of minimax estimations are obtained in the case of ellipsoidal constraints.

Вступ

Ця робота є продовженням [1]. Тут розвивається техніка, застосована у [1] до апріорного оцінювання розв'язків одновимірних крайових задач, на абстрактні лінійні операторні рівняння у гільбертовому просторі.

1. Постановка задачі

Нехай спостерігається реалізація вектора y з гільбертового простору $\mathscr{Y}$ вигляду

$$y = H\varphi + \eta, \tag{1}$$

де η є випадковим вектором зі значеннями у $\mathscr{Y}$, нульовим середнім та кореляційним оператором R_η, H є елементом банахового простору[1] $\mathscr{L}(\mathscr{H}, \mathscr{Y})$, а $\varphi \in \mathscr{H}$ є розв'язком лінійного операторного рівняння

$$L\varphi = f \tag{2}$$

Ми будемо вважати L замкненим оператором, що відображає скрізь щільну підмножину $\mathscr{D}(L)$ гільбертового простору $\mathscr{H}$ у гільбертів простір $\mathscr{F}$.

Припустимо, що вектор $f \in \mathscr{F}$ є деяким, наперед невідомим, елементом множини $\mathscr{G} \subset \mathscr{F}$. Нехай також розподіл випадкового вектора η задовольняє умову $R_\eta \subset \mathscr{R}$, де $\mathscr{R}$ є задана підмножина $\mathscr{L}(\mathscr{Y}, \mathscr{Y})$.

Наша мета полягає в тому, щоб

(ОЗ) для заданого $\ell \in \mathscr{H}$ оцінити скалярний добуток (ℓ, φ), користуючись інформацією про φ: структура рівняння (2), визначена властивостями оператора L; спостереження за φ, задані у вигляді (1); структура множин $\mathscr{G}, \mathscr{R}$.

(ОЗ) є різновидом обернених задач для лінійних операторних рівнянь у гільбертовому просторі в умовах невизначеності. Надалі називатимемо (ОЗ) задачею оцінювання.

[1]Символом $\mathscr{L}(\mathscr{H}, \mathscr{Y})$ позначимо банаховий простір усіх лінійних обмежених операторів з $\mathscr{H}$ у $\mathscr{Y}$

Зауваження 1. Помітимо, що реалізація y визначається не лише кореляційним оператором η, виглядом H та f. У загальному випадку може йтися про невизначеність, породжену неоднозначністю оператора L: кожному $f \in R(L) \cap \mathscr{G}$ ми можемо зіставити множину $\varphi_0 \oplus N(L)$, де $N(L) = \{\varphi \in \mathscr{D}(L) : L\varphi = 0\}$. Таким чином $y = H(\varphi_0 + \varphi) + \eta$ для фіксованих f та реалізації η.

Тут ми не припускаємо, що оператори L, H мають обмежені обернені, відтак незначні відхилення у правій частині (2) та вимірах (1) можуть спричинити необмежено велику похибку оцінювання. Зважаючи на це, а також на цілий ряд невизначеностей, згаданих вище, ми побудуємо операцію оцінювання на основі мінімаксного підходу. Це дозволить запропонувати *гарантовану похибку оцінювання*, яка характеризує найбільше відхилення оцінки від реального значення і для досить[2] широкого класу операторів L, H буде скінченною.

Означення 1. Афінний функціонал $(\hat{u}, \cdot) + \hat{c}$ назвемо *апріорною мінімаксною середньоквадратичною оцінкою* скалярного добутку (ℓ, φ), якщо

$$\sup_{\varphi \in L^{-1}(G), R_\eta \in \mathscr{R}} M((\ell, \varphi) - (\hat{u}, y) - \hat{c})^2 = \inf_{u,c} \sup_{\varphi \in L^{-1}(G), R_\eta \in \mathscr{R}} M((\ell, \varphi) - (u, y) - c)^2 \tag{3}$$

Вираз

$$\hat{\sigma}(\ell) = \sup_{\varphi \in L^{-1}(G), R_\eta \in \mathscr{R}} M((\ell, \varphi) - (\hat{u}, y) - \hat{c})^2$$

назвемо *мінімаксною середньоквадратичною похибкою* у напрямку ℓ.

Нижче приведено теореми про представлення мінімаксніх апріорних та апостеріорних оцінки для того випадку, коли множини обмежень $\mathscr{G}, R$ є опуклими обмеженими і замкненими у відповідних просторах.

2. Мінімаксні оцінки для випадку слабко компактних опуклих обмежень

У подальшому нам знадобляться такі позначення. Введемо множини

$$\mathscr{U}_\ell = \{u \in \mathrm{Y} : L^* z = \ell - H^* u\}, D = \{\ell \in \mathscr{H} : U(\ell) \neq \varnothing\}$$

Тут символом L^* позначено оператор, спряжений до L, який після ототожнення гільбертових просторів $\mathscr{H}, \mathscr{F}$ з їхніми спряженими, діє з F у H. Існування єдиного спряженого L^* забезпечується [4] щільною визначеністю L. Через $\delta(\mathscr{G}, \cdot)$ позначимо індикатор множини $\mathscr{G}$, тобто $\delta(\mathscr{G}, f) = 0, f \in \mathscr{G}$ і $+\infty$ інакше. Покладемо

$$(\delta L)^*(x) = \sup_{\varphi \in \mathscr{D}(L)} \{(\varphi, x) - \delta(\mathscr{G}, L\varphi)\}$$

Функціонал $(\delta L)^*$ є перетворення Юнга-Фенхеля функціоналу (δL), $(\delta L)(x) = \delta(\mathscr{G}, Lx)$, якщо $x \in \mathscr{D}(L)$ і $+\infty$ інакше. Позначимо

$$\mathrm{dom}(\delta L)^* = \{x \in \mathscr{H} : (\delta L)^*(x) < \infty\}$$

[2] Якщо $R(L), H(N(L))$ є замкненими лінійними многовидами, то гарантована похибка оцінювання скінченна у довільному напрямку $\ell \in R(L^*) + R(H^*)$, зокрема для $N(L) \cap N(H) = \{0\}$ гарантована похибка завжди скінченна.

ефективну множину функціоналу $(\delta L)^*$ і нехай

$$(L^*c)(x) = \inf\{c(\mathscr{G}, z)|L^*z = x\}, c(\mathscr{G}, z) = \sup\{(z, f), f \in \mathscr{G}\}$$

Нехай $\mathrm{cl}(f) = f^{**}$. За теоремою Фенхеля-Моро f^{**} збігається з напівнеперервною знизу регуляризацію функціоналу f, якщо f є власним.

Наступна лема лежить в основі доведення основних тверджень про існування, єдиність та представлення мінімаксних оцінок.

Лема 1. Нехай $\mathscr{G}$ є опуклою обмеженою замкненою підмножиною $\mathscr{F}$, L є лінійним щільно визначеним замкненим оператором з $\mathscr{H}$ у $\mathscr{F}$.
Тоді

$$(L^*c)^* = (\delta L), (L^*c)^{**} = (\delta L)^*, R(L^*) \subset \mathrm{dom}(\delta L)^* \subset \overline{R(L^*)} \tag{4}$$

Якщо внутрішність $\mathscr{G}$ має спільні точки з $R(L)$, то $\mathrm{dom}(\delta L)^* = \mathrm{dom}(L^*c) = R(L^*)$, функціонал (L^*c) є власним, $L^*c = (L^*c)^{**}$ і

$$(L^*c)(x) = c(\mathscr{G}, z_0) = \inf\{c(\mathscr{G}, z)|L^*z = x\}, x \in R(L^*)$$

Доведення. Нехай $p \in \mathscr{D}(L)$. Обчислимо[3]

$$(L^*c)^*(p) = \sup_{x \in R(L^*)} \{(p, x) - \inf\{c(\mathscr{G}, z)|L^*z = x\}\} =$$

$$\sup_{x \in R(L^*)} \sup_{z \in L^{*-1}(x)} \{(p, x) - c(\mathscr{G}, z)\} = \sup_{z \in \mathscr{D}(L^*)} \{(p, L^*z) - c(\mathscr{G}, z)\} =$$

$$\sup_{z \in \mathscr{F}} \{(Lp, z) - c(\mathscr{G}, z)\} = c^*(\mathscr{G}, \cdot)(Lp) = \delta(\mathscr{G}, Lp)$$

Розглянемо випадок $p \notin \mathscr{D}(L)$. За означенням спряженого до необмеженого лінійного оператора [4, с.39] лінійний функціонал $z \mapsto p(z) = (p, L^*z)$ є необмеженим. Це означає, що знайдеться послідовність $\{z_n\}$ така, що $\|z_n\| \le 1, z_n \in \mathscr{D}(L^*)$ і $p(z_n) \to +\infty$. З іншого боку опорна функція $c(\mathscr{G}, \cdot)$ обмеженої опуклої множини обмежена в околі довільної точки $z \in \mathscr{F}$ і тому неперервна [3, с.21]. Але тоді $\sup_n c(\mathscr{G}, z_n) = M < +\infty$ і

$$(L^*c)^*(p) = \sup_{z \in \mathscr{D}(L^*)} \{(p, L^*z) - c(\mathscr{G}, z)\} \ge \sup_n \{p(z_n) - M\} = +\infty$$

Таким чином, $(L^*c)^*(p) = +\infty$. З іншого боку $(\delta L)(p) = +\infty$ за означенням. Ми показали, що $(L^*c)^*(p) = (\delta L)(p)$ для всіх p, звідки $(L^*c)^{**} = (\delta L)^*$. Під час доведення ми не використовували інформацію про те, чи мають множини $R(L), \mathscr{G}$ спільні точки чи ні. Тому $(L^*c)^{**} = (\delta L)^*$ виконується для випадку $R(L) \cap \mathscr{G} = \varnothing$.

Нехай $x \notin N(L)^\perp$ і $Lp \in G$ для деякого $p \in \mathscr{D}(L)$. Знайдеться $p_0 \in N(L)$ таке, що $n(p_0, x) > 0, n \in \mathbb{N}$. Але тоді

$$(\delta L)^*(x) = \sup_{q \in \mathscr{D}(L)} \{(q, x) - \delta(G, Lq)\} \ge \sup_{n \in \mathbb{N}} \{(p, x) + n(p_0, x)\} = +\infty$$

Тому $\mathrm{dom}(\delta L)^* \subset N(L)^\perp = \overline{R(L^*)}$.

[3] $p \in \mathscr{D}(L)$ тому лінійний функціонал $z \mapsto p(z) = (p, L^*z)$ є обмеженим, відтак може бути розповсюджений на весь простір $\mathscr{F}$ за неперервністю, звідки $\sup_{z \in \mathscr{D}(L^*)}\{(Lp, z) - c(\mathscr{G}, z)\} = \sup_{z \in \mathscr{F}}\{(Lp, z) - c(\mathscr{G}, z)\}$.

З іншого боку, якщо $x = L^*z$, то

$$(\delta L)^*(x) = \sup_{q\in\mathscr{D}(L)}\{(Lq,x)-\delta(G,Lq)\} \leq \sup_{f\in\mathscr{F}}\{(f,x)-\delta(G,f)\} = c(G,x) < +\infty$$

в силу обмеженості G. Таким чином $R(L^*) \subset \mathrm{dom}(\delta L)^* \subset \overline{R(L^*)}$.

Нехай тепер $\mathrm{int}\mathscr{G} \cap R(L) \neq \varnothing$. Покажемо, що цього достатньо для $(L^*c) \leqslant (\delta L)^*$. Справді

$$x^* \in \mathrm{dom}(\delta L)^*, x \in \mathscr{D}(L) \Rightarrow (x^*,x)-(\delta L)^*(x^*) \leqslant \delta L(x) < +\infty$$

в силу нерівності Юнга-Фенхеля. Зафіксувавши $x^* \in \mathrm{dom}(\delta L)^*$, введемо множину

$$\mathscr{M}(x^*) = \{(z,\mu)|Lx = z, \mu = (x^*,x)-(\delta L)^*(x^*)\}$$

Помітимо, що

$$\mathscr{W} = \mathrm{int}\,\mathrm{epi}(\delta(\mathscr{G},\cdot)) = \mathrm{int}\mathscr{G}\times\{\mu\in\mathbb{R}^1:\mu>0\}\cap\mathscr{M}(x^*) = \varnothing$$

Дійсно, якщо $(z,\mu) \in \mathscr{W}\cap\mathscr{M}$, то

$$\delta(\mathscr{G},Lx) < \mu = (x^*,x)-(\delta L)^*(x^*), Lx = z$$

що суперечить нерівності Юнга-Фенхеля.
Отже, опуклі множини $\mathrm{epi}(\delta(\mathscr{G},\cdot),\mathscr{M}(x^*)$ можна розділити ненульовим лінійним неперервним функціоналом

$$\sup\{(z^0,z)+\beta_0\alpha|(z,\alpha)\in\mathscr{W}\} \leqslant \inf\{(z^0,z)+\beta_0\alpha|(z,\alpha)\in\mathscr{M}(x^*)\} \qquad (*)$$

Легко пересвідчитись, що $\beta_0 < 0$. Дійсно, якщо $\beta_0 > 0$, то sup у $(*)$ дорівнює $+\infty$. З іншого боку sup у $(*)$ завжди відмінна від $-\infty$, що ґарантує скінченність inf у $(*)$. Якщо $\beta_0 = 0$, то згідно $(*)$ $\mathscr{G}$ та $R(L)$ розділяються функціоналом $(z^0,\cdot)$, але тоді $\mathrm{int}\mathscr{G}\cap R(L) = \varnothing$.

За означенням $\mathscr{M}(x^*)$

$$-\infty < (\mathrm{c}(\mathscr{G},z^0) = \sup\{(z^0,z)-\beta_0\delta(\mathscr{G},z)\} \leqslant \inf\{(z^0,Lx)-|\beta_0|(x^*,x)\}+|\beta_0|(\delta L)^*(x^*)$$

звідки

$$-\infty < \inf_x\{(z^0,Lx)-\beta_0(x^*,x)\} \Rightarrow [-|\beta_0|x^*,z^0] \perp \{[x,Lx],x\in\mathscr{D}(L)\}$$

Тому, зважаючи на вигляд [4, с.40] ортогонального доповнення графіка L, дістанемо

$$z_0 \in \mathscr{D}(L^*), L^*z_0 = |\beta_0|x^* \Rightarrow (L^*c)(x^*) \leqslant c(\mathscr{G},\beta_0^{-1}z^0) \leqslant (\delta L)^*(x^*)$$

Ми показали, що на $\mathrm{dom}(\delta L)^*$ виконано $(L^*c) = (\delta L)^*$ і $\mathrm{dom}(\delta L)^* \subset R(L^*)$. За означенням $R(L^*) \subset \mathrm{dom}(L^*c)$. Раніше було доведено, що $R(L^*) \subset \mathrm{dom}(\delta L)^*$. Оскільки, взагалі кажучи, $(L^*c) \geqslant (L^*c)^{**} = (\delta L)^*$, то $\mathrm{dom}(\delta L)^* \subset \mathrm{dom}(L^*c)$. Отже

$$(L^*c) = (\delta L)^*, \mathrm{dom}(\delta L)^* = \mathrm{dom}(L^*c) = R(L^*)$$

За теоремою Фенхеля-Моро $(L^*c) = (L^*c)^{**} = (\delta L)^*$ тоді і лише тоді, коли (L^*c) має замкнений надграфік, що для власних опуклих функціоналів еквівалентно напівнеперервності знизу [5, с.178]. □

Наступна теорема дає загальний вигляд мінімаксної середньо-квадратичної оцінки та визначає критерій скінченності мінімаксної похибки оцінювання.

Теорема 1. Нехай $\mathscr{G}$ є опуклою, замкненою, обмеженою підмножиною $\mathscr{F}$, $\eta \in \{\eta : M(\eta,\eta) \le 1\}$. Гарантована похибка оцінювання має вигляд

$$\sigma(\ell,u) = \frac{1}{4}[\mathrm{cl}(L^*c)(\ell - H^*u) + \mathrm{cl}(L^*c)(-\ell + H^*u)]^2 + \sup_{R_\eta \in \mathscr{R}} (R_\eta u, u), \tag{5}$$

Для заданого $\ell \in \mathscr{H}$ мінімаксна похибка скінчена тоді і лише тоді коли

$$\ell - H^*u \in \mathrm{dom}(\delta L)^* \cap -\mathrm{dom}(\delta L)^*, R(L^*) \subset \mathrm{dom}(\delta L)^* \subset \overline{R(L^*)}$$

для деякого $u \in \mathscr{Y}$. Мінімаксна оцінка дається виразом

$$y \mapsto (\hat{u}, y) + \hat{c}, \hat{u} \in \mathrm{Arginf}_u \sigma(\ell,u), \hat{c} = \frac{1}{2}(\mathrm{cl}(L^*c)(\ell - H^*\hat{u}) - \mathrm{cl}(L^*c)(-\ell + H^*\hat{u}))$$

Доведення. Беручи до уваги рівність $M\xi^2 = M(\xi - M\xi)^2 + (M\xi)^2$ та (1), знайдемо

$$M((\ell,\varphi) - (u,y) - c)^2 = [(\ell - H^*u, \varphi) - c]^2 + M(u,\eta)^2$$

відтак

$$\sup_{\varphi \in L^{-1}(G), R_\eta \in \mathscr{R}} M((\ell,\varphi) - (u,y) - c)^2 = \sup_{\varphi \in L^{-1}(G)} [(\ell - H^*u, \varphi) - c]^2 + \sup_{R_\eta \in \mathscr{R}} (R_\eta u, u)$$

Легко показати, що у цьому випадку

$$\sup_{R_\eta \in \mathscr{R}} (R_\eta u, u) = (u,u)$$

Перетворимо перший доданок

$$\begin{aligned} \sup_{\varphi \in L^{-1}(G)} [(\ell - H^*u, \varphi) - c] = \frac{1}{2}((\delta L)^*(\ell - H^*u) + (\delta L)^*(-\ell + H^*u)) + \\ |c - \frac{1}{2}((\delta L)^*(\ell - H^*u) - (\delta L)^*(-\ell + H^*u))| \end{aligned} \tag{6}$$

Зважаючи на формулу (6) виводимо, що для заданих ℓ, u, c

$$\sup_{\varphi \in L^{-1}(G)} [(\ell - H^*u, \varphi) - c] < +\infty \Leftrightarrow \ell - H^*u \in \mathrm{dom}(\delta L)^* \cap -\mathrm{dom}(\delta L)^*,$$

Множина $\mathrm{dom}(\delta L)^*$ є опуклим конусом з вершиною в нулі, відтак $\mathrm{dom}(\delta L)^* \cap -\mathrm{dom}(\delta L)^*$ є найбільшим лінійним многовидом, що міститься в $\mathrm{dom}(\delta L)^*$. Якщо покласти $c = \frac{1}{2}((\delta L)^*(\ell - H^*u) - (\delta L)^*(-\ell + H^*u))$, то з (6) та леми 1 дістанемо вираз для $\sigma(\ell,u)$.

Якщо $\ell - H^*u \in \mathrm{dom}(\delta L)^* \cap -\mathrm{dom}(\delta L)^*$, то в силу строгої опуклості, коерцитивності та напівнеперервності знизу функціоналу $u \mapsto \sigma(\ell,u)$ існує єдина мінімаксна оцінка $\hat{u}$. □

Наслідок 1. Нехай

$$\mathrm{int}\mathscr{G} \cap R(L) \neq \varnothing, \eta \in \{\eta : M(\eta,\eta) \le 1\}$$

Тоді гарантована похибка оцінювання скінченна для $\ell - H^*u \in R(L^*)$ і лише для них, існує єдина мінімаксна середньоквадратична оцінка $\hat{u}$, що визначається з умови $\sigma(\ell, u) \to \min_u$, де

$$\sigma(\ell, u) = \frac{1}{4}[c(\mathscr{G}, z) + c(\mathscr{G}, -z)]^2 + (u, u)L^*z = \ell - H^*u \tag{7}$$

Доведення. Згідно теореми 1 для заданого $\ell \in \mathscr{H}$ мінімаксна похибка скінчена тоді і лише тоді коли

$$\ell - H^*u \in \mathrm{dom}(\delta L)^* \cap -\mathrm{dom}(\delta L)^*$$

Оскільки $0 \in \mathscr{G} \cap R(L)$, то ми знаходимося в умовах леми 1, звідки $\mathrm{dom}(\delta L)^* = R(L^*)$ і

$$\mathrm{cl}(L^*c)(\ell - H^*u) = (L^*c)(\ell - H^*u) = c(\mathscr{G}, z), L^*z = \ell - H^*u$$

Легко показати, що у цьому випадку

$$\sup_{R_\eta \in \mathscr{R}} (R_\eta u, u) = (u, u)$$

Тоді з (5) виводимо (7). Легко збагнути, що функціонал $u \mapsto \sigma(\ell, u)$ є слабко напівнеперевним строго опуклим та коерцитивним. Справді, слабка напівнеперервність слідує з опуклості та замкнености (лема 1) функціоналу

$$w \mapsto \min\{c(\mathscr{G}, \psi), L^*\psi = w\}$$

Решта властивостей очевидні. Тому на замкненій опуклій множині $\mathscr{U}_\ell$ функціонал $u \mapsto \sigma(\ell, u)$ досягає свого мінімуму в єдиній точці $\hat{u}$. □

Для еліпсоїдів у гільбертовому просторі мінімаксна апріорна середньоквадратична оцінка може бути зображена у вигляді розв'язків системи лінійних операторних рівнянь.

Теорема 2. Нехай

$$\mathscr{G} = \{f : \|f\|^2 \leq 1\}, \eta \in \{\eta : M(\eta, \eta) \leq 1\},$$

множина $R(T) = \{[Lx, Hx], x \in \mathscr{D}(L)\}$ є замкненою. Для $\ell \in R(L^*) + R(H^*)$ і лише для них єдина мінімаксна оцінка $\hat{u}$ може бути подана у вигляді $\hat{u} = H\hat{p}$, де $\hat{p}$ є довільним розв'язком системи

$$\begin{aligned} L^*\hat{z} &= \ell - H^*H\hat{p}, \\ L\hat{p} &= \hat{z} \end{aligned} \tag{8}$$

мінімаксна похибка оцінювання зображується як

$$\sigma(\ell, \hat{u}) = (\ell, \hat{p}) \tag{9}$$

Доведення. Помітимо, що умови наслідку 1 виконано, відтак існує єдина мінімаксна середньоквадратична оцінка $\hat{u}$, що визначається з умови

$$\sigma(\ell, u) = (z, z) + (u, u), L^*z = \ell - H^*u, u \in \mathscr{U}_\ell \tag{10}$$

Можна показати, що оцінка $\hat{u}$, водночас, може бути знайдена як розв'язок задачі проектування

$$\|[z, u]\|^2 \to \inf, T^*[z, u] = \ell \tag{11}$$

Остання має єдиний розв'язок в силу замкненості T, що лежить у множині значень оператора T. Отже, $[\hat{u}, \hat{z}] = Tx$ і, водночас, $T^*[\hat{u}, \hat{z}] = \ell$, звідки, пригадуючи означення T, знаходимо

$$Lx = \hat{z}, Hx = \hat{u}, L^*\hat{z} + H^*\hat{u} = \ell,$$

що і завершує доведення. □

Зауваження 2. Можна показати, що вектор[4] $[\hat{u}, \hat{z}]$, які знаходяться як розв'язок системи (8), є найменшим по нормі розв'язком рівняння $T^*[z, u] = \ell$. Ми можемо ввести оператор

$$\ell \mapsto T^{*+}\ell = [\hat{u}, \hat{z}], \ell \in R(L^*) + R(H^*)$$

З іншого боку вектори $[\hat{\varphi}, \hat{q}]$, що знаходяться як розв'язок системи

$$\begin{aligned} &L^*\hat{q} = H^*(y - H\hat{\varphi}), \\ &L\hat{\varphi} = \hat{q} \end{aligned} \tag{12}$$

дають розв'язок задачі проектування

$$(Lx, Lx) + (y - Hx, y - Hx) \to \min_x,$$

причому $\hat{\varphi}$ – розв'язок цієї задачі з найменшою нормою. Зобразимо цю відповідність у вигляді

$$T^+[0, y] = \hat{\varphi}, y \in \mathscr{Y}$$

Запишемо $(T^{*+}\ell, [y, 0]) = (\hat{u}, y) = (\ell, \hat{\varphi}) = (\ell, T^+[0, y])$. Якщо $R(T^*)$ збігається з усім простором прибуття T^*, то попередня рівність справедлива для довільного ℓ, відтак вектор $T^+[0, y]$ можна прийняти за оцінку x. У загальному випадку $T^+[0, y]$ дає оцінку проекції x на $R(T^*)$.

Висновки

У цій роботі проілюстровано застосування підходу, запропонованого у [2], до проблеми мінімаксного оцінювання розв'язків лінійних операторних рівнянь у випадку слабко компактних опуклих обмежень, виділено множину функціоналів $\mathscr{F}$, для елементів якої і лише для них мінімаксна оцінка існує і єдина, одержано представлення оцінки за допомогою розв'язків системи лінійних операторних рівнянь для випадку квадратичних обмежень.

Список літератури

1. *Деміденко С.В.,Жук С.М.,Наконечний О.Г.* До проблеми мінімаксного оцінювання розв'язків одновимірних крайових задач // Тавр.вісник інформ. та матем.— 2007.— N1— С.7-24.
2. *Жук С.М.* Задачі мінімаксного спостереження для лінійних дескрипторних систем: Автореферат дис.канд-та фіз.-мат. наук / Київ, 2006 – 19 с.
3. *Екланд И.* Выпуклый анализ и вариационные проблемы:Пер.з англ.— М.:Наука, 1979.– 396 с.
4. *Лянце В.Э.,Сторож О.Г.* Методы теории неограниченных операторов. — Киев:Наук.думка, 1983.— 212 с. н
5. *Иоффе А.Д.,Тихомиров В.М.* Теория экстремальных задач. — М.:Наука, 1974.— 477 с.

Статья поступила в редакцию 12.02.2008

[4]Символом $[x, y]$ будемо позначати вектор, що є елементом декартового добутку $H_1 \times H_2$.

УДК 004.9311

ОБОБЩЕНИЕ ИНФОРМАЦИОННОГО КРИТЕРИЯ АКАИКЕ ДЛЯ ВЫБОРА ЗНАЧЕНИЙ НЕПРЕРЫВНЫХ ПАРАМЕТРОВ В МОДЕЛЯХ ДАННЫХ

© Ежова Е.О., Моттль В.В., Красоткина О.В.

Московский Физико-Технический институт
Институтский пер, 9, г. Долгопрудный, 141700, Россия
e-mail: *lena-ezhova@rambler.ru*

ВЦ РАН
ул Вавилова, 40, г. Москва, 117967, Россия
e-mail: *vmottl@yandex.ru*

ТулГУ
пр-т Ленина, 92, г. Тула, 300600, Россия
e-mail: *krasotkina@uic.tula.ru*

Abstract. The crucial restriction of the Akaike Information Criterion (AIC) as means of adjusting a model to the given data set within a succession of nested parametric model classes is the assumption that the classes are rigidly defined by the growing dimension of an unknown vector parameter. We extend the Kullback information maximization principle underlying the classical AIC onto a wider class of data models in which the dimension of the parameter is fixed, but the freedom of its values is softly constrained by a class of continuously nested a priori probability distributions We illustrate the proposed continuous generalization of AIC by its application to the problem of time-varying regression estimation which implies the inevitable necessity to choose the time-variability of regression coefficients treated a nonstationary model of the given signal.

Введение

Широко используемый в современном анализе данных информационный критерий Акаике (AIC) [1] *является простым и эффективным способом выбора наиболее адекватного класса модели из упорядоченного дискретного множества вложенных классов моделей.*

В классической постановке критерия обычно рассматривается выборка $\mathbf{y} = (y_j, j = 1, \ldots, N)$ независимых случайных величин с неизвестной плотностью распределения $\varphi^*(y)$, принадлежащей некоторому параметрическому семейству $\varphi(y \mid \mathbf{c})$, $\mathbf{c} \in \mathbb{R}^m$. Часто размерность вектора параметров m оказывается очень большой и существенно превосходит размер обучающей выборки N, что делает бессмысленным применение для оценивания вектора параметров $\mathbf{c}$ принципа максимального правдоподобия.

$$\hat{\mathbf{c}}(\mathbf{y}) = \arg\max \ln \Phi(\mathbf{y} \,|\, \mathbf{c}), \ln \Phi(\mathbf{y} \,|\, \mathbf{c}) = \sum_{j=1}^{N} \ln \varphi(y_j \,|\, \mathbf{c}). \tag{1}$$

Если же предположить, что элементы вектора $\mathbf{c}$ обладают естественной упорядоченностью по степени значимости и при этом $c_i = 0$, $n < i \leq m$:

$$\mathbf{c} = (\mathbf{c}_n, \mathbf{c}_{m-n}), \mathbf{c}_n \in \mathbb{R}^n, \mathbf{c}_{m-n} = \mathbf{0} \in \mathbb{R}^{m-n}. \tag{2}$$

то это позволит нам рассмотреть параметрическое семейство $\Phi(\mathbf{y}\,|\,\mathbf{c})$ как последовательность вложенных классов моделей $\Phi\big(\mathbf{y}\mid\mathbf{c}=(\mathbf{c}_n\,|\,\mathbf{0})\big)$, $\mathbb{R}^{n_{min}}\subset\cdots\subset\mathbb{R}^{n_{max}}$.

Критерий AIC в классической постановке является способом оценивания подходящей размерности вектора параметров, как меры сложности модели $\hat{n}=\arg\max_n\Big[\ln\Phi\big(\mathbf{y}\mid\mathbf{c}_n(\mathbf{y})\big)-n\Big]$. Однако это формула получена в предположении, что гессиан $\bigtriangledown^2_{\mathbf{c}_n\mathbf{c}_n}\ln\Phi(\mathbf{y}\,|\,\mathbf{c}_n,\mathbf{0})$ в точке максимального правдоподобия имеет полный ранг, а значит и оценка $\hat{\mathbf{c}}_n(\mathbf{y})$ — единственная. В более общем случае заменим штраф n на ранг матрицы

$$\hat{n}=\arg\max_n\Big\{\ln\Phi\big(\mathbf{y}\,|\,\mathbf{c}_n(\mathbf{y}),\mathbf{0}\big)-\mathrm{rank}\big[\bigtriangledown^2_{\mathbf{c}_n\mathbf{c}_n}\ln\Phi(\mathbf{y}\,|\,\mathbf{c}_n,\mathbf{0})\big]\Big\} \qquad (3)$$

В основе классического AIC лежит принцип максимизации информации по Кульбаку между моделью плотности распределения и настоящей гипотетической плотностью распределения.

$$n^*=\arg\max_n\int\big[\ln\Phi(\mathbf{y}\,|\,n,\mathbf{c}^*_n)\big]\,\Phi^*(\mathbf{y})d\mathbf{y} \qquad (4)$$

есть желаемая размерность в предположении, что $\Phi^*(\mathbf{y})=\Phi(\mathbf{y}\mid\mathbf{c}^*_{n^*})$ с некоторым значением $(\mathbf{c}^*_{n^*}\mathbf{0})$, вырезанного из неизвестного $\mathbf{c}^*=(c^*_1,\ldots,c^*_m)$

Одним из первых применений AIC было моделирование нестационарного сигнала на дискретной временной оси, разделенной на на неизвестное количество n интервальных блоков, и проверка локальной стационарности модели авторегрессии с фиксированным порядком k на каждом их них [2].

Со времен публикации первой статьи Акаике было предложено много модификаций этого критерия [3, 4, 5, 6]. *Среди них Байесовский информационный критерий (BIC)* [3] *нашел более широкое применение. Однако все они были нацелены на выбор размерности вектора параметров для случая известной упорядоченности его элементов по степени значимости.*

В данной работе предлагается совершенно новое обобщение критерия Акаике, которое было вызвано необходимостью анализа нестационарного сигнала $(\mathbf{y},\mathbf{x})=\big((y_t,\mathbf{x}_t),t=1,\ldots,N\big)$,*регрессионная модель которого*

$$y_t=\mathbf{c}_t^T\mathbf{x}_t+\eta_t,\mathbf{c}_t,\mathbf{x}_t\in\mathbb{R}^k,\eta_t\sim N(\eta_t\,|\,0,\delta),E(\eta_t,\eta_s)=0, \qquad (5)$$

меняется на интервале наблюдения. Очевидно, что при этом размерность вектора параметров в семействе условных плотностей распределения $\Phi(\mathbf{y}\mid\mathbf{x},\mathbf{c})$ оказывается фиксированной $\mathbf{c}=(\mathbf{c}_1\ldots\mathbf{c}_N)\in\mathbb{R}^{kN}$ и в k раз превосходит количество наблюдений. Вместо этого *предполагается, что искомая последовательность коэффициентов представляет собой случайный марковский процесс*

$$\mathbf{c}_t=\mathbf{c}_{t-1}+\xi_t,\xi_t\sim N\left(\xi\,|\,\mathbf{0},\lambda\delta\mathbf{I}\right),E\left(\xi_t\xi_s^T\right)=\mathbf{0}, \qquad (6)$$

который начинается с неизвестного первого значения $\mathbf{c}_1\sim N(\mathbf{c}_1\,|\,\mathbf{0},\rho\mathbf{I})$, $\rho\to\infty$. *Параметр дисперсии шума λ является структурным параметром априорной модели и отвечает за степень временной нестационарности коэффициентов регрессии.*

Это типичный пример задачи, в которой плавное изменение параметра λ определяет систему непрерывно вложенных априорных плотностей распределения $\Psi(\mathbf{c}\,|\,\lambda)$ вектора параметров модели, начиная от «однородного» распределения в $\mathbb{R}^k$ при $\lambda = 0$ до «однородного» распределения в $\mathbb{R}^{kN}$ при $\lambda \to \infty$. Такая ситуация фактически представляет собой введение вместо дискретной последовательности целочисленных размерностей понятия «размытой размерности» вектора параметров $\mathbf{c}$, непрерывно меняющейся от k до kN при увеличении параметра λ.*Естественно, что классический критерий AIC оказывается неприменимым для выбора наиболее подходящего для данного сигнала* $(\mathbf{y},\mathbf{x})$ *значения параметра* $0 < \lambda < \infty$.

В этой статье мы рассматриваем параметрическую модель плотности распределения неизвестной генеральной совокупности $F^*(\mathbf{y})$ как смесь условной плотности из заданного семейства $\Phi(\mathbf{y}\,|\,\mathbf{c})$, $\mathbf{c} \in \mathbb{R}^m$ и априорной плотности распределения вектора параметров $\Psi(\mathbf{c}\,|\,\lambda)$:

$$F(\mathbf{y}\,|\,\lambda) = \int \Phi(\mathbf{y}\,|\,\mathbf{c})\Psi(\mathbf{c} \mid \lambda)d\mathbf{c}, \mathbf{c} \in \mathbb{R}^m \tag{7}$$

Значение структурного параметра модели λ, *оцененное по наблюдаемой выборке* $\mathbf{y}$, *обеспечивает оптимальную степень сокращения слишком большой размерности вектора параметров* $\mathbf{c}$. Как только значение λ выбрано, результат анализа представляет собой байесовскую оценку вектора параметров $\mathbf{c}$

$$\hat{\mathbf{c}}_\lambda(\mathbf{y}) = \arg\max\left[\ln\Phi(\mathbf{y}\,|\,\mathbf{c}) + \ln\Psi(\mathbf{c}\,|\,\lambda)\right] \tag{8}$$

Мы будем эксплуатировать ту же идею, что и в (4), т.е. будем с помощью варьирования параметра λ пытаться обеспечить наилучшее приближение модельного распределения $F(\mathbf{y} \mid \lambda)$ (15) и неизвестного распределения генеральной совокупности $F^*(\mathbf{y})$

Мы предлагаем в этой статье два способа обобщения критерия AIC, а также покажем, что классический AIC может быть получен как частный случай обоих способов при принятии специальных предположений об априорной плотности $\Psi(\mathbf{c}\,|\,\lambda)$.

Наконец, мы опишем результаты модельного эксперимента, сравнивающего предложенный обобщенный информационный критерий оценки непрерывного структурного параметра модели с критерием скользящего контроля на задаче анализа нестационарного сигнала.

1. Два способа вычисления расстояния по Кульбаку между неизвестной истинной плотностью распределения и ее моделью

С одной стороны, идея максимизации информации по Кульбаку о распределении выборки значений наблюдаемой переменной $F^*(\mathbf{y})$, содержащейся в модельной плотности $F(\mathbf{y}\,|\,\lambda)$, есть математическое ожидание: $\int\left[\ln F(\mathbf{y}\,|\,\lambda)\right]F^*(\mathbf{y})d\mathbf{y}$. Идея максимизации этой информации путем выбора подходящего значения λ приводит к критерию:

$$\lambda^* = \arg\max_{\lambda} \int\left[\ln F(\mathbf{y}\,|\,\lambda)\right]F^*(\mathbf{y})d\mathbf{y}. \tag{9}$$

Этот способ выбора параметра λ подходит для любого вида истинной плотности распределения $F^*(\mathbf{y})$.

С другой стороны, рассматриваемая нами модель (15) включает в себя произвольный параметр $\mathbf{c}$, как скрытую переменную. Таким образом, мы можем выбирать модель совместной плотности распределения скрытой и наблюдаемой переменной $H(\mathbf{c},\mathbf{y} \mid \lambda) = \Psi(\mathbf{c} \mid \lambda)\Phi(\mathbf{y} \mid \mathbf{c})$ наиболее близкую к истинному распределению $H^*(\mathbf{c},\mathbf{y})$.Этот способ имеет смысл, только если неизвестная истинная плотность распределения $F^*(\mathbf{y})$ согласуется с принятым параметрическим семейством распределений $\Phi(\mathbf{y}\,|\,\mathbf{c})$, то есть существует такое распределение $\Psi^*(\mathbf{c})$, что выполняется

$$F^*(\mathbf{y}) = \int \Psi(\mathbf{y}\,|\,\mathbf{c})\Psi^*(\mathbf{c})d\mathbf{c}. \tag{10}$$

Тогда $H^*(\mathbf{c},\mathbf{y}) = \Psi^*(\mathbf{c})\Phi(\mathbf{y}\,|\,\mathbf{c})$, и «идеальный» критерий выбора λ принимает вид

$$\lambda^* = \arg\max_{\lambda} \iint \big[\ln H(\mathbf{c},\mathbf{y}\,|\,\lambda)\big] H^*(\mathbf{c},\mathbf{y}) d\mathbf{c} d\mathbf{y}. \tag{11}$$

Мы увидим, что формализации (9) и (11) приведут к существенно различным обобщениям классического информационного критерия Акаике для выбора значения непрерывного структурного параметра.

2. Основные предположения и некоторые свойства семейств параметрических плотностей

Предположения. Мы ограничимся здесь рассмотрением случая параметрических семейств плотностей распределения $\varphi(y \mid \mathbf{c})$, для которых логарифмическая функция правдоподобия $\ln\Phi(\mathbf{y} \mid \mathbf{c})$ ассимптотически квадратична в окрестности оценки максимальногоправдоподобия $\mathbf{c}$, т.е. для достаточно большого размера N выборки $\mathbf{y} = (y_j, j = 1, \ldots, N)$ можно считать, что

$$\ln\Phi(\mathbf{y}\,|\,\mathbf{c}) = \ln\Phi(\mathbf{y}\,|\,\hat{\mathbf{c}}(\mathbf{y})) + (1/2)(\mathbf{c} - \hat{\mathbf{c}}(\mathbf{y}))^T \mathbf{A}_N(\mathbf{c} - \hat{\mathbf{c}}(\mathbf{y})),\ \bigtriangledown_{\mathbf{c}} \log\Phi(\mathbf{y}\,|\,\mathbf{c}) = A_N(\mathbf{c} - \hat{\mathbf{c}}(\mathbf{y})). \tag{12}$$

Причем гессиан $\mathbf{A}_N = \bigtriangledown^2_{\mathbf{cc}} \ln\Phi(\mathbf{y}\,|\,\mathbf{c})$, называемый информационной матрицей Фишера, не зависит от точки $\mathbf{c}$, в которой определен.

Рассмотрим теперь семейство плотностей априорного распределения скрытой переменной $\Psi(\mathbf{c}\,|\,\lambda)$. Будем полагать, что каждая из этих плотностей является нормальной, возможно вырожденной, с нулевым математическим ожиданием и ковариационной матрицей, определяемой значением структурного параметра λ. Это приводит к тому, что логарифмическая функция правдоподобия $\ln\Psi(\mathbf{c}\,|\,\lambda)$ есть квадратичная функция, достигающая своего максимального значения в нуле $\bigtriangledown_{\mathbf{c}} \log\Psi(\mathbf{0}\,|\,\lambda) = \mathbf{0}$ и определяемая своим Гессианом $\mathbf{B}_\lambda = \bigtriangledown^2_{\mathbf{cc}} \ln\Psi(\mathbf{c}\,|\,\lambda)$, так что

$$\ln\Psi(\mathbf{c}\,|\,\lambda) = const + (1/2)\mathbf{c}^T\mathbf{B}_\lambda\mathbf{c}. \tag{13}$$

Что касается неизвестной плотности распределения выходной переменной $F^*(\mathbf{y})$, то мы будем предполагать, что оно согласуется с семейством плотностей $\Phi(\mathbf{y} \mid \mathbf{c})$ в том смысле, что существует неизвестная плотность $\Psi^*(\mathbf{c})$, которая допускает представление

$$F^*(\mathbf{y}) = \int \Phi(\mathbf{y}\,|\,\mathbf{c})\Psi^*(\mathbf{c})d\mathbf{c} \tag{14}$$

Свойства. Рассмотрим произвольную выборку $\mathbf{y}$, порожденную вероятностным распределением $\Phi(\mathbf{y} \mid \mathbf{c})$ с некоторым фиксированным значением параметра $\mathbf{c}$. Хорошо известно, что для гораздо более широкого класса условных распределений, чем те, что описаны выше (12), таких, что если $\mathbf{A}_N$ есть матрица с полным рангом $\mathrm{rank}(\mathbf{A}_N) = n$, произвольная оценка максимума правдободия $\hat{\mathbf{c}}(\mathbf{y})$оказывается несмещенной

$$\int \hat{\mathbf{c}}(\mathbf{y})\Phi(\mathbf{y}\,|\,\mathbf{c})d\mathbf{y} = \mathbf{c}, \tag{15}$$

а ее условная ковариационная матрица полностью определяется информационной матрицей Фишера:

$$\int \big(\hat{\mathbf{c}}(\mathbf{y}) - \mathbf{c}\big)\big(\hat{\mathbf{c}}(\mathbf{y}) - \mathbf{c}\big)^T \Phi(\mathbf{y}\,|\,\mathbf{c})d\mathbf{y} = -\mathbf{A}_N^{-1}. \tag{16}$$

В более общем случае, когда $\mathrm{rank}(\mathbf{A}_N) < n$, (22) и (16) имеют вид:

$$\int \mathbf{A}_N\big(\hat{\mathbf{c}}(\mathbf{y}) - \mathbf{c}\big)\big(\hat{\mathbf{c}}(\mathbf{y}) - \mathbf{c}\big)^T \Phi(\mathbf{y}\,|\,\mathbf{c})d\mathbf{y} = \mathbf{0}, \tag{17}$$

$$\int \Big[\mathbf{A}_N\big(\hat{\mathbf{c}}(\mathbf{y}) - \mathbf{c}\big)\Big]\Big[\mathbf{A}_N\big(\hat{\mathbf{c}}(\mathbf{y}) - \mathbf{c}\big)\Big]^T \Phi(\mathbf{y}\,|\,\mathbf{c})d\mathbf{y} = -\mathbf{A}_N \tag{18}$$

Если условия (12) и (20) выполнены, то произвольная Байессовская оценка (8) есть линейная функция от оценки максимального правдоподобия

$$\hat{\mathbf{c}}_\lambda(\mathbf{y}) = (\mathbf{A}_N + \mathbf{B}_\lambda)^{-1}\mathbf{A}_N\hat{\mathbf{c}}(\mathbf{y}) \tag{19}$$

с условной ковариационной матрицей относительно фиксированного значения параметра $\mathbf{c}$

$$\int \big(\hat{\mathbf{c}}_\lambda(\mathbf{y}) - \hat{\mathbf{c}}_\lambda(\mathbf{c})\big)\big(\hat{\mathbf{c}}_\lambda(\mathbf{y}) - \hat{\mathbf{c}}_\lambda(\mathbf{c})\big)^T \Phi(\mathbf{y}\,|\,\mathbf{c})d\mathbf{y} = -(\mathbf{A}_N + \mathbf{B}_\lambda)^{-1}\mathbf{A}_N(\mathbf{A}_N + \mathbf{B}_\lambda)^{-1}, \tag{20}$$

где $\hat{\mathbf{c}}_\lambda(\mathbf{c})$ есть условное матеметическое ожидание

$$\hat{\mathbf{c}}_\lambda(\mathbf{c}) = \int \hat{\mathbf{c}}_\lambda(\mathbf{y})\Phi(\mathbf{y}\,|\,\mathbf{c})d\mathbf{y} = (\mathbf{A}_N + \mathbf{B}_\lambda)^{-1}\mathbf{A}_N\hat{\mathbf{c}}. \tag{21}$$

3. Критерий максимума информации о распределении наблюдаемой переменной

Непосредственная реализация критерия (9) невозможна хотя бы потому, что истинное распределение $F^*(\mathbf{y})$ неизвестно. Максимизация функции правдободобия по одной доступной реализации $\ln F(\mathbf{y} \mid \lambda)$, как несмещенной оценки критерия, также бессмысленно, так как при этом будут предпочтительны значения структурного параметра, приводящие к слишком большим размерностям $\mathbf{c} \in \mathbb{R}^m$.

Для того чтобы преодолеть «проклятие единственной выборки», мы применим идею компромисса, обосновывающего классический информационный критерий Акаике ([1]), а именно вообразим существование другой независимой выборки $\tilde{\mathbf{y}}$. Пусть по ней получена произвольная байессовская оценка $\hat{\mathbf{c}}_\lambda(\tilde{\mathbf{y}})$ (8). Заменим $\ln F(\mathbf{y} \mid \lambda)$ в (9) на математическое ожидание $\ln \Phi(\mathbf{y}\,|\,\hat{\mathbf{c}}_\lambda(\tilde{\mathbf{y}}))$:

$$\hat{\lambda} = \arg\max_{\lambda} \int\left\{\int\left\{\int \left[\ln \Phi(\mathbf{y}\,|\,\hat{\mathbf{c}}_\lambda(\tilde{\mathbf{y}}))\right]\Phi(\tilde{\mathbf{y}} \mid \mathbf{c})d\tilde{\mathbf{y}}\right\}\Phi(\mathbf{y}\,|\,\mathbf{c})d\mathbf{y}\right\}\Psi^*(\mathbf{c})d\mathbf{c}. \tag{22}$$

Предложение 1. При предположениях (12) и (20),

$$\int\left\{\int\left\{\int \left[\ln \Phi(\mathbf{y}\,|\,\hat{\mathbf{c}}_\lambda(\tilde{\mathbf{y}}))\right]\Phi(\tilde{\mathbf{y}}\,|\,\mathbf{c})d\tilde{\mathbf{y}}\right\}\Phi(\mathbf{y}\,|\,\mathbf{c})d\mathbf{y}\right\}\Psi^*(\mathbf{c})d\mathbf{c} =$$
$$\int J_1(\lambda\,|\,\mathbf{y})F^*(\mathbf{y})d\mathbf{y}, \tag{23}$$
$$J_1(\lambda\,|\,\mathbf{y})F^*(\mathbf{y}) = \ln \Phi\big(\mathbf{y}\,|\,\hat{\mathbf{c}}_\lambda(\mathbf{y})\big) - Tr\left[\mathbf{A}_N(\mathbf{A}_N + \mathbf{B}_\lambda)^{-1}\right].$$

Доказательство. основано на квадратичном представлении $\ln \Phi(\mathbf{y} \mid \mathbf{c})$ (12) в $\mathbf{c} = \hat{\mathbf{c}}_\lambda(\tilde{\mathbf{y}})$ в выражениях (16)— (20).

Эта теорема указывает на построение непрерывного аналога классического AIC. Хотя распределение $\Psi^*(\mathbf{c})$ в (12) по-прежнему неизвестно, а значит непосредственно применить критерий (22) невозможно, но выражение (23) дает легко вычисляемую функцию $J_1(\lambda \mid \mathbf{y})$, которая является несмещенной оценкой полного критерия. Аналогично рассуждениям Акаике, эту функцию можно также максимизировать по искомому значению структурного параметра:

$$\hat{\lambda} = \arg\max_{\lambda}\left\{\ln \Phi\big(\mathbf{y}\,|\,\hat{\mathbf{c}}_\lambda(\mathbf{y})\big) - Tr\left[\mathbf{A}_N(\mathbf{A}_N + \mathbf{B}_\lambda)^{-1}\right]\right\}. \tag{24}$$

Это и есть обобщенный информационный критерий Акаике (6). Сравнение критериев (24) и (6) позволяет интерпретировать штрафной член $Tr\left[\mathbf{A}_N(\mathbf{A}_N + \mathbf{B}_\lambda)^{-1}\right]$, как условную «размытую размерность» параметра $\mathbf{c}$, выбор которого ограничено распределением $\ln \Psi(\mathbf{c} \mid \lambda)$. □

4. Критерий максимума информации о совместном распределениии наблюдаемой переменной и скрытого параметра

Критерий (11) также невозможно вычислительно реализовать, не только по тому, что совместное распределение $H^*(\mathbf{c},\mathbf{y} \mid \lambda)$ неизвестно, но также из-за того, что призвольный параметр $\mathbf{c}$ скрыт от наблюдателя. Как и в предыдущем разделе мы применим компромисс, заключающийся в использовании независимой воображаемой выборки $\tilde{\mathbf{y}}$ и замене $\ln H(\mathbf{c},\mathbf{y} \mid \lambda)$ на математическое ожидание $\ln H\big(\hat{\mathbf{c}}(\tilde{\mathbf{y}}),\mathbf{y}\,|\,\lambda\big)$:

$$\hat{\lambda} = \arg\max \iint\left\{\int\left[\ln H\big(\hat{\mathbf{c}}(\tilde{\mathbf{y}}),\mathbf{y}\,|\,\lambda\big)\right]\Phi(\tilde{\mathbf{y}}\,|\,\mathbf{c})\right\}H^*(\mathbf{c},\mathbf{y})d\mathbf{c}d\mathbf{y}.$$

Здесь $\ln H\big(\hat{\mathbf{c}}(\tilde{\mathbf{y}}),\mathbf{y}\,|\,\lambda\big) = \ln\Phi\big(\mathbf{y},\hat{\mathbf{c}}(\tilde{\mathbf{y}})\big) + \ln\Psi\big(\hat{\mathbf{c}}(\tilde{\mathbf{y}})\,|\,\lambda\big)$ и $H^*(\mathbf{c}\,|\,\mathbf{y}) = \Phi(\mathbf{y}\,|\,\mathbf{c})\Psi^*(\mathbf{c})$. Мы получаем критерий

$$\hat{\lambda} = \arg\max_{\lambda} \int\left\{\int\left\{\int\Big[\ln\Phi\big(\mathbf{y},\hat{\mathbf{c}}(\tilde{\mathbf{y}})\big) + \ln\Psi\big(\hat{\mathbf{c}}(\tilde{\mathbf{y}})\,|\,\lambda\big)\Big]\Phi(\tilde{\mathbf{y}}\,|\,\mathbf{c})\right\}\Phi(\mathbf{y}\,|\,\mathbf{c})d\mathbf{y}\right\}\Psi^*(\mathbf{c})d\mathbf{c}. \quad (25)$$

который отличается от (22) только наличием дополнительного слагаемого $\Psi\big(\hat{\mathbf{c}}(\tilde{\mathbf{y}})\,|\,\lambda\big)$.

Предложение 2. При принятых предположениях (12) и (20),

$$\int\left\{\int\left\{\int\Big[\ln\Phi\big(\mathbf{y}\,|\,\hat{\mathbf{c}}(\tilde{\mathbf{y}})\big) + \ln\Psi\big(\hat{\mathbf{c}}(\tilde{\mathbf{y}})\,|\,\lambda\big)\Big]\Phi(\tilde{\mathbf{y}}\,|\,\mathbf{c})\right\}\Phi(\mathbf{y}\mid\mathbf{c})d\mathbf{y}\right\}\Psi^*(\mathbf{c})d\mathbf{c} =$$
$$\int J_2(\lambda\,|\,\mathbf{y})F^*(\mathbf{y})d\mathbf{y}, \quad (26)$$
$$J_2(\lambda\,|\,\mathbf{y}) = \ln\Phi\big(\mathbf{y},\hat{\mathbf{c}}(\mathbf{y})\big) + \ln\Psi\big(\hat{\mathbf{c}}(\mathbf{y})\,|\,\lambda\big) - Tr\Big[\mathbf{A}_N(\mathbf{A}_N+\mathbf{B}_\lambda)^{-1}\Big].$$

Доказательство. этого утверждения основано на таких же рассуждениях, которые были сделаны для Предложения 1.

Выражение (26) показывает, что функция $J_2(\lambda\mid\mathbf{y})$ есть несмещенная оценка критерия (25). Его непосредственная максимизация и есть другая версия обобщения классического AIC

$$\hat{\lambda} = \arg\max_{\lambda} = \left\{\ln\Phi\big(\mathbf{y},\hat{\mathbf{c}}(\mathbf{y})\big) + \ln\Psi\big(\hat{\mathbf{c}}(\mathbf{y})\,|\,\lambda\big) - Tr\Big[\mathbf{A}_N(\mathbf{A}_N+\mathbf{B}_\lambda)^{-1}\Big]\right\}. \quad (27)$$

□

5. Частный случай: классический информационный критерий Акаике

Пусть структурный параметр принимает целые положительные числа $0 \le \lambda \le m$ и урезает вектор параметров с упорядоченными элементами $\mathbf{c} = (\mathbf{c}_\lambda,\mathbf{c}_{m-\lambda}) \in \mathbb{R}$, так же как и в (5) с $n = \lambda$, то есть $\mathbf{c}_\lambda \in \mathbb{R}^\lambda$, $\mathbf{c}_{m-\lambda} \in \mathbb{R}^{m-\lambda}$. Никакой априорной информации о векторе $\mathbf{c}$ нет, то есть

$$\Psi(\mathbf{c}_\lambda\,|\,\lambda) = \prod\nolimits_{i=1}^{\lambda}\psi_i(c_i\,|\,n),\, \psi_i(c_i\,|\,\lambda) = N(c_i\,|\,0,\sigma^2),\, \sigma\to\infty$$

Так как только первая часть вектора параметров входит в условную плотность $\Phi(\mathbf{y}\,|\,\mathbf{c}_\lambda,\mathbf{c}_{m-\lambda})$, то Гессиан $\mathbf{A}_{N,\lambda} = \bigtriangledown^2_{\mathbf{c}_\lambda\mathbf{c}_\lambda}\ln\Phi(\mathbf{y}\mid\mathbf{c}_\lambda,\mathbf{0})$ есть матрица $(\lambda\times\lambda)$.

При принятых предположениях, обе версии обобщения AIC (24) и (27) приводят к критерию (6):

$$\max_{\mathbf{c}_\lambda}\ln\Phi(\mathbf{y}\,|\,\mathbf{c}_\lambda,\mathbf{0}) - \mathrm{rank}\,(\mathbf{A}_{N,\lambda}) \to \max_{\lambda}.$$

6. Применение критерия Акаике в задаче оценивания нестационарной регрессии: модельные эксперименты

В задаче оценки нестационарной регрессии (5)–(6) сама байессовская оценка скрытой последовательности коэффициентов регрессии $\mathbf{c} = (\mathbf{c}_1^T \cdots \mathbf{c}_N^T)^T \in \mathbb{R}^{kN}$ зависит только от отношения λ предполагаемых дисперсий шума в уравнениях наблюдения (5) и состояния (6), в том время ее статистические свойства существенно определяется дисперсией шума в модели наблюдения. Байесовская оценка вектора параметров $\mathbf{c}$ может быть получена минимизацией критерия Flexible Least Squares

$$\sum\nolimits_{t=1}^{N} \left(y_t - \mathbf{x}_t^T \mathbf{c}_t\right)^2 + (1/\lambda) \sum\nolimits_{t=2}^{N} (\mathbf{c}_t - \mathbf{c}_{t-1})^T (\mathbf{c}_t - \mathbf{c}_{t-1}) \to \min(\mathbf{c})$$

с помощью фильтра-интерполятора Калмана-Бьюси [7].

Представим модель в явной форме. Мы будем полагать, что $\mathbf{y} = (y_1 \cdots y_N)^T \in \mathbb{R}^N$ и $\mathbf{c} = (\mathbf{c}_1^T \cdots \mathbf{c}_N^T)^T \in \mathbb{R}^{kN}$ есть вектор-столбцы, $\mathbf{X} = (\mathbf{X}_{ts}, t, = 1, \ldots, N)$ есть блочная матрица размера $(kN \times N)$ с блоками $\mathbf{X}_{ts} = (\mathbf{x}_t, \text{если} t \neq s)$ $(k \times 1)$, $\mathbf{B}_{\lambda,\rho}(kN \times kN)$ есть квадратная блочно-трехдиагональная матрица с диагональю $\big((1/\rho + 1/\lambda)\mathbf{I}, (2/\lambda)\mathbf{I}, \ldots, (2/\lambda)\mathbf{I}, (1/\lambda)\mathbf{I}\big)$ и не диагоналями $\big(-(1/\lambda)\mathbf{I}, \ldots, -(1/\lambda)\mathbf{I}\big)$, где $\mathbf{I}$ есть единичная матрица размера $(k \times k)$. Положим также дисперсию наблюдаемого шума равной единице $\delta = 1$, тогда модель (5) будет тогда давать функцию максимального правдоподобия $\ln \Phi(\mathbf{y} \,|\, \mathbf{c}, \mathbf{X}) = \ln N(\mathbf{y} \mid \mathbf{X}^T \mathbf{c}, \mathbf{I}) = const + (1/2)\mathbf{c}^T \mathbf{A}_N \mathbf{c}$, гессиан которой $\mathbf{A}_N = -\mathbf{X}\mathbf{X}^T$ $(kN \times kN)$ всегда вырожден и, если регрессоры $(x_{it}, t = 1, \ldots, N)$ линейно независимы, имеет максимальный ранг $\mathrm{rank}(\mathbf{A}_N) = N$. Скрытый марковская модель коэффициентов регрессии (6) выражается семейством априорных плотностей распределения $\ln \Psi(\mathbf{y} \,|\, \lambda, \rho) = \ln N(\mathbf{c} \,|\, \mathbf{0}, \mathbf{B}_{\lambda,\rho}^{-1}) = const + (1/2)|\mathbf{B}_{\lambda,\rho}| - (1/2)\mathbf{c}^T \mathbf{B}_{\lambda,\rho} \mathbf{c}$.

Мы проанализировали 200 *независимых реализаций случайного процесса* (5) *длинной* $N = 50$ *, полученного как линейная комбинация трех регрессоров* $(x_{it}, t = 1, \ldots, N)$, $i = 1, \ldots, k$, $k = 3$, *представляющих собой случайный белый шум с нулевым средним, с коэффициентами регрессии, взятыми как синусоидальные последовательности* $c_{it}^* = 4\sin\big((2\pi/N)t + (2\pi/3)(i-1)\big)$ *смещенные друг относительно друга по фазе* 10%. Дисперсией шума в модели наблюдения составляла 10 процентов $\delta = 0.1\left((1/N) \sum_{i=1}^{N} (\mathbf{x}_t^T \mathbf{c}_t)^2\right)$.

Предполагается, что нет никакой априорной информации о векторе коэффициентов в первый момент времени, то есть $\rho \to \infty$. Зависимость «эффективной размерности» последовательности коэффициентов регрессии $(\mathbf{c}_1 \cdots \mathbf{c}_N)$ от предполагаемой дисперсии λ, вычисленная по единственной реализации произвольной последовательности регрессоров, изображена на Рис. 1. Эта размерность равна числу регрессоров в случае нулевой дисперсии $\lambda \to 0$ и достигает длины временных серий при $\lambda \to \infty$.

Для каждой из 200 *смоделированных временных серий, были вычислены* 3 *значения параметра дисперсии* $\hat{\lambda}$, *во-первых, по принципу двух версий обобщенного критерия Акаике* (24) и (27), *во-вторых, традиционным методом скользящего контроля* [7]. *Затем мы применили каждое из полученных значений к оставшимся* 199 *временным сериям, как к контрольному множеству, и сравнили истинную последовательность коэффициентов регрессии* $(\mathbf{c}_1^* \cdots \mathbf{c}_N^*)$ *с полученной оценкой* $(\hat{\mathbf{c}}_{1,\hat{\lambda}} \cdots \hat{\mathbf{c}}_{N,\hat{\lambda}})$ *по критерию*

$$\varepsilon_{\hat{\lambda}} = \sum_{t=1}^{N} (\hat{\mathbf{c}}_{t,\hat{\lambda}} - \mathbf{c}_t^*)^T (\hat{\mathbf{c}}_{t,\hat{\lambda}} - \mathbf{c}_t^*) \Big/ \sum\nolimits_{t=1}^{N} (\mathbf{c}_t^*)^T \mathbf{c}_t^*.$$

Мы получили следующие результаты:

Критерий	$\hat{\lambda}$	$\varepsilon_{\hat{\lambda}}$
Максимум близости распределения наблюдаемой переменной	0.010	0.012
Скользящий контроль	0.033	0.034
Максимум близости совместного распределения наблюдаемой переменной и скрытого параметра	0.072	0.055

Формально, первая версия критерия Акаике показала наилучшие результаты, хотя результаты всех трех критериев очень близки друг к другу. Эксперименты выявили важный факт, что два фундаментально различных подхода: новый класс непрерывного обобщения информационного критерия Акаике и традиционный принцип скользящего контроля не превосходят друг друга в вопросе выбора наиболее подходящего значения параметра дисперсии, отвечающего за нестационарность регрессионой модели. В то же время непрерывный AIC несравненно лучше с вычислительной точки зрения.

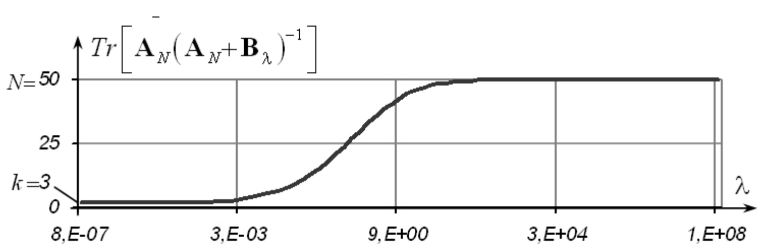

Рис. 1. Эффективная размерность последовательности коэффициентов регрессии как функция от λ (логарифмический масштаб)

Список литературы

1. *Akaike H.* A new look at the statistical model idendification // IEEE Trans. on Automatic Control, Vol. IC-19, No.6, December 1974, pp. 716-723.
2. *Kitagawa G., Akaike H.* A procedure for the modeling of no-stationary time series. // Ann. Inst. Statist. Math., Vol. 30, Part B,1987, pp. 351-363.
3. *Scharz G.* Estimating the dimtnsion of the model. // The Annals of Statistics, Vol. 6,No.2, 1978, pp. 461-464
4. *Bozdogan H.* Model selection fnd Akaik's Information Criterian (AIC): The general theory ands its analitical extensions. // Psychometrica, Vol. 52, No.3, September 1987.
5. *Spiegelhalter D., Best N., Carlin B.* Van der Linde A. Bayesian mesures of model complexity and fit. // Journal of the Royal Statistical Society. Series B (Statistical Methodology), Vol. 64, No.4, 2002, pp. 583-639.
6. *Rodrigues C. C.* The ABC of model selection: AIC, BIC and new CIC. // AIP Conference Proceedings., Vol. 803, November 23, 2005, pp. 80-87.
7. *Markov M., Krasotcina O., Mottl V., Muchnik I.* Time-varying regression model with unknown time-volatility for nonstationary signal analyses. // Proceedings of the 8th IASTED Internation Conference on Signal and Image Processing. Honolulu, Hawaii, USA, August 14-16, 2006.

Статья поступила в редакцию 08.05.2008

УДК 004.81:007.5:612.822

К ТЕОРЕТИКО-МЕТОДОЛОГИЧЕСКИМ ОСНОВАМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

© Воронов А.В.

Институт автоматики и процессов управления ДВО РАН
ул. Радио, 5, г. Владивосток, 690041, Россия
e-mail: *voronov.ia@mail.ru*

Abstract. Actuality of research of reflection by the objective methods is justified. Reacting of different areas of cortex on words has been studied. Activities in the areas F7, F8 and other have been found. The sequence of the states corresponding to increasing of activity of crust structures has been established. Development of works on creation of artificial intelligence is associated with the exposure of neurophysiological mechanisms and construction of neurobiocontrol theory.

Введение

Тематика искусственного интеллекта, в настоящее время, в эпоху информационного общества является приоритетным научным и прикладным направлением в мире. Именно повсеместно применяемые системы искусственного интеллекта и основанные на них интеллектуальные технологии могут стать для человека и цивилизации в целом ключом к обществу, основанному на знаниях. Для разработки эффективных систем искусственного интеллекта необходимо больше внимания уделять исследованию и воспроизведению процессов познавательно-творческой деятельности человека, субъект-субъектного взаимодействия людей в условиях информационного общества и т.п. [1, 2]. Вместе с тем, сегодня в мире интенсивно развивается целый комплекс когнитивных наук по исследованию восприятия, мышления, сознания, поведения, эмоций, языка и других сущностных характеристик человека и мозговых механизмов этих характеристик. Надо эффективно дополнять представления о когнитивных составляющих человека (интеллекте, мышлении, рефлексии, эмоциях и т.п.) и одновременно радикально развивать систему взглядов на разработку и развитие подходов к созданию эффективных систем искусственного интеллекта. Человек, осуществляя и анализируя деятельность, способен рефлексировать, осознавать результаты деятельности, применять, пожалуй, наиболее сложные формы обратной связи. Причем, именно биологической обратной связи нейрофизиологи придают универсальную и решающую роль при когнитивных процессах. Исследование когнитивных процессов и принципов биологической обратной связи, и, особенно, рефлексии, несомненно, актуальнейшая проблема современности, и может стать важнейшим и ключевым фактором при разработке эффективных систем искусственного интеллекта.

1. Об изучении рефлексии

Изучением рефлексии давно занимается философия и психология. В философии рефлексия понимается как один из основных принципов философского мышления, лежащий в основе сознания и направленный на осмысление и обоснование собственных умозаключений. В психологии проблема рефлексии рассматривается как процесс самопознания субъектом внутренних психических актов и состояний. Понятие рефлексии получило развитие в алгебраическом рефлексивном подходе, предложенном В.А. Лефевром. Тем не менее, детальное понимание рефлексии и существенный прогресс в понимании механизмов рефлексии не могут быть достигнуты чисто философскими построениями, методами психологии и математики без систематического исследования явления рефлексии объективными методами. В 90-х годах XX-го века, объявляемых десятилетием мозга, в мире росло экспоненциально количество публикаций в области когнитивных наук [3]. Был получен ряд интересных результатов по различным психическим состояниям и др. Однако изучению рефлексии нейрофизиологи и специалисты других близких дисциплин уделили недостаточное внимание. Понятие рефлексии связано с представлениями о самом субъекте-человеке и неразрывно от его мышления, интеллекта, эмоций, языка и т.д. Одним из подходов к детальному исследованию когнитивных процессов, в том числе явления рефлексии, может стать подход, основанный на знаково-речевой парадигме [4]. При этом существенную роль могут оказать современные методы исследования мозга: компьютерная электроэнцефалография, магнитно-резонансная томография, компьютерная томография и др.

2. Методика исследования и результаты

Некоторые явления когнитивности (восприятие, интеллект и др.) исследуют методами и методиками психологии. Прямое применение психологических наработок при исследовании когнитивных процессов методами исследования мозга (компьютерной электроэнцефалографии, магнитно-резонансной томографии, компьютерной томографии и др.) весьма желательно, но, к сожалению, практически не возможно вследствие специфических условий методов изучения мозга (см., например [5]. Нами в [6]предложена валидная методика исследования когнитивных процессов человека, в том числе рефлексии с помощью метода компьютерной электроэнцефалографии с учетом специфики электроэнцефалографического исследования.

Рассмотрим условия и результаты исследования испытуемого с применением предложенной методики по реагированию на социально значимые стимулы-слова: дружба, честность, совесть, агрессия, доброта и т.д., всего около сотни слов, предъявляемых с помощью компьютера. Испытуемый в начале исследования до предъявления слов находился в состоянии спокойного бодрствования с закрытыми глазами, а после предъявления слов – с открытыми глазами. Для переключения внимания испытуемого после предъявления каждого слова на экране монитора предъявлялось белое пятно. В течение исследования регистрировали электроэнцефалограмму (ЭЭГ) в 16 отведениях на электроэнцефалографической установке «Нейрон-Спектр-2» фирмы «НейроСофт». Электроды располагались по стандартной системе 10-20 в симметричных точках левого и правого полушария (Fp1, Fp2, F3, F4, F7, F8, C3, C4, P3, P4, T3, T4, T5, T6, O1, O2). Референтные электроды (A1, A2) располагались на мочках ушей. Заземляющий электрод располагался в точке Cz. Постоянная времени усилителя была 0,13 с. Квантование КЭЭГ происходило с частотой 256 Гц. Сопротивление электродов было 25-40 кОм. В течение всего времени эксперимента вели запись изображения лица испытуемого и звуков эксперимента (команд врача-оператора КЭЭГ, аудио стимулов-слов и пр.) с помощью стандартной цифровой видеокамеры miniDV. Предъявляемые испытуемому слова классифицировались им до и после эксперимента как «позитивные» или «негативные». Для анализа электроэнцефалографических данных компьютерной программой «Нейрон-Спектр» фирмы «НейроСофт» выбирались 2-секундные отрезки-эпохи ЭЭГ, не содержащие глазодвигательные и мышечные артефакты, спайки и острые волны. Измерение и вычисление амплитуды сигналов ЭЭГ производилось как «пик-пик». Частотно-спектральный анализ выполняли для монополярной, биполярной кольцевой и биполярной перекрестной монтажных схем. Рассматривались следующие частотные диапазоны: дельта=0,5-4 Гц, тэта=4-8 Гц, альфа=8-14 Гц, бета 1 = 14-20 Гц, бета 2 = 20-35 Гц. Спектральная плотность ЭЭГ вычислялась методом быстрого преобразования Фурье. Значения индексов ритмов для каждого электроэнцефалографического отведения вычислялось, как отношение полной мощности спектра в отведении к полной мощности спектра в данном частотном диапазоне этого отведения. Анализировалась электроэнцефалограмма, включающая 80 шт. 2-секундных отрезков эпох, соответствующих реагированию на пятна, а также 80 шт. 2-секундных отрезков эпох, соответствующих реагированию на 80 слов (40 «позитивных» слов и 40 «негативных» слов). Общая длительность анализируемой электроэнцефалограммы составила 320 с.

При фоновом состоянии испытуемого с закрытыми глазами в задних отделах мозга наблюдали типичный хорошо выраженный альфа ритм, который существенно в 1,7 раза уменьшался при открывании глаз испытуемого и его реагировании на стимулы, что свидетельствует об увеличение степени активации коры головного мозга [5, 6]. Причем изменение величины альфа ритма может служить показателем изменения активации корковых областей мозга [5, 6]. В частотном диапазоне дельта и тэта для целого ряда корковых областей (C3, C4, P3, P4, O1, O2 и др.) были выявлены большие в 1,1 раза значения индексов ритмов при реагировании испытуемого на слова, по сравнению с реагированием на пятна. Наблюдали большие в 1,2-1,5 раза значения индексов тета ритма при реагировании на «позитивные» слова, чем на «негативные» слова. Тэта ритм тесно связан с эмоциональным напряжением и его усиление рассматривают как симптом увеличения эмоциональных переживаний [5]. Поэтому наблюдаемые нами большие значения индексов тэта ритма при реагировании на «позитивные» слова, чем «негативные» слова, свидетельствуют об активации корковых областей и могут быть связаны с усилением эмоциональных переживаний испытуемого при реагировании на «позитивные» слова. Меньшие значения индексов тэта ритма при реагировании на «негативные» слова, чем «позитивные» слова могут быть связаны с нейрофизиологическими механизмами торможения при реагировании на «негативные» слова. В частотном диапазоне бета 2 с помощью монополярной, биполярной кольцевой и биполярной перекрестной монтажных схем выявлена локализация максимальной активности во фронтальных областях слева F7 и справа F8. Причем, у человека в области F7 расположена зона Брока, которая отвечает за речедвигательные функции. Было показано [7], что у приматов корковая фронтальная область слева содержит «зеркальные» нейроны, которые отвечают за социальное поведение. Наблюдаемые нами значительные активности во фронтальных областях слева (F7) и справа (F8) у человека при реагировании на социально значимые стимулы-слова могут свидетельствовать о важности указанных областей в процессах рефлексии, связанных с социальным поведением и основами сознания человека. Причем в нейрофизиологических процессах реагирования задействованы и другие, кроме F7 и F8 области коры.

В настоящее время одним из наиболее интенсивно развивающихся направлений, в основном как экспериментальное, техническое и практическое направление медицины, является нейробиоуправление, основанное на методе биологической обратной связи (БОС) и сходных с методом БОС методов регуляции и саморегуляции. Причем метод БОС и методы сходные с методом БОС, часто реализуют с помощью данных электроэнцефалографии. Процессы БОС, имеют место и в наших исследованиях. Сегодня, несмотря на существенные практические результаты применения нейробиоуправления, механизмы и теория нейробиоуправления пока далеки от понимания. Дальнейшее направление исследовательских работ по выявлению принципов создания эффективных систем искусственного интеллекта мы связываем с выявлением нейрофизиологических механизмов нейробиоуправления и построением теории нейробиоуправления.

Заключение

В результате изучения церебрального реагирования на слова в состоянии спокойного бодрствования методом КЭЭГ выяснено следующее.

Изменения значений индексов ритмов происходят в дельта, тэта, альфа, бета 1 и бета 2 диапазонах.

Монополярная, биполярная кольцевая и биполярная перекрестная монтажные схемы свидетельствуют о локализации активности в диапазоне бета 2 в корковых областях F7, F8 и T4.

Последовательность состояний, соответствующих увеличению активности некоторых корковых структур (центральных, теменных, затылочных и височных, особенно справа) может быть представлена как: ФС-СБОГ-РНС-РПС, где ФС – фоновое состояние (состояние спокойного бодрствования с закрытыми глазами), СБОГ – состояние спокойного бодрствования с открытыми глазами, РНС – состояние реагирования на «негативные» слова, а РПС – состояние реагирования на «позитивные» слова, отвечающее наибольшей активности корковых структур.

Дальнейшее направление исследовательских работ по выявлению принципов создания эффективных систем искусственного интеллекта связывается с выявлением нейрофизиологических механизмов и построением теории нейробиоуправления.

Автор благодарит к.м.н, врача-невролога Медицинского объединения ДВО РАН Горбач Т.А. за электроэнцефалографические исследования.

Список литературы

1. *Воронов А.В.* Методологические аспекты искусственного интеллекта // Научно-техническая информация. Сер. 2, 2007, №7, с. 1-6.
2. *Voronov A. V.* Methodological aspects of artificial intelligence // Automatic documentation and mathematical linguistics. 2007, vol. 41, N4, p. 131-137.
3. *Raichle M.E.* Cognitive neuroscience: Bold insights // Nature. 2001, vol.412, N 6843. P. 128-130.
4. *Брушлинский А.В.* Психология индивидуального и группового субъекта в изменяющемся обществе // Вестник РАН. 2002, т. 72, №2, с.162-169.
5. *Гнездицкий В.В.* Обратная задача ЭЭГ и клиническая электроэнцефалография (картирование и локализация источников электрической активности мозга). М.: МЕДпресс-информ, 2004.
6. *Воронов А.В., Горбач Т.А.* Исследование когнитивности человека. Методика и электроэнцефалографические исследования // Информатика и системы управления. 2007, №1, с. 45-56.
7. *Галлезе В., Риццолатти Д., Фогасси Л.* Зеркальная часть мозга // В мире науки. 2007, №3, с. 23-29.

Статья поступила в редакцию 18.04.2008

УДК 517.983

О СТРОЕНИИ ОСТАТОЧНОГО ПОДПРОСТРАНСТВА ПОЛУУНИТАРНОЙ ДИЛАТАЦИИ ЛИНЕЙНОГО ОГРАНИЧЕННОГО ОПЕРАТОРА В БАНАХОВОМ ПРОСТРАНСТВЕ С ВНУТРЕННИМ ПРОИЗВЕДЕНИЕМ

© Д. Л. Тышкевич

Таврический национальный университет им. В. И. Вернадского,
факультет математики и информатики
пр-т Вернадского, 4, г. Симферополь, Крым, Украина, 95007
e-mail: *dtyshk@inbox.ru*

Abstract

In this work we analyze and describe the construction of residual subspace of a semiunitary dilation of a linear bounded operator acting in a Banach space with an indefinite inner product.

Введение

О чём здесь речь. *Целью* данной работы является описание конструкции остаточного подпространства полуунитарной дилатации линейного непрерывного оператора, действующего в банаховом пространстве, наделённом (индефинитным) внутренним произведением, при определённых условиях на оператор и пространство, в котором оператор действует.

Дилатация (dilation) оператора — это такое расширение данного оператора, с выходом из основного пространства, которое «сохраняет степени» (см. ниже (13) на стр. 82). Использование полуунитарной и унитарной дилатации явилось, пожалуй, самым мощным средством для изучения неунитарных сжимающих операторов в гильбертовом пространстве при наличии тесных связей с теорией характеристических функций и функциональных моделей сжимающих операторов (хорошо известны книги [4, 9]). Работа [3], в которой приводилась конструкция уже J–унитарной дилатации но уже и для произвольного оператора (несжатия) в гильбертовом пространстве, открыла путь для исследования операторов, действующих в пространствах с индефинитным внутренним произведением. С тех пор разными авторами были построены различные конструкции полуунитарной и унитарной дилатаций в пространстве Понтрягина и в пространстве Крейна, однако позже выяснилась общность этих конструкций (см. список источников на стр. 82). Однако конструкции полуунитарной (и тем более унитарной) дилатации в пространствах с внутренним произведением, более общих, чем пространства Крейна (насколько известно автору) *не исследовались* (также автором *не найдены* работы, в которых бы изучалось строение остаточного подпространства и в случае пространств Крейна). Так что, насколько позволяет делать вывод осведомленность автора[1], конструкция полуунитарной дилатации оператора, действующего в банаховом пространстве с внутренним

[1]Кроме тщательного исследования литературы, сюда входят и довольно многочисленные беседы автора с известным специалистом в этой области Т. Я. Азизовым, который был оппонентом кандидатской диссертации автора [22], и которому автор, пользуясь случаем, ещё раз выражает свою признательность.

произведением, а также конструктивное описание её остаточного подпространства приведены в данной работе *впервые.*

Соглашения и обозначения. «ПВП» — сокращение для «пространство с внутренним произведением». Под *внешней ортогональной суммой* ПВП $\mathfrak{X}$ и $\mathfrak{Y}$ понимается декартово произведение этих пространств, наделённое внутренним произведением $[\langle x_1, y_1\rangle, \langle x_2, y_2\rangle] := [x_1, x_2]_{\mathfrak{X}} + [y_1, y_2]_{\mathfrak{Y}}$. Знак $[\dot{+}]$ означает (внутреннюю) прямую ортогональную сумму. Все рассматриваемые в данной работе ПВП полагаются *невырожденными.*

$\mathrm{B}(\mathfrak{X}, \mathfrak{Y})$ означает совокупность всех линейных ограниченных всюду заданных операторов из банахова пространства $\mathfrak{X}$ в банахово пространство $\mathfrak{Y}$. $\mathrm{R}(T)$ — область значений линейного оператора T. Сильный предел последовательности $\{x_n\}_{n\in\mathbb{N}}$ обозначен через $s.\lim\limits_{n\to\infty} x_n$.

Через $I_{\mathfrak{X}}$ обозначен единичный оператор в пространстве $\mathfrak{X}$. Под натуральными мы понимаем здесь целые числа, начиная с единицы (а не с нуля); обозначение множества натуральных чисел стандартное – $\mathbb{N}$; $\mathbb{N}_+$ – обозначение для множества $\{0\}\cup\mathbb{N}$ (расширенный натуральный ряд). Через $\overline{a, b}$ обозначается отрезок расширенного натурального ряда $\{a, a+1, \dots, b\}$ $(a \leqslant b)$.

С понятиями, которые в данной работе не оговариваются и особо не разъясняются, можно подробнее ознакомится в [7, 12, 18].

Необходимые определения и факты

Правильные банаховы пространства. Банахово пространство $\mathfrak{X}$ с внутренним произведением $[\cdot,\cdot]$ называется *правильным банаховым пространством* (regular Banach space; далее, коротко, п.б.п. – [6, 16, 20, 22]) если

$$\exists\, b > 0 \quad \forall\, x, y \in \mathfrak{X} \quad |[x, y]| \leqslant b\|x\|\|y\|; \tag{1}$$

$$\exists\, c > 0 \quad \forall\, x \in \mathfrak{X} \quad \sup_{\|y\|\leqslant 1} |[x, y]| \geqslant c\|x\|. \tag{2}$$

В случае, если выполнено лишь условие (1), говорят, что $\mathfrak{X}$ есть пространство *с мажорантой*([7, 20, 22]).

Замечание 1. П.б.п. особенно прозрачно определяется при помощи так называемого *оператора Грама* ([1, 2, 7, 12]) пространства $\mathfrak{X}$, т.е. линейного всюду заданного оператора $G_{\mathfrak{X}}\colon \mathfrak{X} \to \mathfrak{X}^a$, определяемого формулой

$$(G_{\mathfrak{X}} x)(y) := [x, y]_{\mathfrak{X}}\,.$$

Здесь $\mathfrak{X}^a$ – *антисопряжённое* пространство к $\mathfrak{X}$, т.е. совокупность всех сильно непрерывных антилинейных функционалов на $\mathfrak{X}$. Так вот, легко видеть, что условия (1), (2) эквивалентны *ограниченности* (усл. (1)) и *ограниченной обратимости* (усл. (2)) оператора Грама $G_{\mathfrak{X}}$.

Разложение Богнара–Крамли. *Разложением Богнара–Крамли* самосопряжённого оператора A в ПВП $\mathfrak{X}$ называется представление оператора A в форме

$$C^{\sharp}C = A\,, \tag{3}$$

где C – сопрягаемый оператор с областью определения – всем $\mathfrak{X}$ и областью значений, лежащей в некотором ПВП $\mathfrak{Y}$ ([18, 20, 22]). Оператор C называется в [20, 22] *квадратичным расщеплением* самосопряжённого оператора A. Особенно важны квадратичные расщепления с нулевым ядром сопряжённого:

$$\ker C^{\sharp} = \{0\} \tag{4}$$

((4) равносильно плотности образа C в определённых топологиях, [7, 8, 22]). Из (4) и (3) следует равенство

$$\ker C = \ker A\,.$$

В случае банаховых ПВП, как правило, необходима ограниченность C и его сопряжённого (см. таблицу категорий в следующем пункте).

Категории ПВП. В [19, 20, 22] вводятся понятия *категории с сопряжением* и *категории с квадратичным расщеплением*. Охарактеризовать эти категории, не вдаваясь в детали, можно следующим образом. Категория с сопряжением, в которой объектами являются ПВП, а стрелками – линейные всюду определённые операторы между ПВП – замкнуты относительно образования матричных операторов, относительно сопряжения операторов и содержат нулевой объект – нулевое пространство. Категория с квадратичным расщеплением – это категория с сопряжением, в которой произвольный самосопряжённый объект должен обладать квадратичным расщеплением с нулевым ядром сопряжённого, и это квадратичное расщепление – стрелка категории. Мы приведём три примера категорий с сопряжением (банаховых ПВП) в следующей табличке:

Категория	*Объекты*	*Стрелки*
$\mathbf{B}^{(\mathbf{mutadj})}$	Банаховы пространства с мажорантой	Ограниченные сопрягаемые операторы, для которых сопряжённый ограничен
$\mathbf{RegB}^{(\mathbf{adj})}$	Правильные банаховы пространства	Ограниченные сопрягаемые операторы
$\mathbf{Kr}$	Пространства Крейна	Ограниченные операторы

Снизу вверх они образуют башню полных подкатегорий. Из них $\mathbf{Kr}$ и $\mathbf{RegB}^{(\mathbf{adj})}$ являются категориями с квадратичным расщеплением (доказательство того, что $\mathbf{Kr}$ – категория с квадратичным расщеплением содержится, вне рамок теории категорий, в [18]; доказательства для $\mathbf{RegB}^{(\mathbf{adj})}$, в теоретико–категорных рамках, – в [22] и частично в [20]).

О пространстве $\ell_2(\mathfrak{X})$. Для банахова пространства $\mathfrak{X}$ линейное пространство $\ell_2(\mathfrak{X})$ определяется (см., например, [17]) как совокупность всех бесконечных последовательностей $(x_1, x_2, \ldots)$ элементов из $\mathfrak{X}$, для которых сходится ряд $\sum\limits_{n=1}^{\infty} \|x_n\|^2$. Норма в $\ell_2(\mathfrak{X})$ определяется как

$$\|(x_1, x_2 \ldots)\|_{\ell_2(\mathfrak{X})} := \Big(\sum_{n=1}^{\infty} \|x_n\|^2\Big)^{1/2}.$$

С такой нормой $\ell_2(\mathfrak{X})$ становится банаховым пространством. Если $A \in \mathrm{B}(\mathfrak{X}, \mathfrak{Y})$, то естественным образом определяется оператор $\ell_2(A) \in \mathrm{B}\big(\ell_2(\mathfrak{X}), \ell_2(\mathfrak{Y})\big)$:

$$\ell_2(A)(x_1, x_2 \ldots) := (Ax_1, Ax_2, \ldots)\,.$$

При этом для произвольных $A \in \mathrm{B}(\mathfrak{X}, \mathfrak{Y})$, $B \in \mathrm{B}(\mathfrak{Z}, \mathfrak{X})$

$$\ell_2(I_{\mathfrak{X}}) = I_{\ell_2(\mathfrak{X})}, \quad \ell_2(AB) = \ell_2(A)\ell_2(B), \quad \|\ell_2(A)\| = \|A\|\,;$$

таким образом, $\ell_2(\cdot)$ *является* (ковариантным) *функтором* в категории банаховых пространств. Также, в частности, это влечёт *одновременную ограниченную обратимость* операторов A и $\ell_2(A)$. Существует естественный изометрический изоморфизм между $\ell_2(\mathfrak{X}^*)$ и $\ell_2(\mathfrak{X})^*$, осуществляемый отображением

$$\begin{aligned} &\big(\Phi(f_1, f_2, \ldots)\big)\big((x_1, x_2, \ldots)\big) := \sum_{n=1}^{\infty} f_n(x_n)\,, \\ &(x_1, x_2, \ldots) \in \ell_2(\mathfrak{X}), \quad (f_1, f_2, \ldots) \in \ell_2(\mathfrak{X}^*)\,. \end{aligned} \tag{5}$$

Эта же конструкция приводит к изометрической изоморфности $\ell_2(\mathfrak{X}^a)$ и $\ell_2(\mathfrak{X})^a$.

Если $\mathfrak{X}$ – банахово ПВП, то на $\mathfrak{X}$ естественным образом индуцируется внутреннее произведение

$$[(x_1, x_2, \ldots), (y_1, y_2, \ldots)]_{\ell_2(\mathfrak{X})} := \sum_{n=1}^{\infty} [x_n, y_n]_{\mathfrak{X}}\,.$$

В таком случае имеет место связь между сопряжёнными:

$$\ell_2(A)^\sharp = \ell_2(A^\sharp), \tag{6}$$

а операторы Грама в $\mathfrak{X}$ и $\ell_2(\mathfrak{X})$, как нетрудно видеть, связаны соотношением

$$G_{\ell_2(\mathfrak{X})} = \Phi \ell_2(G_{\mathfrak{X}}) \tag{7}$$

(для антисопряжённой версии Φ). Действительно, согласно (5)

$$\begin{aligned} &\big(G_{\ell_2(\mathfrak{X})}(x_1, x_2, \ldots)\big)\big((y_1, y_2, \ldots)\big) = [(x_1, x_2, \ldots), (y_1, y_2, \ldots)]_{\ell_2(\mathfrak{X})} = \sum_{n=1}^{\infty} [x_n, y_n]_{\mathfrak{X}} = \\ &= \sum_{n=1}^{\infty} (G_{\mathfrak{X}} x_n)(y_n) = \big(\Phi(G_{\mathfrak{X}} x_1, G_{\mathfrak{X}} x_2, \ldots)\big)\big((y_1, y_2, \ldots)\big) = \\ &\qquad = \big(\Phi \ell_2(G_{\mathfrak{X}})(x_1, x_2, \ldots)\big)\big((y_1, y_2, \ldots)\big)\,. \end{aligned}$$

Из (6), (7) и замечания 1 следует, что если $\mathfrak{X}$ – п.б.п. (пространство с мажорантой), то и $\ell_2(\mathfrak{X})$ – п.б.п. (пространство с мажорантой). Кроме того, ясно, что если $\mathfrak{X}$ – пространство Крейна, то и $\ell_2(\mathfrak{X})$ является пространством Крейна.

Замечание 2. Более точно, все эти рассуждения показывают, что $\ell_2(\cdot)$ является функтором в категориях $\mathbf{B}^{\mathbf{(mutadj)}}$, $\mathbf{RegB}^{\mathbf{(adj)}}$ и $\mathbf{Kr}$.

Полуунитарные операторы и сдвиги в ПВП. Всюду определённый оператор W, действующий из ПВП $\mathfrak{X}$ в ПВП $\mathfrak{Y}$, и сохраняющий внутреннее произведение: $[Wx_1, Wx_2]_{\mathfrak{Y}} = [x_1, x_2]_{\mathfrak{X}}$ (для всех $x_1, x_2 \in \mathfrak{X}$) (т.е. *изометрический* относительно внутреннего произведения), но образ которого *не совпадает* с $\mathfrak{Y}$, называется *полуунитарным* (semiunitary; ср. [8, 12, 21]). Отметим без доказательства[2] следующие свойства полуунитарного оператора W, которые далее нам понадобятся:

$$\mathfrak{Y} = \mathrm{R}(W)[\dot{+}]\ker W^{\sharp}\,; \tag{8}$$

$$\ker W^{\sharp\, n} = \underset{k\in\overline{0,\,n-1}}{[\dot{+}]} W^k \ker W^{\sharp} \quad (n \in \mathbb{N})\,. \tag{9}$$

Далее, подпространство $\bigcap\limits_{n\in\mathbb{N}} W^n\mathfrak{X}$ пространства $\mathfrak{Y}$ называется *остаточным подпространством* полуунитарного оператора W (residual subspace; см. [8, 10, 11, 20]).

Пусть далее $\mathfrak{X} = \mathfrak{Y}$. Полуунитарный оператор называется *односторонним сдвигом* (unilateral shift), если его остаточное подпространство – нулевое[3]. Сопрягаемый односторонний сдвиг W назовём *проекционно устойчивым* ([22]), если $s.\lim\limits_{n\to\infty} \|W^n W^{\sharp n} x\| = 0$ для любого $x \in \mathfrak{X}$. Специальный класс проекционно устойчивых односторонних сдвигов составляют *правосторонние сдвиги* в пространствах $\ell_2(\mathfrak{X})$:

$$W(x_1, x_2, \ldots) := (0, x_1, x_2, \ldots) \tag{10}$$

(правосторонний сдвиг служит моделью абстрактного одностороннего сдвига в случае гильбертовых пространств: [10]). При этом W является одновременно изометрическим относительно нормы и внутреннего произведения. Сопряжённый к W есть *левосторонний сдвиг*:

$$W^{\sharp}(x_1, x_2, x_3, \ldots) := (x_2, x_3, \ldots). \tag{11}$$

Из (10) и (11) непосредственно видна проекционная устойчивость W:

$$\|W^n W^{\sharp n}(x_1, x_2, \ldots)\|^2 = \|(0, \ldots, \underset{n}{0}, x_{n+1}, x_{n+2}, \ldots)\|^2 = \sum_{k=n+1}^{\infty} \|x_k\|^2 \to 0, \quad n \to \infty.$$

[2]Доказательства данных свойств – чисто технические, при этом не требуется никаких особых топологических выкладок – лишь «линейные» рассуждения и работа понятия сопряжённости. Эти доказательства имеются в наших работах [21, предл. 3, лем. 1], [22, предл. 5.3, лем. 5.17] (соответственно).

[3]Это – косвенное определение. Прямое (равносильное приведённому) определение в случае индефинитности внутреннего произведения требует определённых топологических выкладок и дополнительных определений, которых мы стремимся в данной работе избегать, так как на формулировку и доказательство основного результата они не имеют прямого влияния.

Отметим следующий важный момент: из (11) непосредственно вытекает равенство

$$\ker W^\sharp = \{(x,0,0,\ldots) \mid x \in \mathfrak{X}\} \tag{12}$$

(т.е. $\ker W^\sharp$ изометрически изоморфно $\mathfrak{X}$).

Предложение 3. Пусть W – правосторонний сдвиг (10), и $\{x_{kn}\}_{\substack{n\in\mathbb{N}\\ k\in\overline{0,n-1}}}$ – двойная последовательность «треугольного вида» векторов из $\ker W^\sharp$. Тогда

$$\Big\{\sum_{k=0}^{n} W^k x_{kn}\Big\}_{n\in\mathbb{N}} \text{— \textit{сильно сходится}} \;\Leftrightarrow\; \lim_{n\to\infty}\sum_{k=0}^{n}\|x_{kn}\|^2 < \infty.$$

Доказательство. Справедливость этого утверждения обеспечивается цепочкой:

$$\Big\|\sum_{k=m}^{n} W^k x_{kn}\Big\|^2 \overset{(10),(12)}{=} \|(0,\ldots,\underset{m}{0},x_{mn},x_{m+1,n},\ldots,x_{nn},0,0,\ldots)\|^2 = \sum_{k=m}^{n}\|x_{kn}\|^2.$$

□

Полуунитарная дилатация. Известное определение дилатации линейного непрерывного (или, по меньшей мере, всюду заданного) оператора ПВП с гильбертовым носителем ([3, 4, 5, 9, 10, 12, 13, 14, 15]) без труда может быть перенесено на случай произвольных ПВП.

Пусть T – всюду заданный линейный оператор в ПВП $\mathfrak{X}$, φ – инъективное изометрическое (относительно внутреннего произведения) вложение $\mathfrak{X}$ в некоторое ПВП $\mathfrak{K}$. Линейный всюду заданный оператор V в $\mathfrak{K}$ называется *дилатацией* оператора T, если

$$[T^n x_1, x_2]_{\mathfrak{X}} = [V^n \varphi x_1, \varphi x_2]_{\mathfrak{K}} \quad (n\in\mathbb{N};\ x_1,x_2\in\mathfrak{X}). \tag{13}$$

В таком общем виде определение (13) ещё слишком «сыро́»; естественно, в зависимости от класса рассматриваемых ПВП, нужно наделить вложение φ дополнительными свойствами (в основном, непрерывностью в той или иной топологии). Например, в случае банаховых ПВП естественно требовать *гомеоморфности* φ относительно *сильных* топологий, что далее будем считать выполненным.

Приведём конструкцию полуунитарной дилатации для операторов из категории $\mathbf{B^{(mutadj)}}$ (самой широкой для описываемой ниже конструкции). Пусть T – оператор, – стрелка в $\mathbf{B^{(mutadj)}}$ с началом и концом в $\mathfrak{X}$, – для которого его (самосопряжённый) *дефект по полуунитарности* $\delta_T := I - T^\sharp T$ – *ненулевой* (т.е. T *не является полуунитарным*); причём δ_T обладает сопрягаемым квадратичным расщеплением $C \in \mathrm{B}(\mathfrak{X},\mathfrak{Y})$ с ограниченным сопряжённым: $C^\sharp \in \mathrm{B}(\mathfrak{Y},\mathfrak{X})$. Если T является стрелкой некоторой подкатегории $\mathbf{B^{(mutadj)}}$ с квадратичным расщеплением (например, $\mathbf{RegB^{(adj)}}$ или $\mathbf{Kr}$), то такие требования излишни, однако в общем случае мы вынуждены ограничиться специальным классом операторов T в $\mathfrak{X}$. Пусть $\mathfrak{G}$ – некоторый объект в $\mathbf{B^{(mutadj)}}$ (т.е. банахово ПВП с мажорантой), и ψ – *изометрическое* относительно внутреннего произведения и *гомеоморфное* вложение $\mathfrak{Y}$ в $\mathfrak{G}$, обладающее следующим свойством:

$$\psi \text{ – \textit{сопрягаемый оператор}}, \tag{14}$$

откуда автоматически следует, что $\psi^\sharp \in \mathrm{B}(\mathfrak{G},\mathfrak{Y})$, ибо в данном случае, как показывают простые выкладки, $\psi^\sharp = \psi^{-1} \mid \mathrm{R}(\psi)$.

Пусть W – стрелка в $\mathbf{B}^{(\mathbf{mutadj})}$ с началом и концом в $\mathfrak{G}$ $\big($т.е. $W \in \mathrm{B}(\mathfrak{G})$, W сопрягаем и $W^\sharp \in \mathrm{B}(\mathfrak{G})\big)$, являющаяся полуунитарным оператором со свойством:

$$\psi\mathfrak{Y} = \ker W^\sharp\,. \tag{15}$$

Пусть $\mathscr{C}$ – распространение на $\mathfrak{G}$ оператора C:

$$\mathscr{C} := \psi C. \tag{16}$$

Тогда из (15) и (16) тривиально следует равенство

$$W^\sharp\mathscr{C} = 0. \tag{17}$$

Так как ψ – изометрическое (относительно внутреннего произведения) и гомеоморфное вложение, то, как легко проверить, $\mathscr{C} \in \mathrm{B}(\mathfrak{X},\mathfrak{G})$, $\mathscr{C}$ – сопрягаем, $\mathscr{C}^\sharp \in \mathrm{B}(\mathfrak{G},\mathfrak{X})$, и $\mathscr{C}$ является квадратичным расщеплением дефекта δ_T:

$$\mathscr{C}^\sharp\mathscr{C} = \delta_T\,. \tag{18}$$

Пусть $\mathfrak{K}$ – внешняя ортогональная сумма $\mathfrak{X}$ и $\mathfrak{G}$, φ – естественное изометрическое (относительно внутреннего произведения) и гомеоморфное вложение $\mathfrak{X}$ в $\mathfrak{K}$. Рассмотрим оператор V, заданный матрицей

$$V := \begin{bmatrix} T & 0 \\ \mathscr{C} & W \end{bmatrix}. \tag{19}$$

Индукцией легко доказывается формула

$$V^n = \begin{bmatrix} T^n & 0 \\ \sum\limits_{k=0}^{n-1} W^k C T^{n-1-k} & W^n \end{bmatrix} \ (n \in \mathbb{N}). \tag{20}$$

Покажем теперь, что V – полуунитарная дилатация T. Действительно, оператор V – полуунитарный:

$$V^\sharp V = \begin{bmatrix} T^\sharp & \mathscr{C}^\sharp \\ 0 & W^\sharp \end{bmatrix} \begin{bmatrix} T & 0 \\ \mathscr{C} & W \end{bmatrix} = \begin{bmatrix} T^\sharp T + \mathscr{C}^\sharp\mathscr{C} & \mathscr{C}^\sharp W \\ W^\sharp\mathscr{C} & W^\sharp W \end{bmatrix} \overset{(17),(18)}{=} \begin{bmatrix} I_{\mathfrak{X}} & 0 \\ 0 & I_{\mathfrak{G}} \end{bmatrix} = I_{\mathfrak{K}}$$

и является дилатацией T:

$$[V^n\varphi x_1, \varphi x_2]_{\mathfrak{K}} \overset{(20)}{=} [\begin{bmatrix} T^n & 0 \\ \sum\limits_{k=0}^{n-1} W^k\mathscr{C}T^{n-1-k} & W^n \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix}, \begin{bmatrix} x_2 \\ 0 \end{bmatrix}]_{\mathfrak{K}} =$$

$$= [\begin{bmatrix} T^n x_1 \\ \sum\limits_{k=0}^{n-1} W^k\mathscr{C}T^{n-1-k}x_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ 0 \end{bmatrix}]_{\mathfrak{K}} = [T^n x_1, x_2]_{\mathfrak{X}} \quad (n \in \mathbb{N};\ x_1, x_2 \in \mathfrak{X})\,.$$

Дилатация V оператора T (13) называется *минимальной* (см. начало этого пункта и список источников там), если ортогональное дополнение линейной оболочки множеств $V^n\varphi\mathfrak{X}$ $(n \in \mathbb{N}_+)$ – нулевое[4].

[4]Это – косвенное определение. См. сноску на стр. 81.

Важность условия (4) и понятия одностороннего сдвига при работе с полуунитарной дилатацией показывает следующая теорема.

Теорема 1. *Дилатация* V *оператора* T (15), (16), (19) *минимальна тогда и только тогда, когда выполняются условия: ядро сопряжённого к оператору* C, *порождающему* $\mathscr{C}$ *– нулевое, и* W *– односторонний сдвиг.*

Доказательство. Пусть $g \in \mathfrak{G}$. Тогда справедливо равенство (см. чуть выше последнюю цепочку равенств):

$$[V^n \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ g \end{bmatrix}]_{\mathfrak{K}} = \sum_{k=0}^{n-1} [W^k \mathscr{C} T^{n-1-k} x, g]_{\mathfrak{G}} \,. \tag{21}$$

Необходимость. **I**. Пусть $y \in \ker C^\sharp$. Тогда $\mathscr{C}^\sharp \psi y \overset{(14),(16)}{=} C^\sharp \psi^\sharp \psi y = C^\sharp y = 0$. Отсюда

$$[V^n \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \psi y \end{bmatrix}]_{\mathfrak{K}} \overset{(21)}{=} \sum_{k=0}^{n-1} [\mathscr{C} T^{n-1-k} x, W^{\sharp k} \psi y]_{\mathfrak{G}} \overset{(15)}{=}$$
$$= [\mathscr{C} T^{n-1} x, \psi y]_{\mathfrak{G}} = [T^{n-1} x, \mathscr{C}^\sharp \psi y]_{\mathfrak{X}} = 0 \quad (x \in \mathfrak{X}).$$

Из последней цепочки по определению минимальной дилатации вытекает равенство $\psi y = 0$, откуда $y = 0$. Итак, показано, что $\ker C^\sharp = \{0\}$.

II. Пусть $g \in \bigcap\limits_{n \in \mathbb{N}} W^n \mathfrak{G}$. Тогда существует такая последовательность $\{g_m\}_{m \in \mathbb{N}}$ из $\mathfrak{G}$, что $g = W^m g_m$. Имеем цепочку:

$$[V^n \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ g \end{bmatrix}]_{\mathfrak{K}} \overset{(21)}{=} \sum_{k=0}^{n-1} [W^k \mathscr{C} T^{n-1-k} x, W^m g_m]_{\mathfrak{G}} =$$
$$= \sum_{k=0}^{n-1} [W^{\sharp(m-k)} \mathscr{C} T^{n-1} x, g_m]_{\mathfrak{G}} \overset{(17)}{=} 0 \quad (x \in \mathfrak{X},\ n \in \mathbb{N},\ m \geqslant n).$$

Тогда по определению минимальной дилатации получаем равенство $g = 0$. Итак, показано, что W – односторонний сдвиг.

Достаточность. Пусть вектор $\begin{bmatrix} x_0 \\ g_0 \end{bmatrix}$ ортогонален каждому из линеалов $V^n \mathfrak{K}$ $(n \in \mathbb{N}_+)$. Тогда, в частности, будем иметь:

$$0 = [V^0 \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} x_0 \\ g_0 \end{bmatrix}]_{\mathfrak{K}} = [\begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} x_0 \\ g_0 \end{bmatrix}]_{\mathfrak{K}} = [x, x_0]_{\mathfrak{X}} \quad (x \in \mathfrak{X}).$$

Таким образом, $x_0 = 0$. Далее, по (21) при $n = 1$

$$0 = [V^1 \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ g_0 \end{bmatrix}]_{\mathfrak{K}} = [\mathscr{C} x, g_0]_{\mathfrak{G}} \quad (x \in \mathfrak{X}).$$

Покажем при помощи индукции, что

$$[W^n \mathscr{C} x, g_0]_{\mathfrak{G}} = 0 \quad (n \in \mathbb{N}_+,\ x \in \mathfrak{X}) \,. \tag{22}$$

Базис индукции уже обоснован. Пусть $n > 0$ и

$$\forall\, x \in \mathfrak{X} \quad \forall\, k \in \overline{0, n} \quad [W^k \mathscr{C} x, g_0]_{\mathfrak{G}} = 0. \tag{23}$$

Тогда

$$0 = [V^{n+2} \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ g_0 \end{bmatrix}]_{\mathfrak{K}} \overset{(21)}{=} \sum_{k=0}^{n+1} [W^k \mathscr{C} T^{n+1-k} x, g_0]_{\mathfrak{G}} \overset{(23)}{=} [W^{n+1} \mathscr{C} x, g_0]_{\mathfrak{G}} \quad (x \in \mathfrak{X}).$$

Индуктивный переход обоснован, и (22) доказано. Далее, в силу обратимости $C^\sharp$ (по условию теоремы)

$$\mathrm{R}(\mathscr{C})^{[\perp]} = \ker \mathscr{C}^\sharp \overset{(14),(16)}{=} \ker C^\sharp \psi^\sharp = \ker \psi^\sharp = \mathrm{R}(\psi)^{[\perp]} \overset{(15)}{=} \left(\ker W^\sharp\right)^{[\perp]},$$

откуда, в свою очередь, получим цепочку

$$W^{\sharp n} g_0 \overset{(22)}{\in} \mathrm{R}(\mathscr{C})^{[\perp]} = (\ker W^\sharp)^{[\perp]} \ \sim\ g_0 \in (W^n \ker W^\sharp)^{[\perp]} \quad (n \in \mathbb{N}_+)\,. \tag{24}$$

И, наконец,

$$\begin{gathered} \forall\, n \in \mathbb{N}_+ \quad g_0 \in (W^n \ker W^\sharp)^{[\perp]} \overset{(9)}{\Leftrightarrow} \forall\, n \in \mathbb{N} \quad g_0 \in (\ker W^{\sharp n})^{[\perp]} \overset{(8)}{\Leftrightarrow} \\ \Leftrightarrow\ \forall\, n \in \mathbb{N} \quad g_0 \in \mathrm{R}(W^n) \ \Leftrightarrow\ g_0 \in \bigcap_{n \in \mathbb{N}} W^n \mathfrak{G}\,. \end{gathered} \tag{25}$$

Так как по условию теоремы W – односторонний сдвиг, то согласно (22), (24), (25) $g_0 = 0$. Итак, показано, что вектор $\begin{bmatrix} x_0 \\ g_0 \end{bmatrix}$, ортогональный каждому из линеалов $V^n \mathfrak{K}\ (n \in \mathbb{N}_+)$ – нулевой, т.е. дилатация V – минимальна. $\square$

✓ **Осуществимость конструкции.** Конструкция полуунитарной дилатации с условиями (14), (15) легко осуществима (при наличии квадратичного расщепления дефекта в общем случае), если задействовать правосторонний сдвиг в $\ell_2(\cdot)$–пространстве. Действительно, пусть $\mathfrak{G} := \ell_2(\mathfrak{Y})$, ψ – естественное вложение $\mathfrak{Y}$ в $\mathfrak{G}$:

$$\psi y := (y, 0, 0, \ldots) \quad (y \in \mathfrak{Y}),$$

(изометрическое по внутреннему произведению и норме), а W – правосторонний сдвиг (10). Тогда свойства (14), (15) легко проверяются (см. (12)).

Формулировка и доказательство основного результата

Положим

$$\mathscr{N} := \bigcup_{n \in \mathbb{N}} \overline{0, n-1} \times \{n\}$$

(т.е. $\mathscr{N}$ есть совокупность всех пар $\langle k, n \rangle$, $n \in \mathbb{N}$, $0 \leqslant k < n$ – «бесконечный нижний треугольник» декартова квадрата $\mathbb{N}^2$).

Замечание 3. В этом подразделе мы будем считать, что односторонний сдвиг W, используемый в конструкции полуунитарной дилатации, является *проекционно устойчивым.* Это, снижая общность, не снижает силу построений (см. ниже ключевой момент (31)); наоборот, согласно рассуждениям пункта ✓, для любого оператора T, имеющего соответствующее квадратичное расщепление дефекта, всегда можно построить полуунитарную дилатацию «максимальной силы».

Рассмотрим множество

$$\mathscr{X}_V(x) := \left\{\mathfrak{h} \in \prod_{\langle k,n\rangle \in \mathscr{N}} \left(T^{k+1}\right)^{-1}\bigl(\{x\}\bigr) \;\Big|\; \Bigl\{\sum_{k=0}^{n-1} W^k C\mathfrak{h}\bigl(\langle k,n\rangle\bigr)\Bigr\}_{n\in\mathbb{N}} \text{— \textit{сильно сходится}}\right\}$$

(иначе говоря, $\mathfrak{h}$ представляет собой двойную последовательность «треугольного вида» $\{h_{kn}\}_{\substack{n\in\mathbb{N}\\ k\in\overline{0,n-1}}}$, в которой $T^{k+1}h_{kn} = x$). Теперь определим множество

$$\mathscr{X}_V^{nd} := \Bigl\{x \in \bigcap_{n\in\mathbb{N}} T^n\mathfrak{X} \mid \mathscr{X}_V(x) \neq \varnothing\Bigr\}.$$

Предложение 4. Множество $\mathscr{X}_V^{nd}$ является линейным многообразием.

Доказательство. Множество $\bigcap\limits_{n\in\mathbb{N}} T^n\mathfrak{X}$ является линейным многообразием. Далее, совокупность всех двойных последовательностей элементов из $\mathfrak{X}$ «треугольного вида», очевидно, является линейным многообразием относительно обычных операций сложения и умножения на число. При этом, если $\widetilde{0}$ – нулевая последовательность, то $T^{k+1}\widetilde{0}\bigl(\langle k,n\rangle\bigr) = 0$, поэтому $0 \in \mathscr{X}_V^{nd}$. Далее, если $T^{k+1}\mathfrak{h}_1\bigl(\langle k,n\rangle\bigr) = x_1$ и $T^{k+1}\mathfrak{h}_2\bigl(\langle k,n\rangle\bigr) = x_2$, то для произвольных $\alpha_1, \alpha_2 \in \mathbb{C}$

$$\begin{aligned} T^{k+1}\Bigl((\alpha_1\mathfrak{h}_1 + \alpha_2\mathfrak{h}_2)\bigl(\langle k,n\rangle\bigr)\Bigr) &= T^{k+1}\Bigl(\alpha_1\mathfrak{h}_1\bigl(\langle k,n\rangle\bigr)\Bigr) + T^{k+1}\Bigl(\alpha_2\mathfrak{h}_2\bigl(\langle k,n\rangle\bigr)\Bigr) = \\ &= \alpha_1 T^{k+1}\mathfrak{h}_1\bigl(\langle k,n\rangle\bigr) + \alpha_2 T^{k+1}\mathfrak{h}_2\bigl(\langle k,n\rangle\bigr) = \alpha_1 x_1 + \alpha_2 x_2 \quad (n \in \mathbb{N},\;\; k \in \overline{0, n-1}), \end{aligned}$$

причём (предел суммы двух сходящихся последовательностей равен…)

$$\begin{aligned} &s.\lim_{n\to\infty}\sum_{k=0}^{n-1} W^k C(\alpha_1\mathfrak{h}_1 + \alpha_2\mathfrak{h}_2)\bigl(\langle k,n\rangle\bigr) = \\ &= \alpha_1\, s.\lim_{n\to\infty}\sum_{k=0}^{n-1} W^k C\mathfrak{h}_1\bigl(\langle k,n\rangle\bigr) + \alpha_2\, s.\lim_{n\to\infty}\sum_{k=0}^{n-1} W^k C\mathfrak{h}_2\bigl(\langle k,n\rangle\bigr). \end{aligned}$$

Таким образом, линейная комбинация векторов из $\mathscr{X}_V^{nd}$ лежит в $\mathscr{X}_V^{nd}$. □

Основным результатом данной работы является следующая теорема.

Теорема 2. *Остаточное подпространство полуунитарной дилатации V оператора T имеет следующий вид:*

$$\bigcap_{n\in\mathbb{N}} V^n\mathfrak{K} = \left\{ \begin{bmatrix} x \\ g \end{bmatrix} \mid x \in \mathscr{X}_V^{nd},\; g = s.\lim_{n\to\infty}\sum_{k=0}^{n-1} W^k C\mathfrak{h}\bigl(\langle k,n\rangle\bigr),\; \mathfrak{h} \in \mathscr{X}_V(x)\right\} \tag{26}$$

Доказательство. Из (20) следует эквиваленция

$$\begin{bmatrix} x \\ g \end{bmatrix} \in V^n\mathfrak{K} \sim \exists x_n \in \mathfrak{X} \quad \exists g_n \in \mathfrak{G} \qquad \begin{bmatrix} x \\ g \end{bmatrix} = \begin{bmatrix} T^n x_n \\ \sum\limits_{k=0}^{n-1} W^k C T^{n-1-k} x_n + W^n g_n \end{bmatrix} . \tag{27}$$

Согласно (27) определим последовательность $\mathfrak{h}$, полагая

$$\mathfrak{h}\big(\langle k,n\rangle\big) := T^{n-1-k}x_n\,, \quad \langle k,n\rangle \in \mathscr{N}\,. \tag{28}$$

Покажем, что

$$\mathfrak{h} \in \mathscr{X}_V(x)\,. \tag{29}$$

Согласно правой части последнего равенства в (27) (первая строка столбца)

$$T^{k+1}\mathfrak{h}\big(\langle k,n\rangle\big) \overset{(28)}{=} T^{k+1}T^{n-k-1}x_n = T^n x_n = x \tag{30}$$

и (вторая строка)

$$W^n W^{\sharp n} g = \sum_{k=0}^{n-1} W^n W^{\sharp(n-k)} C T^{n-1-k} x_n + W^n W^{\sharp n} W^n g_n \overset{(17)}{=} W^n g_n\,, \tag{31}$$

откуда из проекционной устойчивости W следует, что $s.\lim\limits_{n\to\infty} W^n g_n = 0$. Последнее равенство влечёт цепочку

$$g \overset{(27)}{=} s.\lim_{n\to\infty} \sum_{k=0}^{n-1} W^k C T^{n-1-k} x_n \overset{(28)}{=} s.\lim_{n\to\infty} \sum_{k=0}^{n-1} W^k C \mathfrak{h}\big(\langle k,n\rangle\big). \tag{32}$$

Из (32) и (30) непосредственно следует (29). Тогда согласно (27) – (29) справедлива эквиваленция

$$\begin{bmatrix} x \\ g \end{bmatrix} \in \bigcap_{n\in\mathbb{N}} V^n\mathfrak{K} \sim \Big(x \in \bigcap_{n\in\mathbb{N}} T^n\mathfrak{X}\Big) \& \Big(\exists\, \mathfrak{h} \in \mathscr{X}_V(x) \quad g = s.\lim_{n\to\infty} \sum_{k=0}^{n-1} W^k C \mathfrak{h}\big(\langle k,n\rangle\big)\Big),$$

которая равносильна равенству (26). □

В случае обратимости оператора T, вид остаточного подпространства его полуунитарной дилатации V заметно упрощается (в частности, отпадает необходимость рассмотрения промежуточных множеств $\mathscr{X}_V(x)$).

Следствие 1. *Пусть T – обратимый оператор. Тогда*

$$\mathscr{X}_V^{nd} = \Big\{x \in \bigcap_{n\in\mathbb{N}} T^n\mathfrak{X} \;\Big|\; \sum_{k=0}^{\infty} W^k C T^{-(k+1)} x \text{ — (сильно) сходится}\Big\}\,;$$

$$\bigcap_{n\in\mathbb{N}} V^n\mathfrak{K} = \Big\{ \begin{bmatrix} x \\ g \end{bmatrix} \;\Big|\; x \in \mathscr{X}_V^{nd},\ g = \sum_{k=0}^{\infty} W^k C T^{-(k+1)} x \Big\}\,.$$

Доказательство. Из уравнения $T^{k+1}\mathfrak{h}\big(\langle k,n\rangle\big) = x$ однозначно определяется вектор $\mathfrak{h}\big(\langle k,n\rangle\big)$: он равен $T^{-(k+1)}x$. □

Описание остаточного подпространства ещё более упрощается, если дилатация V построена по образцу пункта ✓.

Следствие 2. *Пусть V – полуунитарная дилатация оператора T конструкции пункта ✓. Тогда*

$$\mathscr{X}_V(x) = \left\{ \mathfrak{h} \in \prod_{\langle k,n\rangle \in \mathscr{N}} \left(T^{k+1}\right)^{-1}\left(\{x\}\right) \;\middle|\; \lim_{n\to\infty} \sum_{k=0}^{n-1} \|C\mathfrak{h}\left(\langle k,n\rangle\right)\|^2 < \infty \right\},$$

а в случае обратимости T

$$\mathscr{X}_V^{nd} = \left\{ x \in \bigcap_{n\in\mathbb{N}} T^n \mathfrak{X} \mid \sum_{k=0}^{\infty} \|CT^{-(k+1)}x\|^2 < \infty \right\}.$$

Доказательство. См. предложение 3, теорему 2 и следствие 1. □

Иллюстрации

Замечание 4. Дилатации, рассматриваемые в данном подразделе, имеют конструкцию пункта ✓.

Пример 1. Положим $\mathfrak{X} := \mathbb{C}^2$, и внутреннее произведение на $\mathfrak{X}$ зададим как

$$[(x_1, y_1), (x_2, y_2)]_{\mathfrak{X}} := x_1\overline{y_2} + y_1\overline{x_2}\,. \tag{33}$$

Будем отождествлять операторы с порождающими их матрицами относительно базиса $\{(1,0),(0,1)\}$. Тогда внутреннее произведение (33), очевидно, порождено симметрией $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Наше $\mathfrak{X}$ является объектом $\mathbf{Kr}$ (если, на всякий случай, мыслить о пространстве Крейна с возможностью *конечномерности* носителя и равными положительным и отрицательным индексами инерции).

Рассмотрим оператор $T := \begin{bmatrix} 0 & \alpha \\ 0 & \beta \end{bmatrix}$. Прямыми вычислениями находим степень и дефект:

$$T^n = \beta^{n-1}T \quad (n \in \mathbb{N}), \qquad \delta_T = \begin{bmatrix} 0 & -2\operatorname{Re}(\overline{\alpha}\beta) \\ 0 & 0 \end{bmatrix}.$$

Чтобы дефект был ненулевым, числа α и β должны удовлетворять условиям:

$$\alpha \neq 0, \quad \beta \neq 0, \quad \operatorname{Re}(\overline{\alpha}\beta) \neq 0,$$

которые и будем далее полагать выполненными (формально, первые два неравенства излишни, так как следуют из третьего).

Легко находится квадратичное расщепление C дефекта δ_T с условием $\ker C^\sharp = \{0\}$. Действительно, положим $\mathfrak{Y} := \mathbb{C}$ и зададим внутреннее произведение:

$$[z_1, z_2]_{\mathfrak{Y}} := -\operatorname{sgn}\left(\operatorname{Re}(\overline{\alpha}\beta)\right) z_1\overline{z_2}\,.$$

Тогда C и его сопряжённый имеют вид:

$$C = \begin{bmatrix} 0 & \sqrt{2|\alpha\beta|} \end{bmatrix}, \qquad C^\sharp = \begin{bmatrix} -\operatorname{sgn}\left(\operatorname{Re}(\overline{\alpha}\beta)\right)\sqrt{2|\alpha\beta|} \\ 0 \end{bmatrix}.$$

Отсюда видно, что пространство $\mathfrak{G}$, в котором реализуется односторонний сдвиг, есть просто ℓ_2 с внутренним произведением равным $\pm$ скалярное произведение в зависимости от знака числа $\mathrm{Re}(\overline{\alpha}\beta)$.

Далее, вполне очевидно, что

$$\bigcap_{n\in\mathbb{N}} T^n\mathfrak{X} = T\mathfrak{X} = \left\{ \begin{bmatrix} z \\ \alpha^{-1}\beta z \end{bmatrix} \mid z \in \mathbb{C} \right\},$$

и уравнение $T^{k+1}\begin{bmatrix} h_{1kn} \\ h_{2kn} \end{bmatrix} = \begin{bmatrix} z \\ \alpha^{-1}\beta z \end{bmatrix}$ даёт решение: h_{1kn} – любое, $h_{2kn} = \alpha^{-1}\beta^{-k}z$, откуда находим:

$$C\begin{bmatrix} h_{1kn} \\ h_{2kn} \end{bmatrix} = \sqrt{2|\alpha\beta|}\,\alpha^{-1}\beta^{-k}z, \qquad \|C\begin{bmatrix} h_{1kn} \\ h_{2kn} \end{bmatrix}\|^2 = 2|\alpha|^{-1}|\beta|^{1-2k}|z|^2. \tag{34}$$

Из (34) следует, что $\lim\limits_{n\to\infty}\sum\limits_{k=0}^{n}\|C\begin{bmatrix} h_{1kn} \\ h_{2kn} \end{bmatrix}\|^2 < \infty$ в том и только том случае, если $|\beta| > 1$ (независимо от z). Таким образом, если $|\beta| \leqslant 1$, то остаточное подпространство минимальной полуунитарной дилатации V оператора T – *нулевое* (т.е. V представляет собой в этом случае односторонний сдвиг). Если же $|\beta| > 1$, то согласно (34) и следствию 2 остаточное подпространство имеет следующее строение:

$$\bigcap_{n\in\mathbb{N}} V^n\mathfrak{K} \quad \textit{натянуто на вектор} \quad \begin{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \\ \sqrt{2|\alpha\beta|}\,(1, \beta^{-1}, \beta^{-2}, \ldots) \end{bmatrix}.$$

Пример 2. Положим теперь $\mathfrak{X} := C[-1,1]$, и определим внутреннее произведение:

$$[x,y]_{\mathfrak{X}} := \int_{-1}^{1} x(-t)\,\overline{y(t)}\,dt \tag{35}$$

(симметричность формы (35) обеспечивается симметричностью интервала $[-1,1]$). По стандартным свойствам интеграла (Римана) выполняется (1) (где в качестве b может выступать любое число $\geqslant 2$), т.е. $\mathfrak{X}$ – *пространство с мажорантой*. Однако $\mathfrak{X}$, как можно догадаться, *не является п.б.п.*: отрицание свойства (2) –

$$\forall\, c > 0 \quad \exists\, x_c \in \mathfrak{X} \quad \sup_{\|y\|\leqslant 1} |[x_c, y]| < c\|x_c\| \tag{36}$$

– легко доказать, рассмотрев, например, функции $x_c(t) := |t|^{\frac{2}{\eta(c)}-1}$, где η – произвольная функция, удовлетворяющая условию $0 < \eta(c) < c \leqslant 2$ (справедливость (36) достаточно доказать, в частности, для $c \leqslant 2$). Действительно, тогда

$$\sup_{\|y\|\leqslant 1} |[x_c, y]| \leqslant \int_{-1}^{1} |x_c(-t)|\,dt = 2\int_{0}^{1} t^{\frac{2}{\eta(c)}-1}\,dt = \eta(c) < c = c\|x_c\|$$

($\|x_c\| = 1$ так как $\frac{2}{\eta(c)} - 1 > 0$ в силу наложенных на η и c условий). Таким образом $\mathfrak{X}$ – объект $\mathbf{B^{(mutadj)}}$, но не является объектом $\mathbf{RegB^{(adj)}}$.

Рассмотрим интегральный оператор $(Tx)(t) := \int\limits_{-1}^{t} x(\tau)\,d\tau$. Оператор T оказывается *самосопряжённым* (!) относительно введённого внутреннего произведения:

$$[Tx,y]_{\mathfrak{X}} = \int\limits_{-1}^{1}\Big(\int\limits_{-1}^{-t} x(\tau)\,d\tau\Big)\overline{y(t)}\,dt = \int\limits_{-1}^{-t} x(\tau)\,d\tau \int\limits_{-1}^{t}\overline{y(\tau)}\,d\tau\,\Bigg|_{-1}^{1} - \int\limits_{-1}^{1}\big(-x(-t)\big)\Big(\int\limits_{-1}^{t}\overline{y(\tau)}\,d\tau\Big)\,dt =$$

$$= \int\limits_{-1}^{1} x(-t)\overline{\Big(\int\limits_{-1}^{t} y(\tau)\,d\tau\Big)}\,dt = [x,Ty]_{\mathfrak{X}}\,,$$

и это позволяет (при помощи трюка) построить квадратичное расщепление его дефекта $\delta_T = I - T^2$. А именно, пусть $\mathfrak{Y}$ – то же пространство $C[-1,1]$, но с внутренним произведением $[x,y]_{\mathfrak{Y}} := -[x,y]_{\mathfrak{X}}$. Обозначим через $\mathscr{T}$ тот же оператор T, но действующий из $\mathfrak{X}$ в $\mathfrak{Y}$. Тогда $\mathscr{T}^\sharp = -\mathscr{T}$:

$$[\mathscr{T}x,y]_{\mathfrak{Y}} := -[Tx,y]_{\mathfrak{X}} = -[x,Ty]_{\mathfrak{X}} = -[x,\mathscr{T}y]_{\mathfrak{X}} \quad (x\in\mathfrak{X},\ y\in\mathfrak{Y})\,.$$

Положим $C := I + \mathscr{T}$. Тогда C – квадратичное расщепление δ_T :

$$C^\sharp C = (I+\mathscr{T})^\sharp(I+\mathscr{T}) = (I-\mathscr{T})(I+\mathscr{T}) = (I-T)(I+T) = I - T^2 = \delta_T\,.$$

Ядро $C^\sharp$ – нулевое. Действительно, если $x_0 \in \ker C^\sharp$, то

$$0 = (I+\mathscr{T})^\sharp x_0 = (I-\mathscr{T})x_0 = (I-T)x_0 \;\Leftrightarrow\; x_0(t) = \int\limits_{-1}^{t} x_0(\tau)\,d\tau \;\Leftrightarrow$$

$$\Leftrightarrow\; x_0'(t) = x_0(t),\ x_0(-1) = 0 \;\Leftrightarrow\; x_0 = 0\,.$$

Далее, нетрудно видеть, что $\bigcap\limits_{n\in\mathbb{N}} T^n\mathfrak{X}$ состоит из всех бесконечно дифференцируемых функций x, удовлетворяющих условию $x^{(n)}(-1) = 0$ $(n\in\mathbb{N}_+)$. Оператор T обратим, и $CT^{-(k+1)}x = (I+T)T^{-(k+1)}x = x^{(k+1)} + x^{(k)}$ для $x \in \bigcap\limits_{n\in\mathbb{N}} T^n\mathfrak{X}$. Тогда $\sum\limits_{k=0}^{\infty}\|CT^{-(k+1)}x\|^2 < \infty$ в том и только том случае, если $\sum\limits_{k=0}^{\infty}\|(x'+x)^{(k)}\|^2 < \infty$. Но для интегрального оператора известно, что $\|T^n\| \to 0,\ n\to\infty$, и учитывая сходимость $\|(x'+x)^{(n)}\| \to 0,\ n\to\infty$, получим цепочку соотношений:

$$\|x'+x\| = \|T^n(x'+x)^{(n)}\| \leqslant \|T^n\|\|(x'+x)^{(n)}\| \to 0,\ n\to\infty,$$

из которой следует равенство $x'+x = 0$. А функция x, удовлетворяющая этому дифференциальному уравнению с начальным условием $x'(-1) = 0$ – нулевая.

Итак, мы выяснили, что для данного случая множество $\mathscr{X}_V^{nd}$ состоит лишь из нуль–вектора (см. следствие 2). Таким образом, остаточное подпространство минимальной полуунитарной дилатации V интегрального оператора T – *нулевое*, и V представляет собой односторонний сдвиг.

Замечание 5. Рассуждения этого пункта без труда переносятся на чуть более общий случай $\mathfrak{X} = C[a,b]$ с внутренним произведением $[x,y]_{\mathfrak{X}} := \int\limits_a^b x(\frac{a+b}{2} - t)\,\overline{y(t)}\,dt$ и оператором $(Tx)(t) := \int\limits_a^t x(\tau)\,d\tau$.

Заключение

Итак, дано конструктивное описание остаточного подпространства полуунитарной дилатации линейного непрерывного оператора, действующего в банаховом пространстве; а точнее, оператора, являющегося стрелкой категории $\mathbf{B}^{(\mathbf{mutadj})}$ и имеющего квадратичное расщепление своего дефекта из $\mathbf{B}^{(\mathbf{mutadj})}$ (см. стр. 82). *Основным результатом* работы является теорема 2 (стр. 86). *Дальнейшие перспективы* исследований в таком направлении – это попытки хотя бы наметить те же вехи в «банаховой области», которые были «пройдены гигантами» ([4, 9]) в «гильбертовой»: построение унитарной дилатации, построение и изучение характеристической функции, построение и изучение модели и пр.

Список литературы

1. Гинзбург Ю. П., Иохвидов И. С. *Исследования по геометрии бесконечномерных пространств с билинейной метрикой* // Успехи мат. наук. – 1962. – Т. 17, Вып. 4. – С. 3–56
2. Ароншайн Р. *Квадратичные формы на векторных пространствах* // Математика (сб. переводов). – 1964. – Т. 8, № 5. – С. 105–168
3. Davis Ch. *J–unitary dilation of a general operator* // Acta Sci. Math. (Szeged). – 1970. – Vol. 31. – P. 75–86
4. Сёкефальви–Надь Б., Фояш Ч. *Гармонический анализ операторов в гильбертовом пространстве.* – М.: Мир, 1970. – 431 с.
5. Davis Ch., Foiaş C. *Operators with bounded characteristic function and their J–unitary dilation* // Acta Sci. Math. (Szeged). – 1971. – Vol. 32. – P. 127–139
6. Штраус В. А. *Некоторые вопросы геометрии и спектральной теории операторов в банаховых пространствах с эрмитовой формой*: Дис... канд. физ.–мат. наук: 01.01.01. – Воронеж, 1972. – 126 с.
7. Bognar J. *Indefinite inner product spaces.* – Berlin: Springer, 1974. – 225 p.
8. McEnnis B. W. *Shifts on indefinite inner product spaces* // Pacific J. Math. – 1979. – Vol. 81. – P. 113–130
9. Сёкефальви–Надь Б. *Унитарные дилатации операторов в гильбертовом пространстве и смежные вопросы*// Рисс Ф., Сёкефальви–Надь Б. *Лекции по функциональному анализу. Добавление 2.* – М.: Мир, 1979. – С. 511–560
10. Никольский Н. К. *Лекции об операторе сдвига.* – М.: Наука, 1980. – 384 с.
11. McEnnis B. W. *Shifts on indefinite inner product spaces. II.* // Pacific J. Math. – 1982. – Vol. 100. – P. 177–183
12. Азизов Т. Я., Иохвидов И. С. *Основы теории линейных операторов в пространствах с индефинитной метрикой.* – М.: Наука, 1986. – 352 с.
13. Никольский Н. К., Хрущёв С. В. *Функциональная модель и некоторые задачи спектральной теории функций* // Труды Математического института АН СССР им. Стеклова. – 1987 – Т. 176. – С. 97–210

14. Constantinescu T., Gheondea A. *On unitary dilations and characteristic functions in indefinite inner product spaces* // Oper. Theory: Adv. Appl. – Vol. 24. – Basel–Boston–Berlin: Birkhäuser, 1987. – P. 87–102
15. Bruinsma P., Dijksma A., de Snoo H. S. V. *Unitary dilations of contractions in* Π_κ // Oper. Theory: Adv. Appl. – Vol. 28. – Basel–Boston–Berlin: Birkhäuser, 1988. – P. 27–42
16. Штраус В. А. *Модельное представление и функциональное исчисление операторов в пространствах с индефинитной метрикой*: Вар–нт дисс... д–ра физ.–мат. наук: 01.01.01. –Челябинск, 1993. – 363 с.
17. Кутателадзе С. С. *Основы функционального анализа.* – 4–е изд., испр. – Новосибирск: Изд–во Ин–та математики, 2001. – xii+354 с.
18. Rovnyak J. *Methods of Krein space operator theory* // Oper. Theory: Adv. Appl. – Vol. 134. – Basel–Boston–Berlin: Birkhäuser, 2002. – P. 31–66
19. Тышкевич Д. Л. *Элементарные ротации операторов в категориях с квадратичным расщеплением* // Таврический Вестник Математики и Информатики (ТВИМ). – 2004, Вып. 1. – С. 112–124
20. Tyshkevich D. L. *Elementary rotation of a semiunutary operator in regular Banach spaces* // Fundamental and Applied Mathematics. – 2006. – vol. 12, №6. – P. 175–192
21. Тышкевич Д. Л. *О разложении Вольда полуунитарного оператора в банаховых пространствах с индефинитной метрикой* // Учёные записки Таврического национального университета им. В. И. Вернадского. – 2006. – Т.19(58), №1. – С. 98–124
22. Тышкевич Д. Л. *Об ортогонализации систем векторов и разложении типа Вольда в линейных пространствах с внутренним произведением*: Дис... канд. физ.–мат. наук: 01.01.01. – Харьков, 2008. – 187 с.

Статья поступила в редакцию 25.12.2008

УДК 519.8

ОЦІНЮВАННЯ ЗА ЗАШУМЛЕННИМИ СПОСТЕРЕЖЕННЯМИ НЕВІДОМИХ ДАНИХ ЛІНІЙНИХ ЕЛІПТИЧНИХ РІВНЯНЬ, ЩО ДОПУСКАЮТЬ ЗМІШАНЕ ВАРІАЦІЙНЕ ФОРМУЛЮВАННЯ

© Горбатенко М.Ю.

Чернівецький національний університет ім.Ю.Федьковича
факультет прикладної математики
вул. Університетська 28, м. Чернівці, 58000, Україна
e-mail: *Mikola.Gorbatenko@gmail.com*

Abstract. We obtain a new class of systems of variational equations via whose solutions the minimax estimates of values of functionals from unknown right-hand sides of the second order linear elliptic equations are expressed.

Вступ

Більшість результатів в галузі мінімаксного оцінювання була одержана з використанням традиційної постановки відповідних варіаційних крайових задач, доведення існування і єдиності розв'язків яких істотно спирається на відому лему Лакса–Мільграма (див. [1] і вказану там літературу).

Такий підхід дозволяє, наприклад, в стаціонарних задачах теплопровідності оцінювати невідомий розподіл щільності джерел за спостереженнями температури. Однак не менший інтерес являє собою також задача оцінювання цього розподілу за спостереженнями теплового потоку. Відмітимо, що відомі на цей час методи оцінювання не дозволяють розв'язувати подібні задачі.

В роботі запропоновано новий метод, який дає змогу оцінювати невідомий розподіл щільністі джерел, як за спостереженнями температури, так і за спостереженнями теплового потоку.

Розроблений метод спирається на змішані варіаційні постановки, започатковані в роботах І. Бабушки і Ф. Брецці (див. [12]).

Названі задачі оцінювання мають важливе прикладне значення в багатьох галузях, тому їх теоретичний аналіз є актуальним.

1. Допоміжні факти

В роботі використовуються наступні позначення:

H – гільбертовий простір над $\mathbb{R}$ із скалярним добутком $(\cdot,\cdot)_H$ і нормою $\|\cdot\|_H$;

$J_H \in \mathscr{L}(H,H')$ – оператор, що називається ізометричним ізоморфізмом, який діє з H на його спряжений простір H' та визначається рівністю[1] $(v,u)_H = < v, J_H u >_{H\times H'}$ $\forall u,v \in H$, де $< x, f >_{H\times H'} := f(x)$ для $x \in H$, $f \in H'$;

$x = (x_1,\dots,x_n)$ – просторова змінна, що змінюється в обмеженій відкритій множині $D \subset \mathbb{R}^n$, з ліпшицевою границею Γ; $dx = dx_1 \cdots dx_n$ – міра Лебега в $\mathbb{R}^n$;

$L^2(D)$ – простір функцій, сумовних з квадратом в області D;

[1] Цей оператор існує в силу теореми Рісса.

$H^k(D)$ і $H^k_0(D)$ – стандартні простори Соболєва цілого порядку $k > 0$ в області D з відповідною нормою;

$\mathbf{grad}\, p := (\frac{\partial p}{\partial x_1}, \dots, \frac{\partial p}{\partial x_n})^T$; $\mathrm{div}\,\mathbf{v} := \sum_{i=1}^n \frac{\partial v_i}{\partial x_i}$;[2]

$H(\mathrm{div}; D) := \{\mathbf{v} \in L^2(D)^n, \mathrm{div}\,\mathbf{v} \in L^2(D)\}$ – гільбертів простір з нормою $\|\mathbf{v}\|_{H(\mathrm{div};D)} := \{\|\mathbf{v}\|^2_{L^2(D)^n} + \|\mathrm{div}\,\mathbf{v}\|^2_{L^2(D)}\}^{1/2}$;

Позначимо через $L^2(\Omega, H)$ простір Бохнера, що складається з випадкових елементів $\xi = \xi(\omega)$, визначених на деякому ймовірнісному просторі $(\Omega, \mathscr{B}, P)$ зі значеннями в H таких, що

$$\|\xi\|^2_{L^2(\Omega,H)} = \int_\Omega \|\xi(\omega)\|^2_H dP(\omega) < \infty. \tag{1}$$

В цьому випадку існує інтеграл Бохнера $\mathbb{E}\xi := \int_\Omega \xi(\omega)\, dP(\omega) \in H$, що називається математичним сподіванням або середнім випадкового елемента $\xi(\omega)$, що задовольняє умову

$$(h, \mathbb{E}\xi)_H = \int_\Omega (h, \xi(\omega))_H\, dP(\omega) \quad \forall h \in H. \tag{2}$$

Застосовуючи до випадкової величини ξ це визначення приводить до традиційного означення її математичного сподівання, оскільки інтеграл Бохнера (1) переходить у звичайний інтеграл Лебега по ймовірнісній мірі $dP(\omega)$. У $L^2(\Omega, H)$ можна ввести скалярний добуток:

$$(\xi, \eta)_{L^2(\Omega,H)} := \int_\Omega (\xi(\omega), \eta(\omega))_H\, dP(\omega) \quad \forall \xi, \eta \in L^2(\Omega, H). \tag{3}$$

Використовуючи знак математичного сподівання для випадкових величин, рівності (1)–(3) можна записати у вигляді:

$$\|\xi\|^2_{L^2(\Omega,H)} = \mathbb{E}\|\xi(\omega)\|^2_H, \tag{4}$$

$$(h, \mathbb{E}\xi)_H = \mathbb{E}(h, \xi(\omega))_H \quad \forall h \in H, \tag{5}$$

$$(\xi, \eta)_{L^2(\Omega,H)} := \mathbb{E}(\xi(\omega), \eta(\omega))_H \quad \forall \xi, \eta \in L^2(\Omega, H). \tag{6}$$

Простір $L^2(\Omega, H)$, оснащений нормою (4) і скалярним добутком (6), є гільбертовим.

Постановка задачі оцінювання. Нехай стан системи характеризується функцією $\varphi(x)$, яка визначаються як узагальнений розв'язок крайової задачі Діріхле:

$$-\mathrm{div}\,(\mathbf{A}\,\mathbf{grad}\,\varphi) = f \quad \text{в} \quad D, \tag{7}$$

$$\varphi = 0 \quad \text{на} \quad \Gamma, \tag{8}$$

яку, увівши змінну $\mathbf{j} = -\mathbf{A}\,\mathbf{grad}\,\varphi$, можна записати у вигляді еквівалентної системи першого порядку:

$$\mathbf{A}^{-1}\mathbf{j} = \mathbf{grad}\,\varphi \quad \text{в} \quad D, \tag{9}$$

$$\mathrm{div}\,\mathbf{j} = f \quad \text{в} \quad D, \quad \varphi = 0 \quad \text{на} \quad \Gamma, \tag{10}$$

де $\mathbf{A} = \mathbf{A}(x) = (a_{ij}(x))$ симетрична $n \times n$-матриця з елементами $a_{ij} \in L^\infty(D)$, для якої існують такі додатні числа μ_1 і μ_2, що виконується нерівність

[2]Частинні похідні, що входять до виразів $\mathbf{grad}\, p$ і $\mathrm{div}\,\mathbf{v}$ слід розуміти у сенсі розподілів у D.

$\mu_1 \sum_{i=1}^{n} \xi_i^2 \leq \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \leq \mu_2 \sum_{i=1}^{n} \xi_i^2 \quad \forall x \in D, \xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$, через $\mathbf{A}^{-1}$ позначена, обернена до $\mathbf{A}$. У відповідності з [12], під узагальненим розв'язком задачі (9)–(10) будемо розуміти пару функцій $(\mathbf{j}, \varphi) \in H(\mathrm{div}; D) \times L^2(\Omega)$, що задовольняє співвідношенням

$$\int\limits_D (\mathbf{A}^{-1}(x)\mathbf{j}(x), \mathbf{q}(x))_{\mathbb{R}^n} dx + \int\limits_D \varphi(x) \mathrm{div}\, \mathbf{q}(x)\, dx = 0 \quad \forall \mathbf{q} \in H(\mathrm{div}; D) \tag{11}$$

$$\int\limits_D v \mathrm{div}\, \mathbf{j}(x)\, dx = \int\limits_D f(x) v(x)\, dx \quad \forall v \in L^2(D). \tag{12}$$

Зауважимо, що із (11) і (12) випливає $\varphi \in H_0^1(D)$.

З фізичної токи зору, крайова задача (4)–(6), або еквівалента до неї задача (9)–(10), моделює усталений процесс розповсюдження тепла в області D, при цьому функції $\varphi(x)$, $A^{-1}(x)\mathbf{j}(x)$ і $f(x)$ відповідно мають смисл температури, теплового потоку і об'ємної щільності теплових джерел в точці x.

Відзначимо ще, що для знаходження узагальненого розв'язку в [12] на базі так званого змішаного методу скінченних елементів, розроблені ефективні чисельні алгоритми.

Вважається, що функція $f(x)$ у рівняннях (10) і (12) – невідома і належить множині

$$G_0 := \left\{ \tilde{f} \in L^2(D) : \Big(Q(\tilde{f} - f_0), \tilde{f} - f_0 \Big)_{L^2(D)} \leq 1 \right\}, \tag{13}$$

де $f_0 \in L^2(D)$ – задана функція, $Q : L^2(D) \to L^2(D)$ – неперервний додатньо визначений самоспряжений оператор, обернений для якого обмежений.

Задача, що досліджується в даній роботі, полягає в тому, щоб за спостереженням випадкових елементів вигляду

$$y_1(\mathbf{j}; \eta_1) = C_1 \mathbf{j} + \eta_1, \quad y_2(\varphi; \eta_2) = C_2 \varphi + \eta_2, \tag{14}$$

що належать сепарабельним гільбертовим просторам H_1 і H_2 над $\mathbb{R}$ відповідно, оцінити значення лінійного функціоналу

$$l(f) := \int\limits_D l_0(x) f(x)\, dx \tag{15}$$

в класі оцінок вигляду

$$\widehat{l(f)} := (y_1(\mathbf{j}; \eta_1), u_1)_{H_1} + (y_2(\varphi; \eta_2), u_2)_{H_2} + c, \tag{16}$$

де $(\mathbf{j}, \varphi)$ – невідомий узагальнений розв'язок задачі (9)–(10), l_0 – заданий елемент із $L^2(D)^n$ і $L^2(D)$, $u_1 \in H_1$, $u_2 \in H_2$, $c \in \mathbb{R}$, $C_1 \in \mathscr{L}(L^2(D)^n, H_1)$ і $C_2 \in \mathscr{L}(L^2(D), H_2)$– лінійні неперервні оператори, $(\eta_1, \eta_2) \in G_1$, а через G_1 позначено множину випадкових елементів $\tilde{\eta}_1 = \tilde{\eta}_1(\omega) \in L^2(\Omega, H_1)$ і $\tilde{\eta}_2 = \tilde{\eta}_2(\omega) \in L^2(\Omega, H_2)$ з нульовими середніми, що задовольняють умову

$$\mathbb{E}(\tilde{Q}_1 \tilde{\eta}_1, \tilde{\eta}_1)_{H_1} + \mathbb{E}(\tilde{Q}_2 \tilde{\eta}_2, \tilde{\eta}_2)_{H_2} \leq 1, \tag{17}$$

в якій $\tilde{Q}_1$ і $\tilde{Q}_2$ – задані в H_1 і H_2 обмежені самоспряжені додатньо визначені оператори, що мають обмежені обернені.

Означення 1. Оцінку вигляду

$$\widehat{\widehat{l(f)}} = (y_1(\mathbf{j};\eta_1), \hat{u}_1)_{H_1} + (y_2(\varphi;\eta_2), \hat{u}_2)_{H_2} + \hat{c} \tag{18}$$

будемо називати мінімаксною оцінкою $l(f)$, якщо елементи $\hat{u}_1 \in H_1$, $\hat{u}_2 \in H_2$ і число $\hat{c}$ визначаються із умови

$$\sup_{\tilde{f}\in G_0,(\tilde{\eta}_1,\tilde{\eta}_2)\in G_1} \mathbb{E}|l(\tilde{f}) - \widehat{l(\tilde{f})}|^2 \to \inf_{u_1\in H_1,u_2\in H_2,c\in\mathbb{R}}, \tag{19}$$

де $\widehat{l(\tilde{f})} := (y_1(\tilde{\mathbf{j}};\tilde{\eta}_1), u_1)_{H_1} + (y_2(\tilde{\varphi};\tilde{\eta}_2), u_2)_{H_2} + c$, $(\tilde{\mathbf{j}}, \tilde{\varphi})$ – розв'язок задачі (9)–(10) при $f(x) = \tilde{f}(x)$. Величину $\varrho := \{\mathbb{E}|l(\tilde{f}) - \widehat{\widehat{l(\tilde{f})}}|^2\}^{1/2}$ будемо називати похибкою мінімаксного оцінювання виразу $l(f)$.

2. Представлення для мінімаксних оцінок і похибок оцінювання

Введемо до розгляду, при фіксованому $u := (u_1, u_2) \in H_1 \times H_2 := H$, пару функцій $(\mathbf{z}_1(\cdot;u), z_2(\cdot;u)) \in H(\mathrm{div};D) \times L^2(D)$, як єдиний розв'язок наступної крайової задачі:

$$\mathbf{A}^{-1}\,\mathbf{z}_1(\cdot;u) = \mathbf{grad}\, z_2(\cdot;u) - C_1^t J_{H_1} u_1 \quad \text{в} \quad D, \tag{20}$$

$$\mathrm{div}\,\mathbf{z}_1(\cdot;u) = -C_2^t J_{H_2} u_2 \quad \text{в} \quad D, \tag{21}$$

$$z_2(\cdot;u) = 0 \quad \text{на} \quad \Gamma, \tag{22}$$

під яким слід розуміти розв'язок варіаційної задачі

$$\int_D (\mathbf{A}^{-1}(x)\,\mathbf{z}_1(x;u), \mathbf{q}(x))_{\mathbb{R}^n} dx + \int_D z_2(x;u)\mathrm{div}\,\mathbf{q}(x)\,dx =$$
$$= -\int_D ((C_1^t J_{H_1} u_1)(x), \mathbf{q}(x))_{\mathbb{R}^n} dx \quad \forall \mathbf{q} \in H(\mathrm{div}, D), \tag{23}$$

$$\int_D v(x)\mathrm{div}\,\mathbf{z}_1(x;u)\,dx = -\int_D (C_2^t J_{H_2} u_2)(x) v(x)\,dx \quad \forall v \in L^2(D), \text{в} \quad D, \tag{24}$$

де $C_1^t : H_1' \to L^2(D)^n$ і $C_2^t : H_2' \to L^2(D)$ – оператори, транспоновані до C_1 і C_2, що визначаються співвідношеннями $\int_D (v(x), C_1^t w(x))_{\mathbb{R}^n} dx =< Cv, w >_{H_1\times H_1'}$ для всіх $v \in L^2(D)^n$, $w \in H_1'$ і $\int_D v(x) C_2^t w(x)\,dx =< Cv, w >_{H_2\times H_2'}$ для всіх $v \in L^2(D)$, $w \in H_2'$. Із (23) і (24) маємо $z_2 \in H_0^1(D)$.

З теорії варіаційних задач, які допускають змішане формулювання (див., наприклад, [12]), випливає, що функції $\mathbf{z}_1(x;u_1,u_2), z_2(x;u_1,u_2)$ визначаються із рівнянь (23)– (24) єдиним чином.

Теорема 1. Задача знаходження мінімаксної оцінки виразу $l(f)$ еквівалентна задачі оптимального керування системою, що описується варіаційною задачею (23) – (24) з функцією вартості виду

$$I(u) = \big(Q^{-1}(l_0 + z_2(\cdot;u)), l_0 + z_2(\cdot;u)\big)_{L^2(D)} + \\ + (\tilde{Q}_1^{-1}u_1, u_1)_{H_1} + (\tilde{Q}_2^{-1}u_2, u_2)_{H_2} \to \inf_{u\in H}. \quad (25)$$

Доведення. В наслідок (14)–(16) маємо при $u \in H$

$$l(\tilde{f}) - \widehat{l(\tilde{f})} = (l_0, \tilde{f})_{L^2(D)} - (y_1(\mathbf{j};\eta_1), u_1)_{H_1} - (y_2(\varphi;\eta_2), u_2)_{H_2} - c =$$
$$= (l_0, \tilde{f})_{L^2(D)} - (u_1, C_1\tilde{\mathbf{j}} + \tilde{\eta}_1)_{H_1} - (u_2, C_2\tilde{\varphi} + \tilde{\eta}_2)_{H_2} - c =$$
$$= (l_0, \tilde{f})_{L^2(D)} - < J_{H_1}u_1, C_1\tilde{\mathbf{j}} >_{H_1'\times H_1} - < J_{H_2}u_2, C_2\tilde{\varphi} >_{H_2'\times H_2} - (u_1, \tilde{\eta}_1)_{H_1} - (u_2, \tilde{\eta}_2)_{H_2} - c =$$
$$= -(C_1^t J_{H_1}u_1, \tilde{\mathbf{j}})_{L^2(D)^n} - (C_2^t J_{H_2}u_2, \tilde{\varphi})_{L^2(D)} + (l_0, \tilde{f})_{L^2(D)} - (u_1, \tilde{\eta}_1)_{H_1} - (u_2, \tilde{\eta}_2)_{H_2} - c. \quad (26)$$

Далі враховуючи, що операторні рівняння (20)–(22) і (9)–(10) при $\mathbf{f} = \tilde{\mathbf{f}}$, в силу (23)–(24) і (11)–(12) еквівалентні відповідно наступним системам варіаційних рівнянь:

$$\int_D (\mathbf{A}^{-1}(x)\boldsymbol{\chi}_1(x), \mathbf{z}_1(x;u))_{\mathbb{R}^n} dx + \int_D z_2(x;u)\,\mathrm{div}\,\boldsymbol{\chi}_1(x)\,dx = \\ = -\int_D ((C_1^t J_{H_1}u_1)(x), \boldsymbol{\chi}_1(x))_{\mathbb{R}^n}\,dx \quad \forall \boldsymbol{\chi}_1 \in H(\mathrm{div}; D), \quad (27)$$

$$\int_D \chi_2(x)\,\mathrm{div}\,\mathbf{z}_1(x;u)\,dx = -\int_D (C_2^t J_{H_2}u_2)(x)\chi_2(x)\,dx \quad \forall \chi_2 \in L^2(D), \quad (28)$$

і

$$\int_D (\mathbf{A}^{-1}(x)\tilde{\mathbf{j}}(x), \boldsymbol{\psi}_1(x))_{\mathbb{R}^n} dx + \int_D \tilde{\varphi}(x)\,\mathrm{div}\,\boldsymbol{\psi}_1(x)\,dx = 0 \quad \forall \boldsymbol{\psi}_1 \in H(\mathrm{div}; D), \quad (29)$$

$$\int_D \psi_2(x)\,\mathrm{div}\,\tilde{\mathbf{j}}(x)\,dx = \int_D \tilde{f}(x)\psi_2(x)\,dx \quad \forall \psi_2 \in L^2(D), \quad (30)$$

перетворимо третій і четвертий доданок в (26). Для цього покладемо в (27) і (28) $\boldsymbol{\chi}_1 = \tilde{\mathbf{j}}$ і $\chi_2 = \tilde{\varphi}$. Тоді отримаємо

$$\int_D (\mathbf{A}^{-1}(x)\tilde{\mathbf{j}}(x), \mathbf{z}_1(x;u))_{\mathbb{R}^n} dx + \int_D z_2(x;u)\,\mathrm{div}\,\tilde{\mathbf{j}}(x)\,dx = \\ = -\int_D ((C_1^t J_{H_1}u_1)(x), \tilde{\mathbf{j}}(x))_{\mathbb{R}^n} dx, \quad (31)$$

$$\int_D \tilde{\varphi}(x)\mathrm{div}\,\mathbf{z}_1(x;u)\,dx = -\int_D (C_2^t J_{H_2}u_2)(x)\tilde{\varphi}(x)\,dx. \quad (32)$$

З іншого боку, підставляючи в (29) і (30) $\boldsymbol{\psi}_1 = \mathbf{z}_1(\cdot; u)$ і $\psi_2 = z_2(\cdot; u)$, знаходимо

$$\int\limits_D (\mathbf{A}^{-1}(x)\tilde{\mathbf{j}}(x), \mathbf{z}_1(x; u))_{\mathbb{R}^n} dx + \int\limits_D \tilde{\varphi}(x) \operatorname{div} \mathbf{z}_1(x; u_1, u_2)\, dx = 0, \tag{33}$$

$$\int\limits_D z_2(x; u) \operatorname{div} \tilde{\mathbf{j}}(x)\, dx = \int\limits_D \tilde{f}(x) z_2(x; u)\, dx. \tag{34}$$

Із (31)–(34) отримуємо

$$-(C_1^t J_{H_1} u_1, \tilde{\mathbf{j}})_{L^2(D)^n} - (C_2^t J_{H_2} u_2, \tilde{\varphi})_{L^2(D)} = \int\limits_D (\mathbf{A}^{-1}(x)\tilde{\mathbf{j}}(x), \mathbf{z}_1(x; u))_{\mathbb{R}^n} dx+$$

$$+\int\limits_D z_2(x; u) \operatorname{div} \tilde{\mathbf{j}}(x)\, dx + \int\limits_D \tilde{\varphi}(x) \operatorname{div} \mathbf{z}_1(x; u)\, dx = \int\limits_D (\mathbf{A}^{-1}(x)\tilde{\mathbf{j}}(x), \mathbf{z}_1(x; u))_{\mathbb{R}^n} dx+$$

$$+\int\limits_D \tilde{\varphi}(x) \operatorname{div} \mathbf{z}_1(x; u)\, dx + \int\limits_D z_2(x; u) \operatorname{div} \tilde{\mathbf{j}}(x)\, dx = (\tilde{f}, z_2(\cdot; u)_{L^2(D)} = (\tilde{f}, z_2(\cdot; u)_{L^2(D)}.$$

Звідси та з (26) випливає, що

$$l(f) - \widehat{l(f)} = (\tilde{f}, l_0 + z_2(\cdot; u)_{L^2(D)} - (u_1, \tilde{\eta}_1)_{H_1} - (u_2, \tilde{\eta}_2)_{H_2} - c =$$

$$= (\tilde{f} - f_0, l_0 + z_2(\cdot; u))_{L^2(D)} + (f_0, l_0 + z_2(\cdot; u))_{L^2(D)} - (u_1, \tilde{\eta}_1)_{H_1} - (u_2, \tilde{\eta}_2)_{H_2} - c.$$

В силу (5) знаходимо

$$\mathbb{E}\left|l(f) - \widehat{l(f)}\right|^2 = \left|(\tilde{f}_2 - f_0, l_0 + z_2(\cdot; u))_{L^2(D)} + (f_0, l_0 + z_2(\cdot; u))_{L^2(D)} - c\right|^2 +$$

$$+\mathbb{E}[(u_1, \tilde{\eta}_1)_{H_1} + (u_2, \tilde{\eta}_2)_{H_2}]^2.$$

З останньої рівності отримуємо

$$\inf_{c\in\mathbb{R}} \sup_{\tilde{f}\in G_0, (\tilde{\eta}_1, \tilde{\eta})\in G_1} \mathbb{E}|l(f) - \widehat{l(f)}|^2 =$$

$$= \inf_{c\in\mathbb{R}} \sup_{\tilde{f}\in G_0} \left[(\tilde{f} - f_0, l_0 + z_2(\cdot; u))_{L^2(D)} + (f_0, l_0 + z_2(\cdot; u))_{L^2(D)} - c\right]^2 +$$

$$+ \sup_{(\tilde{\eta}_1, \tilde{\eta}_2)\in G_1} \mathbb{E}[(\tilde{\eta}_1, u_1)_{H_1} + (\tilde{\eta}_2, u_2)_{H_2}]^2$$

$$= \sup_{\tilde{f}\in G_0} \left[(\tilde{f}_2 - f_2^{(0)}, l_0 + z_2(\cdot; u))_{L^2(D)}\right]^2 + \sup_{(\tilde{\eta}_1, \tilde{\eta}_2)\in G_1} \mathbb{E}[(\tilde{\eta}_1, u_1)_{H_1} + (\tilde{\eta}_2, u_2)_{H_2}]^2, \tag{35}$$

де інфімум по c досягається при $c = (f_0, l_0 + z_2(\cdot; u))_{L^2(D)}$. Далі, застосовуючи нерівність Коші-Буняківського, з (13) отримаємо

$$\left|(\tilde{f} - f_0, l_0 + z_2(\cdot; u))_{L^2(D)}\right|^2 \leq (Q^{-1}(l_0 + z_2(\cdot; u)), l_0 + z_2(\cdot; u))_{L^2(D)} (Q(\tilde{f} - f_0), \tilde{f} - f_0)_{L^2(D)} \leq$$

$$\leq (Q^{-1}(l_0 + z_2(\cdot; u)), l_0 + z_2(\cdot; u))_{L^2(D)},$$

причому рівність досягається при

$$\tilde{f} = f_0 + \frac{Q^{-1}(l_0 + z_2(\cdot; u))}{(Q^{-1}(l_0 + z_2(\cdot; u)), l_0 + z(\cdot; u))^{1/2}_{L^2(D)}}.$$

Звідси, отримуємо

$$\sup_{\tilde{f} \in G_0} \left[(\tilde{f}_2 - f_2^{(0)}, l_0 + z_2(\cdot; u))_{L^2(D)} \right]^2 = (Q^{-1}(l_0 + z_2(\cdot; u)), l_0 + z_2(\cdot; u))_{L^2(D)}.$$

Аналогічно, в силу (17), знаходимо

$$\sup_{(\tilde{\eta}_1, \tilde{\eta}_2) \in G_1} \mathbb{E}[(\tilde{\eta}_1, u_1)_{H_1} + (\tilde{\eta}_2, u_2)_{H_2}]^2 = (\tilde{Q}_1^{-1} u_1, u_1)_{H_1} + (\tilde{Q}_2^{-1} u_2, u_2)_{H_2}.$$

Із останніх двох рівностей та з (30) знаходимо

$$\inf_{c \in \mathbb{R}} \sup_{\tilde{f} \in G_0, (\tilde{\eta}_1, \tilde{\eta}) \in G_1} \mathbb{E}|l(\tilde{f}) - \widehat{l(\tilde{f})}|^2 = I(u),$$

при $c = (l_0 + z_2(\cdot; u), f_0)_{L^2(D)}$, а $I(u)$ визначається формулою (25). □

В результаті розв'язування задачі оптимального керування (23) – (25) приходимо до наступного результату.

Теорема 2. Існує єдина мінімаксна оцінка значення $l(f)$, яка має вигляд

$$\widehat{\widehat{l(f)}} = (y_1(\mathbf{j}; \eta_1), \hat{u}_1)_{H_1} + (y_2(\varphi; \eta_2), \hat{u}_2)_{H_2} + \hat{c}$$

де

$$\hat{c} = \int_D (l_0(x) + \hat{z}_2(x)) f_0(x)\, dx, \quad \hat{u}_1 = \tilde{Q}_1 C_1 \mathbf{p}_1, \quad \hat{u}_2 = \tilde{Q}_2 C_2 p_2, \tag{36}$$

а функції $\hat{\mathbf{z}}_1, \mathbf{p}_1 \in H(\mathrm{div}, D)$ і $\hat{z}_2, p_2 \in L^2(D)$ знаходяться з розв'язку наступної системи варіаційних рівнянь:

$$\int_D (\mathbf{A}^{-1}(x)\, \hat{\mathbf{z}}_1(x), \mathbf{q}_1(x))_{\mathbb{R}^n} dx + \int_D \hat{z}_2(x) \mathrm{div}\, \mathbf{q}_1(x)\, dx =$$

$$= -\int_D (C_1^t J_{H_1} \tilde{Q}_1 C_1 \mathbf{p}_1(x), \mathbf{q}_1(x))_{\mathbb{R}^n}\, dx \quad \forall \mathbf{q}_1 \in H(\mathrm{div}, D), \tag{37}$$

$$\int_D v_1(x) \mathrm{div}\, \hat{\mathbf{z}}_1(x)\, dx = -\int_D (C_2^t J_{H_2} \tilde{Q}_2 C_2 p_2(x) v_1(x)\, dx \quad \forall v_1 \in L^2(D), \tag{38}$$

$$\int_D (\mathbf{A}^{-1}(x)\, \mathbf{p}_1(x), \mathbf{q}_2(x))_{\mathbb{R}^n} dx + \int_D p_2(x) \mathrm{div}\, \mathbf{q}_2(x)\, dx = 0 \quad \forall \mathbf{q}_2 \in H(\mathrm{div}, D), \tag{39}$$

$$\int_D v_2(x) \mathrm{div}\, \mathbf{p}_1(x)\, dx = \int_D v_2(x) Q^{-1}(l_0 + \hat{z}_2(\cdot))(x)\, dx \quad \forall v_2 \in L^2(D). \tag{40}$$

Задача (37) – (40) однозначно розв'язна. Похибка мінімаксного оцінювання ϱ визначається формулою

$$\varrho = l(Q^{-1}(l_0 + \hat{z}_2)). \tag{41}$$

Доведення. Покажемо, що розв'язок задачі оптимального керування (23)–(25) зводиться до розв'язку системи рівнянь (37)-(40). Для цього спочатку зауважимо, що із вигляду функціоналу $I(u)$ існує єдиний елемент $\hat{u} :=\in H$, на якому досягається мінімум цього функціоналу, тобто $I(\hat{u}) = \inf_{u \in H_0} I(u)$. Тому, для будь-яких $\tau \in \mathbb{R}$ і $w = (w_1, w_2) \in H_0$ виконується співвідношення

$$0 = \frac{d}{dt} I(\hat{u} + \tau w)\Big|_{\tau=0} =$$

$$= (Q^{-1}(l_0 + z_2(\cdot; \hat{u})), \tilde{z}_2(\cdot; w))_{L^2(D)} + (\tilde{Q}_1^{-1}\hat{u}_1, w_1)_{H_1} + (\tilde{Q}_2^{-1}\hat{u}_2, w_2)_{H_2}, \tag{42}$$

де через $(\tilde{\mathbf{z}}_1(\cdot; w), \tilde{z}_2(\cdot; w))$ позначино єдиний розв'язок системи рівнянь (23), (24) при $u = w$.

Далі, ввівши функції $\mathbf{p}_1 \in H(\mathrm{div}, D)$ і $p_2 \in L^2(D)$ як єдиний розв'язок задачі

$$\int\limits_D (\mathbf{A}^{-1}(x)\,\mathbf{p}_1(x), \mathbf{q}_2(x))_{\mathbb{R}^n} dx + \int\limits_D p_2(x) \mathrm{div}\, \mathbf{q}_2(x)\, dx = 0 \quad \forall \mathbf{q}_2 \in H(\mathrm{div}, D), \tag{43}$$

$$\int\limits_D v_2(x) \mathrm{div}\, \mathbf{p}_1(x)\, dx = \int\limits_D v_2(x) Q^{-1}(l_0 + z_2(\cdot; \hat{u}))(x)\, dx \quad \forall v_2 \in L^2(D). \tag{44}$$

і провівши міркування подібні до тих, які використовувались при доведенні теореми 1, прийдемо до наступної рівності

$$(Q^{-1}(l_0 + z_2(\cdot; \hat{u})), \tilde{z}_2(\cdot; w))_{L^2(D)} = -(w_1, C_1\mathbf{p}_1)_{H_1} - (w_2, C_2 p_2)_{H_2}.$$

звідки, внаслідок (42), знайдемо

$$(w_1, C_1\mathbf{p}_1)_{H_1} + (w_2, C_2 p_2)_{H_2} = (\tilde{Q}_1^{-1}\hat{u}_1, w_1)_{H_1} + (\tilde{Q}_2^{-1}\hat{u}_2, w_2)_{H_2},$$

Звідси випливає, що $\hat{u}_1 = \tilde{Q}_1 C_1 \mathbf{p}_1, \quad \hat{u}_2 = \tilde{Q}_2 C_2 p_2$. Замінюючи в правих частинах рівностей (23), (24) u_1 і u_2 відповідно на знайдені вирази $\hat{u}_1$ і $\hat{u}_2$ і, вводячи позначення $\mathbf{z}_1(x; \hat{u}_1, \hat{u}_2) =: \hat{\mathbf{z}}_1(x)$, $z_2(x; \hat{u}_1, \hat{u}_2) =: \hat{z}_2(x)$, отримуємо, що функції $(\hat{\mathbf{z}}_1, \hat{z}_2)$ і $(\mathbf{p}_1, p_2)$ задовольняють системі рівнянь (37) – (40), однозначна розв'язність якої випливає з єдиності елементу $\hat{u} = (\hat{u}_1, \hat{u}_2)$.

Знайдемо далі похибку оцінювання. Підставляючи значення $\hat{u}_1 = \tilde{Q}_1 C_1 \mathbf{p}_1$ і $\hat{u}_2 = \tilde{Q}_2 C_2 p_2$ у вираз для $I(\hat{u})$, отримуємо

$$\varrho = I(\hat{u}) = (Q^{-1}(l_0 + z_2(\cdot; \hat{u}_1, \hat{u}_2)), l_0 + z_2(\cdot; \hat{u}_1, \hat{u}_2))_{L^2(D)} + (\tilde{Q}_1^{-1}\hat{u}_1, \hat{u}_1)_{H_1} + (\tilde{Q}_2^{-1}\hat{u}_2, \hat{u}_2)_{H_2} =$$

$$= (Q^{-1}(l_0 + \hat{z}_2), l_0 + \hat{z}_2)_{L^2(D)} + (C_1\mathbf{p}_1, \tilde{Q}_1 C_1 \mathbf{p}_1)_{H_1} + (C_2 p_2, \tilde{Q}_2 C_2 p_2)_{H_2}. \tag{45}$$

Підставляючи в (39) і (40) $\mathbf{q}_2 = \hat{\mathbf{z}}_1$ і $v_2 = \hat{z}_2$, знаходимо

$$\int\limits_D (\mathbf{A}^{-1}(x)\,\mathbf{p}_1(x), \hat{\mathbf{z}}_1(x))_{\mathbb{R}^n} dx + \int\limits_D p_2(x) \mathrm{div}\, \hat{\mathbf{z}}_1(x)\, dx = 0,$$

$$\int_D \hat{z}_2(x)\mathrm{div}\,\mathbf{p}_1(x)\,dx = \int_D \hat{z}_2(x)Q^{-1}(l_0+\hat{z}_2)(x)\,dx.$$

З останніх двох співвідношень і з системи варіаційних рівнянь (37) і (38), в яких покладено $\mathbf{q}_1 = \mathbf{p}_1$ і $v_1 = p_2$ матимемо

$$(Q^{-1}(l_0+\hat{z}_2), l_0+\hat{z}_2)_{L^2(D)} = (Q^{-1}l_0, l_0+\hat{z}_2)_{L^2(D)} + \int_D \hat{z}_2(x)\mathrm{div}\,\hat{\mathbf{p}}_1(x)\,dx+$$

$$+\int_D (\mathbf{A}^{-1}(x)\,\mathbf{p}_1(x), \hat{\mathbf{z}}_1(x))_{\mathbb{R}^n}dx + \int_D p_2(x)\mathrm{div}\,\hat{\mathbf{z}}_1(x)\,dx =$$

$$= (l_0, Q^{-1}(l_0+\hat{z}_2))_{L^2(D)} + \int_D (\mathbf{A}^{-1}(x)\,\mathbf{p}_1(x), \hat{\mathbf{z}}_1(x))_{\mathbb{R}^n}dx + \int_D z_2(x)\mathrm{div}\,\mathbf{p}_1(x)\,dx+$$

$$+\int_D p_2(x)\mathrm{div}\,\hat{\mathbf{z}}_1(x)\,dx = (l_0, Q^{-1}(l_0+\hat{z}_2))_{L^2(D)} - \int_D (C_1^t J_{H_1}\tilde{Q}_1 C_1\mathbf{p}_1(x), \mathbf{p}_1(x))_{\mathbb{R}^n}\,dx-$$

$$-\int_D C_2^t J_{H_2}\tilde{Q}_2 C_2 p_2(x)p_2(x)\,dx = l(Q^{-1}(l_0+\hat{z}_2)) - (C_1\mathbf{p}_1, \tilde{Q}_1 C_1\mathbf{p}_1)_{H_1} - (C_2p_2, \tilde{Q}_2C_2p_2)_{H_2}.$$

Звідси і з (45) випливає співвідношення (41) для виразу похибки оцінювання. □

Інше представлення для мінімаксної оцінки, яке не залежить від кокретного вигляду функционала $l(f)$, міститься в наступному твердженні.

Теорема 3. Мінімаксна оцінка виразу $l(f)$ має вигляд $\widehat{\widehat{l(f)}} = l(\hat{f})$, де $\hat{f}(x,\omega) = Q^{-1}\hat{p}_2(x,\omega) + f^{(0)}(x)$, а випадкове поле $\hat{p}_2 \in L^2(\Omega, L^2(D))$ знаходиться з розв'язку наступної системи варіаційних стохастичних рівнянь:

$$\int_D (\mathbf{A}^{-1}(x)\,\hat{\mathbf{p}}_1(x,\omega), \mathbf{q}_1(x))_{\mathbb{R}^n}dx + \int_D \hat{p}_2(x,\omega)\mathrm{div}\,\mathbf{q}_1(x)\,dx =$$

$$= \int_D (C_1^t J_{H_1}\tilde{Q}_1(y_1(\mathbf{j},\eta_1(\omega)) - C_1\hat{\mathbf{j}}(\cdot,\omega))(x), \mathbf{q}_1(x))_{\mathbb{R}^n}dx \;\forall \mathbf{q}_1 \in H(\mathrm{div}, D), \quad (46)$$

$$\int_D v_1(x)\mathrm{div}\,\hat{\mathbf{p}}_1(x,\omega)\,dx = \int_D (C_2^t J_{H_2}\tilde{Q}_2(y_2(\varphi,\eta_2(\omega))-$$

$$- C_2\hat{\varphi}(\cdot,\omega))(x)v_1(x)\,dx \quad \forall v_1 \in L^2(D), \quad (47)$$

$$\int_D (\mathbf{A}^{-1}(x)\,\hat{\mathbf{j}}(x,\omega), \mathbf{q}_2(x))_{\mathbb{R}^n}\,dx + \int_D \hat{\varphi}(x,\omega)\mathrm{div}\,\mathbf{q}_2(x)\,dx = 0 \quad \forall \mathbf{q}_2 \in H(\mathrm{div}, D), \quad (48)$$

$$\int_D v_2(x)\mathrm{div}\,\hat{\mathbf{j}}(x,\omega)\,dx = \int_D v_2(x)(Q^{-1}\hat{p}_2(x,\omega) + f_0(x))\,dx \quad \forall v_2 \in V_2, \quad (49)$$

в яких рівності (46) – (49) виконуються з ймовірністю 1. Задача (46) – (49) має єдиний розв'язок.

Доведення цієї теореми є аналогічним доведенню теореми 2.

На завершення зауважимо, що користуючись запропонованими в [12] змішаними методами скінченних елементів, для знаходження розвязків задач (37) – (40) і (46) – (49) можливо розробити наближені методи їх розвязання.

Заключення

В роботі отримані наступні результати: *встановлена еквівалентність задачі мінімаксного оцінювання деякій задачі оптимізації; доведені нові твердження про загальний вигляд мінімаксних середньоквадратичних оцінок функціоналів від невідомих правих частин рівнянь, що входять у постановку розглядуваних в роботі крайових задач, і отримані представлення для похибок оцінювання.*

Список літератури

1. *Наконечний О.Г.* Оптимальне керування та оцінювання в рівняннях із частинними похідними // Київський університет, Київ, 2004 г., 103 с.
2. *Brezzi F., Fortin M.* Mixed and hybrid finite element methods // Springer-Verlag, New York, 1991, 350 p.
3. *Подлипенко Ю.К., Грищук Н.В.* Оптимальное прогнозировании решений параболических уравнений по наблюдениям, распределенным на системе поверхностей // Доповіді НАН України. – 2003. – №9. с. 107–112.
4. *Подлипенко Ю.К., Грищук Н.В.* Оцінювання параметрів вироджених еліптичних крайових задач Неймана в умовах невизначеності //Вісник Київського університету. Серія: фізико-математичні науки. – 2004. – №1. с. 262–269.
5. *Hiptmair R., Schwab C.* / Numerical treatment of partial differential equations. Lecture notes for course held by R. Hiptmair in WS03/04. pp. 1–219
6. *Наконечний О.Г.* Оптимальне керування та оцінювання в рівняннях із частинними похідними. – Київський університет, Київ 2004. – 103 с.
7. *Наконечный А.Г.* Минимаксное оценивание функционалов от решений вариационных уравнений в гильбертовых пространствах. – Киев: КГУ, 1985. – 82 с.
8. *Красовский Н.Н.* Теория управления движением. – М.: Наука, 1968. – 476 с.
9. *Подлипенко Ю.К.* Задачи минимаксного оценивания для нетеровых уравнений в гильбертовом пространстве // Доповіді НАН України.Серія: Математика. – 2005. – №12. с. 36-44.
10. *Подлипенко Ю.К., Грищук Н.В.* Минимаксное оценивание решений вырожденных краевых задач Неймана для эллиптических уравнений по наблюдениям, распределенным на системе поверхностей // Системні дослідження і інформаційні технології. – 2004. – №2 с. 104–128.
11. *Подлипенко Ю.К., Грищук Н.В.* Оцінювання параметрів вироджених еліптичних крайових задач Неймана в умовах невизначеності // Вісник Київського університету. Серія: фізико-математичні науки. – 2004. – №1. с. 262–269.
12. *F. Brezzi, M. Fortin* Mixed and hybrid finite element methods. – Springer-Verlag, New York, 1991. – 350 p.

Статья поступила в редакцию 16.09.2008

УДК 519.8

МИНИМАКСНОЕ ОЦЕНИВАНИЕ НЕИЗВЕСТНЫХ ДАННЫХ КРАЕВОЙ ЗАДАЧИ ДЛЯ БИГАРМОНИЧЕСКОГО УРАВНЕНИЯ С ГРАНИЧНЫМИ УСЛОВИЯМИ ТИПА НЕЙМАНА

© Перцов А.С.

Черновицкий национальный университет им. Ю.Федьковича
факультет прикладной математики
ул. Коцюбинского, 2, г. Черновцы, 58012, Украина
e-mail: *pertsov@ukr.net*

Abstract. We find minimax estimates of functionals from unknown deterministic data of the boundary value problem for biharmonic equation with Neumann type boundary conditions.

Введение

Задачам минимаксного оценивания состояний систем, описываемых обыкновенными дифференциальными уравнениями и уравнениями в частных производных при условии их однозначной разрешимости посвящено значительное число работ (см., например, [1]).

Однако, в ситуации, когда решения краевых задач не определены однозначно и существуют лишь, если данные этих краевых задач удовлетворяют некоторым условиям совместности, вопросы их минимаксного оценивания разработаны недостаточно полно (в этом направлении см., например, [3]). Исследуемая ниже задача минимаксного оценивания относится к описанному кругу проблем.

В данной статье по зашумленным наблюдениям решений и при специальных ограничениях на неизвестные правые части уравнений и граничных условий типа Неймана, входящих в постановку краевых задач для бигармонического уравнения, а также на шумы в наблюдениях, найдены минимаксные оценки функционалов от этих правых частей.

Нахождение минимаксных оценок сведено к решению некоторых систем интегро-дифференциальных уравнений и доказана их однозначная разрешимость.

1. Вспомогательные факты

Обозначим через H – гильбертово пространство над $\mathbb{R}$ со скалярным произведением $(\cdot,\cdot)_H$ и нормой $\|\cdot\|_H$. Через $J_H \in \mathscr{L}(H,H')$ будем обозначать оператор, называемый изометричным изоморфизмом, действующий из H на его сопряженное пространство H', и определяемый равенством $(u,v)_H = < J_H u, v >_{H'\times H}$ $\forall u,v \in H$, где $< f,x >_{H'\times H} := f(x)$ для $x \in H$, $f \in H'$. Этот оператор существует в силу теоремы Рисса.

Обозначим через $L^2(\Omega,H)$ пространство Бохнера, состоящее из случайных элементов $\xi=\xi(\omega)$, определенных на некотором вероятностном пространстве $(\Omega,\mathscr{B},P)$

со значениями в H таких, что

$$\|\xi\|^2_{L^2(\Omega,H)} = \int\limits_{\Omega} \|\xi(\omega)\|^2_H dP(\omega) < \infty. \tag{1}$$

В этом случае существует интеграл Бохнера $\mathbb{E}\xi := \int_\Omega \xi(\omega)\, dP(\omega) \in H$, называемый математическим ожиданием или средним случайного элемента $\xi(\omega)$, удовлетворяющий условию

$$(h, \mathbb{E}\xi)_H = \int\limits_{\Omega} (h, \xi(\omega))_H \, dP(\omega) \quad \forall h \in H. \tag{2}$$

Применение к случайной величине ξ этого определения приводит к традиционному определению ее математического ожидания, поскольку интеграл Бохнера (1) переходит в обычный интеграл Лебега по вероятностной мере $dP(\omega)$. В $L^2(\Omega, H)$ можно ввести скалярное произведение:

$$(\xi, \eta)_{L^2(\Omega,H)} := \int\limits_{\Omega} (\xi(\omega), \eta(\omega))_H \, dP(\omega) \quad \forall \xi, \eta \in L^2(\Omega, H). \tag{3}$$

Пространство $L^2(\Omega, H)$, оснащенное нормой (1) и скалярным произведением (3), является гильбертовым.

Введем также следующие обозначения: $x = (x_1, \ldots, x_n)$ – пространственная переменная, принадлежащая ограниченной открытой области $D \subset \mathbb{R}^n$ с липшицевой границей Γ; $dx = dx_1 \cdots dx_n$ – мера Лебега в $\mathbb{R}^n$; $L^2(D)$ – пространство функций, суммируемых с квадратом в области D; для целого числа m обозначим через $H^m(D)$ – стандартные пространства Соболева с естественными нормами.

Пусть состояние $\varphi(x)$ системы определяется как решение краевой задачи

$$\varphi \in H^2(D), \tag{4}$$

$$\Delta^2 \varphi(x) = f(x) \quad \text{в} \quad D, \tag{5}$$

$$N\varphi = h_1, \quad M\varphi = h_2 \qquad \text{на} \quad \Gamma, \tag{6}$$

где

$$M\varphi = \sigma \Delta \varphi + (1-\sigma) \frac{\partial^2 \varphi}{\partial \nu^2}, \tag{7}$$

$$N\varphi = -\frac{\partial}{\partial \nu}(\Delta) + (1-\sigma) \frac{\partial}{\partial s} \left[\frac{\partial^2 \varphi}{\partial x_1^2} \nu_1 \nu_2 - \frac{\partial^2 \varphi}{\partial x_1 \partial x_2} (\nu_1^2 - \nu_2^2) - \frac{\partial^2 \varphi}{\partial x_2^2} \nu_1 \nu_2 \right], \tag{8}$$

$f \in L^2(D)$, $h_1, h_2 \in L^2(\Gamma)$, $0 \le \sigma < 1$, ν – единичная нормаль к Γ, внешняя по отношению к области D, ν_i – i-я координата единичной нормали ν, $\frac{\partial}{\partial s} = -\frac{\partial}{\partial x_1}\nu_2 + \frac{\partial}{\partial x_2}\nu_1$ – производная по направлению касательной к кривой Γ. При этом под обобщенным решением этой задачи понимается функция $\varphi \in H^2(D)$, для которой справедливо интегральное тождество ([2], стр. 435),

$$\int\limits_D \left[\left(\frac{\partial^2 v}{\partial x_1^2} + \sigma \frac{\partial^2 v}{\partial x_2^2} \right) \frac{\partial^2 \varphi}{\partial x_1^2} + 2(1-\sigma) \frac{\partial^2 v}{\partial x_1 \partial x_2} \frac{\partial^2 \varphi}{\partial x_1 \partial x_2} + \left(\frac{\partial^2 v}{\partial x_2^2} + \sigma \frac{\partial^2 v}{\partial x_1^2} \right) \frac{\partial^2 \varphi}{\partial x_2^2} \right] dx$$

$$= \int\limits_D v(x)f(x)\,dx + \int\limits_\Gamma v h_1\,d\Gamma + \int\limits_\Gamma \frac{\partial v}{\partial \nu} h_2\,d\Gamma \quad \forall v \in H^2(D). \tag{9}$$

Как известно (см., например, [2]), для существования решения задачи (4) – (6) необходимо и достаточно, чтобы удовлетворялись условия совместности

$$\int\limits_D f(x)\,dx + \int\limits_\Gamma h_1\,d\Gamma = 0, \quad \int\limits_D x_1 f(x)\,dx + \int\limits_\Gamma x_1 h_1\,d\Gamma + \int\limits_\Gamma \frac{\partial x_1}{\partial \nu} h_2\,d\Gamma = 0, \tag{10}$$

$$\int\limits_D x_2 f(x)\,dx + \int\limits_\Gamma x_2 h_1\,d\Gamma + \int\limits_\Gamma \frac{\partial x_2}{\partial \nu} h_2\,d\Gamma = 0. \tag{11}$$

Если условия (10), (11) выполняются, то существует бесконечное множество решений данной задачи, причем любые два решения отличаются друг от друга на полином вида $p(x_1, x_2) = a + bx_1 + cx_2$ почти всюду в D. Обозначим множество этих полиномов через P_1.

Постановка задачи минимаксного оценивания. Задача состоит в том, чтобы по наблюденим вида

$$y = y(\varphi; \eta) = C\varphi + \eta \tag{12}$$

найти оптимальную, в некотором смысле, оценку значения функционала

$$l(F) = \int\limits_D l_0(x) f(x)\,dx + \int\limits_\Gamma l_1 h_1\,d\Gamma + \int\limits_\Gamma l_2 h_2\,d\Gamma \tag{13}$$

в классе линейных оценок

$$\widehat{l(F)} = (y(\varphi; \eta), u)_{H_0} + c, \tag{14}$$

где $\varphi(x)$ – решение краевой задачи (4) – (6), $F := (f, h_1, h_2) \in L^2(D) \times L^2(\Gamma) \times L^2(\Gamma)$, элемент u принадлежит гильбертову пространству H_0, $c \in \mathbb{R}$, $l_0 \in L^2(D)$, $l_1 \in L^2(\Gamma)$, $l_2 \in L^2(\Gamma)$ – заданные функции, в предположении, что правые части $f(x)$, h_1, h_2 уравнений (5), (6) и погрешности $\eta = \eta(\omega)$ в наблюдениях (12), являющиеся случайными элементами, определенными на некотором вероятностном пространстве $(\Omega, \mathscr{B}, P)$ со значениями в H_0, неизвестны, а известно лишь, что элемент $F := (f, h_1 h_2) \in G_0$ и $\eta \in G_1$. Здесь $C \in \mathscr{L}(L^2(D), H_0)$ – линейный непрерывный оператор, такой что его ограничение на подпространство P_1 инъективно; через G_0 обозначено множество функций $\tilde{F} := (\tilde{f}, \tilde{h}_1, \tilde{h}_2) \in L^2(D) \times L^2(\Gamma) \times L^2(\Gamma)$, удовлетворяющих условиям

$$\int\limits_D Q(\tilde{f} - f_0)(x)(\tilde{f}(x) - f_0(x))\,dx + \int\limits_\Gamma Q_1(\tilde{h}_1 - h_1^{(0)})(\tilde{h}_1 - h_1^{(0)})\,d\Gamma$$
$$+ \int\limits_\Gamma Q_2(\tilde{h}_2 - h_2^{(0)})(\tilde{h}_2 - h_2^{(0)})\,d\Gamma \le 1, \tag{15}$$

и

$$\int\limits_D \tilde f(x)\,dx + \int\limits_\Gamma \tilde h_1\,d\Gamma = 0, \quad \int\limits_D x_1 \tilde f(x)\,dx + \int\limits_\Gamma x_1\tilde h_1\,d\Gamma + \int\limits_\Gamma \frac{\partial x_1}{\partial \nu}\tilde h_2\,d\Gamma = 0, \tag{16}$$

$$\int\limits_D x_2 \tilde f(x)\,dx + \int\limits_\Gamma x_2\tilde h_1\,d\Gamma + \int\limits_\Gamma \frac{\partial x_2}{\partial \nu}\tilde h_2\,d\Gamma = 0, \tag{17}$$

а через G_1 обозначено множество случайных элементов $\tilde\eta = \tilde\eta(\omega) \in L^2(\Omega, H_0)$, с нулевыми средними, удовлетворяющими неравенству

$$\mathbf{M}(Q_0\tilde\eta, \tilde\eta)_{H_0} \le 1, \tag{18}$$

где Q, Q_1, Q_2 і Q_0 – ограниченные самосопряженные положительно-определенные операторы в $L^2(D)$, $L^2(\Gamma)$ и H_0 соответсвенно, для которых существуют ограниченные обратные операторы Q^{-1}, Q_1^{-1}, Q_2^{-1} и Q_0^{-1}, $F_0 := (f_0, h_1^{(0)}, h_2^{(0)}) \in L^2(D) \times L^2(\Gamma) \times L^2(\Gamma)$ – заданные функции, удовлетворяющие условиям (16)–(17).

Определение 1. Оценку вида $\widehat{\widehat{l(F)}} = (y(\varphi;\eta), \hat u)_{H_0} + \hat c$ будем называть минимаксной оценкой $l(F)$, если элемент $\hat u$ и число $\hat c$ определяются из условия

$$\sigma(u,c) := \sup_{\tilde F \in G_0, \tilde\eta\in G_1} \mathbf{M}|l(\tilde F) - \widehat{l(\tilde F)}|^2 \to \inf_{u\in H_0, c\in\mathbb{C}} := \sigma^2,$$

где $\tilde\varphi$ – любое решение краевой задачи (4)– (6) при $f(x) = \tilde f(x)$, $h_1 = \tilde h_1$, $h_2 = \tilde h_2$, $\widehat{l(\tilde F)} = (y(\tilde\varphi;\tilde\eta), u)_{H_0} + c$. Величину σ будем называть погрешностью минимаксного оценивания выражения (13).

2. Основные результаты

Теорема 1. *Задача нахождения минимаксной оценки $\widehat{l(F)}$ значения функционала $l(F)$ эквивалентна задаче оптимального управления системой, описываемой краевой задачей*

$$z(\cdot;u) \in H^2(D), \tag{19}$$

$$\Delta^2 z(x;u) = -(C^* J_{H_0} u)(x) \quad \text{в} \quad D, \tag{20}$$

$$Nz(\cdot;u) = 0, \quad Mz(\cdot;u) = 0, \quad \text{на} \quad \Gamma, \tag{21}$$

$$\int\limits_D Q^{-1}(l_0(x) + z(x;u))\,dx + \int\limits_\Gamma Q_1^{-1}(l_1 + z(\cdot;u))\,d\Gamma = 0, \tag{22}$$

$$\int\limits_D x_1 Q^{-1}(l_0(x)+z(x;u))\,dx + \int\limits_\Gamma x_1 Q_1^{-1}\left(l_1 + z(\cdot;u)\right)\,d\Gamma + \int\limits_\Gamma \frac{\partial x_1}{\partial\nu} Q_2^{-1}\left(l_2 + \frac{\partial z(\cdot;u)}{\partial \nu}\right)\,d\Gamma = 0, \tag{23}$$

$$\int\limits_D x_2 Q^{-1}(l_0(x)+z(x;u))\,dx + \int\limits_\Gamma x_2 Q_1^{-1}\left(l_1 + z(\cdot;u)\right)\,d\Gamma + \int\limits_\Gamma \frac{\partial x_2}{\partial\nu} Q_2^{-1}\left(l_2 + \frac{\partial z(\cdot;u)}{\partial \nu}\right)\,d\Gamma = 0. \tag{24}$$

с функцией стоимости вида

$$I(u) = \int_D Q^{-1}(z(x;u) + l_0(x))(z(x;u) + l_0(x))dx + \int_\Gamma Q_1^{-1}(z(\cdot;u) + l_1)(z(\cdot;u) + l_1)\, d\Gamma$$

$$+ \int_\Gamma Q_2^{-1}\left(\frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right)\left(\frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right) d\Gamma + (Q_0^{-1}u, u)_{H_0} \to \inf_{u\in V}, \quad (25)$$

где[1]

$$V = \{u \in H_0 : \int_D (C^* J_{H_0}u)(x) p_0(x)\, dx = 0 \quad \forall p_0 \in P_1\},$$

$C^* : H_0' \to L_2(a,b)$ *– оператор, сопряженный к* C*, который определяется соотношением* $< Cv, w >_{H_0\times H_0'} = \int_D v(x)C^*w(x)\, dx$ *для всех* $v \in L^2(D)$, $w \in H_0'$.

Доказательство. Сначала заметим, что задача (19)–(24), в силу условий совместности (10)–(11), в которых положено $f(x) = -(C^* J_{H_0}u)(x)$, $h_1 = 0$, $h_2 = 0$, однозначно разрешима при $u \in V$.

Обозначим через $\tilde\varphi_\perp$ единственное решение задачи (4)-(6) при $f(x) = \tilde f(x)$, $h_1 = \tilde h_1$, $h_2 = \tilde h_2$, ортогональное ко всем полиномам из множества P_1.

Тогда, поскольку любое решение $\tilde\varphi$ этой задачи можно представить в виде $\tilde\varphi = \tilde\varphi_\perp + \varphi_0$, где $\tilde\varphi_0 \in P_1$, имеем

$$\sup_{\tilde F\in G_0, \tilde\eta\in G_1} \mathbf{M}|l(\tilde F) - \widehat{l(\tilde F)}|^2 = \sup_{\tilde F\in G_0, \tilde\eta\in G_1} \sup_{\varphi_0\in P_1} \mathbf{M}|l(\tilde\varphi_\perp + \varphi_0) - l(\widehat{\tilde\varphi_\perp + \varphi_0})|^2.$$

Учитывая (12)-(14), для любого $u \in H_0$ получим

$$\widehat{l(\tilde F)} = (y(\tilde\varphi_\perp + \varphi_0; \tilde\eta), u)_{H_0} + c = (C(\tilde\varphi_\perp + \varphi_0), u)_{H_0} + (\tilde\eta, u)_{H_0} + c$$

$$=< C(\tilde\varphi_\perp + \varphi_0), J_{H_0}u >_{H_0\times H_0'} + (\tilde\eta, u)_{H_0} + c$$

$$= \int_D (\tilde\varphi_\perp(x) + \varphi_0(x))(C^* J_{H_0}u)(x)\, dx + (\tilde\eta, u)_{H_0} + c$$

$$= \int_D \tilde\varphi_\perp(x)(C^* J_{H_0}u)(x)\, dx + \int_D \varphi_0(x)(C^* J_{H_0}u)(x)\, dx + (\tilde\eta, u)_{H_0} + c,$$

откуда

$$l(\tilde F) - \widehat{l(\tilde F)} = \int_D l_0(x)\tilde f(x)\, dx + \int_\Gamma l_1\tilde h_1\, d\Gamma + \int_\Gamma l_2\tilde h_2\, d\Gamma - l(\widehat{\tilde\varphi_\perp + \varphi_0})$$

$$= \int_D l_0(x)\tilde f(x)\, dx + \int_\Gamma l_1\tilde h_1\, d\Gamma + \int_\Gamma l_2\tilde h_2\, d\Gamma$$

$$- \int_D \tilde\varphi_\perp(x)(C^* J_{H_0}u)(x)\, dx - \int_D \varphi_0(x)(C^* J_{H_0}u)(x)\, dx - (\tilde\eta, u)_{H_0} - c.$$

[1]Нетрудно видеть,что V – непустое, замкнутое, выпуклое множество.

Отсюда, принимая во внимание соотношение $\mathbf{D}\xi = \mathbf{M}|\xi - \mathbf{M}\xi|^2 = \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2$ между дисперсией $\mathbf{D}\xi$ случайной величины ξ и ее математическим ожиданием $\mathbf{M}\xi$, находим

$$\mathbf{M}\left|l(\tilde F) - \widehat{l(\tilde F)}\right|^2 = \sup_{\varphi_0 \in P_1}\left|\int\limits_D l_0(x)\tilde f(x)\,dx + \int\limits_\Gamma l_1\tilde h_1\,d\Gamma + \int\limits_\Gamma l_2\tilde h_2\,d\Gamma\right.$$

$$\left. - \int\limits_D \tilde\varphi_\perp(x)(C^* J_{H_0}u)(x)\,dx - \int\limits_D \varphi_0(x)(C^* J_{H_0}u)(x)\,dx - c\right|^2 + \mathbf{M}|(\tilde\eta, u)_{H_0}|^2. \qquad (26)$$

Поскольку функция $\varphi_0(x)$ под знаком последнего интеграла в правой части этого равенства пробегает все пространство P_1, то величина $\mathbf{M}\left|l(\tilde F) - \widehat{l(\tilde F)}\right|^2$ будет ограниченной тогда и только тогда, когда $u \in V$. Учитывая теперь, что функция $z(x;u)$, которая определена лишь тогда, когда $u \in V$, удовлетворяет уравнениям (19)–(21), имеем

$$\int\limits_D \left[\left(\frac{\partial^2 w(x)}{\partial x_1^2} + \sigma\frac{\partial^2 w(x)}{\partial x_2^2}\right)\frac{\partial^2 z(x;u)}{\partial x_1^2} + 2(1-\sigma)\frac{\partial^2 w(x)}{\partial x_1\partial x_2}\frac{\partial^2 z(x;u)}{\partial x_1\partial x_2}\right.$$

$$\left. + \left(\frac{\partial^2 w(x)}{\partial x_2^2} + \sigma\frac{\partial^2 w(x)}{\partial x_1^2}\right)\frac{\partial^2 z(x;u)}{\partial x_2^2}\right]dx = -\int\limits_D w(x)C^* J_{H_0}u(x)\,dx \quad \forall w \in H^2(D). \quad (27)$$

Положив в последнем тождестве $w = \tilde\varphi_\perp$, получим равенство

$$\int\limits_D \left[\left(\frac{\partial^2 \tilde\varphi_\perp(x)}{\partial x_1^2} + \sigma\frac{\partial^2 \tilde\varphi_\perp(x)}{\partial x_2^2}\right)\frac{\partial^2 z(x;u)}{\partial x_1^2} + 2(1-\sigma)\frac{\partial^2 \tilde\varphi_\perp(x)}{\partial x_1\partial x_2}\frac{\partial^2 z(x;u)}{\partial x_1\partial x_2}\right.$$

$$\left. + \left(\frac{\partial^2 \tilde\varphi_\perp(x)}{\partial x_2^2} + \sigma\frac{\partial^2 \tilde\varphi_\perp(x)}{\partial x_1^2}\right)\frac{\partial^2 z(x;u)}{\partial x_2^2}\right]dx = -\int\limits_D \tilde\varphi_\perp(x)C^* J_{H_0}u(x)\,dx. \qquad (28)$$

С другой стороны, поскольку $\tilde\varphi_\perp$ является решением краевой задачи (4)–(6) при $f = \tilde f$, $h_1 = \tilde h_1$, $h_2 = \tilde h_2$, то для этой функции выполняется интегральное тождество (9), положив в котором $v = z(\cdot;u)$, находим

$$\int\limits_D \left[\left(\frac{\partial^2 z(x;u)}{\partial x_1^2} + \sigma\frac{\partial^2 z(x;u)}{\partial x_2^2}\right)\frac{\partial^2 \tilde\varphi_\perp}{\partial x_1^2} + 2(1-\sigma)\frac{\partial^2 z(x;u)}{\partial x_1\partial x_2}\frac{\partial^2 \tilde\varphi_\perp}{\partial x_1\partial x_2}\right.$$

$$\left. + \left(\frac{\partial^2 z(x;u)}{\partial x_2^2} + \sigma\frac{\partial^2 z(x;u)}{\partial x_1^2}\right)\frac{\partial^2 \tilde\varphi_\perp}{\partial x_2^2}\right]dx = \int\limits_D z(x;u)\tilde f(x)\,dx + \int\limits_\Gamma z(\cdot;u)\tilde h_1\,d\Gamma + \int\limits_\Gamma \frac{\partial z(\cdot;u)}{\partial\nu}\tilde h_2\,d\Gamma. \qquad (29)$$

Учитывая, что левые части (28) и (29) совпадают, получаем

$$-\int\limits_D \tilde\varphi_\perp(x)C^* J_{H_0}u(x)\,dx = \int\limits_D z(x;u)\tilde f(x)\,dx + \int\limits_\Gamma z(\cdot;u)\tilde h_1\,d\Gamma + \int\limits_\Gamma \frac{\partial z(\cdot;u)}{\partial\nu}\tilde h_2\,d\Gamma.$$

Отсюда, в силу (26), находим

$$\inf_{c\in\mathbb{R}} \sup_{\tilde{F}\in G_0,\tilde{\eta}\in G_1} \mathbf{M}|l(\tilde{F}) - \widehat{l(\tilde{F})}|^2 =$$

$$= \inf_{c\in\mathbb{R}} \sup_{\tilde{F}\in G_0} \left| (\tilde{f}, z(\cdot;u)+l_0)_{L^2(D)} + (\tilde{h}_1, z(\cdot;u)+l_1)_{L^2(\Gamma)} + \left(\tilde{h}_2, \frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right)_{L^2(\Gamma)} - c\right|^2$$

$$+ \sup_{\tilde{\eta}\in G_1} \mathbf{M}|(\tilde{\eta}, u)_{H_0}|^2. \tag{30}$$

Для вычисления первого члена в правой части (30) применим обобщенное неравенсто Коши-Буняковского и (15). Имеем

$$\inf_{c\in\mathbb{R}} \sup_{\tilde{F}\in G_0,} \left| (\tilde{f}, z(\cdot;u)+l_0)_{L^2(D)} + (\tilde{h}_1, z(\cdot;u)+l_1)_{L^2(\Gamma)} + \left(\tilde{h}_2, \frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right)_{L^2(\Gamma)} - c\right|^2$$

$$= \inf_{c\in\mathbb{R}} \sup_{\tilde{F}\in G_0,} \Bigg| (z(\cdot;u)+l_0, \tilde{f}-f_0)_{L^2(D)} + (\tilde{h}_1 - h_1^{(0)}, z(\cdot;u)+l_1)_{L^2(\Gamma)} + \left(\tilde{h}_2 - h_2^{(0)}, \frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right)_{L^2(\Gamma)}$$

$$+ (z(\cdot;u)+l_0, \tilde{f}_0)_{L^2(D)} + (h_1^{(0)}, z(\cdot;u)+l_1)_{L^2(\Gamma)} + \left(h_2^{(0)}, \frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right)_{L^2(\Gamma)} - c\Bigg|^2$$

$$\leq \Big\{ (Q^{-1}(z(\cdot;u)+l_0), z(\cdot;u)+l_0)_{L^2(D)} + (Q_1^{-1}(z(\cdot;u)+l_1), z(\cdot;u)+l_1)_{L^2(\Gamma)}$$

$$+ \left(Q_2^{-1}\left(\frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right), \frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right)_{L^2(\Gamma)} \Bigg\}$$

$$\times \Big\{ (Q(\tilde{f}-f^{(0)}), \tilde{f}-f^{(0)})_{L^2(D)} + (Q_1(\tilde{h}_1 - h_1^{(0)}), \tilde{h}_1 - h_1^{(0)})_{L^2(\Gamma)} + (Q_2(\tilde{h}_2 - h_2^{(0)}), \tilde{h}_2 - h_2^{(0)})_{L^2(\Gamma)} \Big\}$$

$$\leq (Q^{-1}(z(\cdot;u)+l_0), z(\cdot;u)+l_0)_{L^2(D)} + (Q_1^{-1}(z(\cdot;u)+l_1, z(\cdot;u)+l_1)_{L^2(\Gamma)}$$

$$+ \left(Q_2^{-1}\left(\frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right), \frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right)_{L^2(\Gamma)}. \tag{31}$$

Непосредсвенной подстановкой убеждаемся, что неравенство (31) обращается в равенство на элементе $\tilde{F}^{(0)} := (\tilde{f}^{(0)}, \tilde{h}_1^{(0)}, \tilde{h}_2^{(0)}) \in G_0$, где

$$\tilde{f}^{(0)}(x) := \frac{1}{d} Q^{-1}(z(x,u)) + l_0) + f_0(x),$$

$$\tilde{h}_1^{(0)} := \frac{1}{d} Q_1^{-1}(z(\cdot;u)+l_1) + h_1^{(0)},\ \tilde{h}_2^{(0)} := \frac{1}{d} Q_2^{-1}\left(\frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right) + h_2^{(0)},$$

$$d = \Big\{ (Q^{-1}(z(\cdot;u)+l_0), z(\cdot;u)+l_0)_{L^2(D)} + (Q_1^{-1}(z(\cdot;u)+l_1), z(\cdot;u)+l_1)_{L^2(\Gamma)}$$

$$+ \left(Q_2^{-1}\left(\frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right), \frac{\partial z(\cdot;u)}{\partial\nu} + l_2\right)_{L^2(\Gamma)} \Bigg\}^{1/2}$$

(тот факт, что элемент $\tilde{F}^{(0)} \in G_0$ следует, в силу равенств (22) – (24)). Поэтому

$$\inf_{c\in\mathbb{R}} \sup_{\tilde{F}\in G_0,} \left| (\tilde{f}, z(\cdot;u)+l_0)_{L^2(D)} + (\tilde{h}_1, z(\cdot;u)+l_1)_{L^2(\Gamma)} + \left(\tilde{h}_2, \frac{\partial z(\cdot;u)}{\partial\nu}+l_2\right)_{L^2(\Gamma)} - c\right|^2$$

$$= \int\limits_D Q^{-1}(z(x;u)+l_0(x))(z(x;u)+l_0(x))dx + \int\limits_\Gamma Q_1^{-1}(z(\cdot;u)+l_1)(z(\cdot;u)+l_2)\,d\Gamma$$

$$+\int\limits_\Gamma Q_2^{-1}\left(\frac{\partial z(\cdot;u)}{\partial\nu}+l_2\right)\left(\frac{\partial z(\cdot;u)}{\partial\nu}+l_2\right)d\Gamma \qquad (32)$$

при

$$c = \int\limits_D (l_0+z(x;u))f^{(0)}(x)dx + \int\limits_\Gamma (l_1+z(\cdot;u))h_1^{(0)}\,d\Gamma + \int\limits_\Gamma \left(l_2+\frac{\partial z(\cdot;u)}{\partial\nu}\right)h_2^{(0)}\,d\Gamma.$$

Аналогично, вычисляя второе слагаемое в правой части (30), получим $\sup_{\tilde{\eta}\in G_1}\mathbf{M}|(\tilde{\eta},u)_{H_0}|^2 = (Q_0^{-1}u,u)_{H_0}$. Отсюда и из (30) и (18) придем к утверждению теоремы. □

Решая задачу оптимального управления (19) – (25), придем к следующему результату.

Теорема 2. *Существует единственная минимаксная оценка выражения $l(\varphi)$, которая может быть представлена в виде*

$$\widehat{\widehat{l(\varphi)}} = (y(\varphi,\eta),\hat{u})_{H_0} + \hat{c}, \qquad (33)$$

где

$$\hat{u} = Q_0Cp, \quad \hat{c} = \int\limits_D (l_0(x)+\hat{z}(x))f^{(0)}(x)\,dx + \int\limits_\Gamma (l_1+\hat{z})h_1^{(0)}\,d\Gamma + \int\limits_\Gamma\left(l_2+\frac{\partial\hat{z}}{\partial\nu}\right)h_2^{(0)}\,d\Gamma, \quad (34)$$

а функции $p \in H^2(D)$ и $\hat{z} \in H^2(D)$ определяются из системы интегро-дифференциальных уравнений

$$\Delta^2\hat{z}(x) = -C^*J_{H_0}Q_0Cp(x) \quad \textit{в} \quad D, \qquad (35)$$

$$N\hat{z}=0, \quad M\hat{z}=0 \quad \textit{на} \quad \Gamma, \qquad (36)$$

$$\int\limits_D Q^{-1}(l_0(x)+\hat{z}(x))\,dx + \int\limits_\Gamma Q_1^{-1}(l_1+\hat{z})\,d\Gamma = 0, \qquad (37)$$

$$\int\limits_D x_1Q^{-1}(l_0+\hat{z}(x))\,dx + \int\limits_\Gamma x_1Q_1^{-1}(l_1+\hat{z}))\,d\Gamma + \int\limits_\Gamma \frac{\partial x_1}{\partial\nu}Q_2^{-1}\left(l_2+\frac{\hat{\partial z}}{\partial\nu}\right)d\Gamma = 0, \quad (38)$$

$$\int\limits_D x_2Q^{-1}(l_0(x)+\hat{z}(x))\,dx + \int\limits_\Gamma x_2Q_1^{-1}\,(l_1+\hat{z})\;d\Gamma + \int\limits_\Gamma \frac{\partial x_2}{\partial\nu}Q_2^{-1}\left(l_2+\frac{\partial\hat{z}}{\partial\nu}\right)d\Gamma = 0, \quad (39)$$

$$\Delta^2p(x) = Q^{-1}(l_0(x)+\hat{z}(x)) \quad \textit{в} \quad D, \qquad (40)$$

$$Np = Q_1^{-1}(l_1 + \hat{z}), \quad Mp = Q_2^{-1}\left(l_2 + \frac{\partial \hat{z}}{\partial \nu}\right) \quad \text{на} \quad \Gamma, \tag{41}$$

$$\int_D (C^* J_{H_0} Q_0 Cp)(x)\, dx = 0, \quad \int_D (C^* J_{H_0} Q_0 Cp)(x) x_1\, dx = 0, \tag{42}$$

$$\int_D (C^* J_{H_0} Q_0 Cp)(x) x_2\, dx = 0. \tag{43}$$

Задача (35)–(43) *однозначно разрешима. Погрешность оценивания* σ *определяется формулой* $\sigma = l(\hat{\hat{P}})^{1/2}$, *где* $\hat{\hat{P}} = \left(Q^{-1}(l_0 + z), Q_1^{-1}(l_1 + z), Q_1^{-1}\left(l_2 + \frac{\partial \hat{z}}{\partial \nu}\right)\right)$.

Альтернативное представление для минимаксной оценки через решение системы интегро-дифференциальных уравнений специального вида, не зависящее от конкретного вида функционала (13), получено в следующей теореме.

Теорема 3. *Минимаксная оценка выражения (13) имеет вид* $\widehat{\widehat{l(F)}} = l(\hat{F})$ *где* $\hat{F} = (\hat{f}, \hat{h}, \hat{h}_2)$, $\hat{f}(x) = Q^{-1}\hat{p}(x) + f^{(0)}(x)$, $\hat{h}_1 = Q_1^{-1}\hat{p} + h_1^{(0)}$, $\hat{h}_2 = Q_2^{-1}\hat{p} + h_2^{(0)}$, *а функция* $\hat{p} \in L^2(\Omega, H^2(D))$ *определяется из решения следующей системы стохастических интегро-дифференциальных уравнений:*

$$\Delta^2 \hat{p}(x;\omega) = C^* J_{H_0} Q_0 (y(\varphi;\eta) - C\hat{\varphi})(x;\omega) \quad \text{в} \quad D, \tag{44}$$

$$N\hat{p}(\cdot;\omega) = 0, \quad M\hat{p}(\cdot;\omega) = 0 \quad \text{на} \quad \Gamma, \tag{45}$$

$$\int_D Q^{-1}\hat{p}(x;\omega)\, dx + \int_\Gamma Q_1^{-1}\hat{p}(\cdot;\omega)\, d\Gamma = 0, \tag{46}$$

$$\int_D x_1 Q^{-1}\hat{p}(x;\omega)\, dx + \int_\Gamma x_1 Q_1^{-1}\hat{p}(\cdot;\omega)\, d\Gamma + \int_\Gamma \frac{\partial x_1}{\partial \nu} Q_2^{-1} \frac{\partial \hat{p}(\cdot;\omega)}{\partial \nu}\, d\Gamma = 0, \tag{47}$$

$$\int_D x_2 Q^{-1}\hat{p}(x;\omega)\, dx + \int_\Gamma x_2 Q_1^{-1}\hat{p}(\cdot;\omega)\, d\Gamma + \int_\Gamma \frac{\partial x_2}{\partial \nu} Q_2^{-1} \frac{\partial \hat{p}(\cdot;\omega)}{\partial \nu}\, d\Gamma = 0, \tag{48}$$

$$\Delta^2 \hat{\varphi}(x;\omega) = Q^{-1}\hat{p}(x;\omega) + f_0(x) \quad \text{в} \quad D, \tag{49}$$

$$N\hat{\varphi}(\cdot;\omega) = Q_1^{-1}\hat{p}(\cdot;\omega) + h_1^{(0)}, \quad M\hat{\varphi} = Q_2^{-1}\frac{\partial \hat{p}(\cdot;\omega)}{\partial \nu} + h_2^{(0)} \quad \text{на} \quad \Gamma, \tag{50}$$

$$\int_D C^* J_{H_0} Q_0 (y(\varphi;\eta) - C\hat{\varphi})(x;\omega)\, dx = 0, \quad \int_D x_1 C^* J_{H_0} Q_0 (y(\varphi;\eta) - C\hat{\varphi})(x;\omega)\, dx = 0, \tag{51}$$

$$\int_D x_2 C^* J_{H_0} Q_0 (y(\varphi;\eta) - C\hat{\varphi})(x;\omega)\, dx = 0, \tag{52}$$

где $\hat{\varphi} \in L^2(\Omega, H^2(D))$, *и равенства* (44) – (52) *выполняются с вероятностью 1. Задача* (44)–(52) *однозначно разрешима.*

Доказательство этой теоремы аналогично доказательству теоремы 1.

Заключение

Результаты данной статьи, таким образом, *состоят в том, что для систем, описываемых краевыми задачами для бигармонического уравнения с граничными условиями типа Неймана, при сформулированных выше ограничениях на параметры этих задач, получены представления для минимаксных прогнозных оценок функционалов от детерминированных данных этих задач и погрешностей оценивания через решения однозначно разрешимых интегро-дифференциальных уравнений специального вида.*

Методика и результаты работы могут быть использованы в дальнейших теоретических и прикладных исследованиях процессов в теории упругости, а также при использовании системного анализа этих процессов.

Список литературы

1. *Наконечний О.Г.* Оптимальне керування та оцінювання в рівняннях із частинними похідними // Київський університет, Київ, 2004 г., 103 с.
2. *Ректорис К.* Вариационные методы в математической физике и технике //М.: Мир, 1985 г., 589 с.
3. *Подлипенко Ю.К., Грищук Н.В.* Минимаксное оценивание решений вырожденных краевых задач Неймана для эллиптических уравнений по наблюдениям, распределенным на системе поверхностей // Системні дослідження і інформаційні технології // №2, 2004 г., с. 104-128.

Статья поступила в редакцию 22.09.2008

Демиденко С.В., Наконечний О.Г. ***Мінімаксні середньоквадратичні оцінки тренду в задачах регресії*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 23-30.

УДК 519.962.22

У цій роботі пропонується підхід до лінійного мінімаксного оцінювання узагальнених поліномів з невідомими параметрами. Лінійні мінімаксні оцінки будуються на основі вимірів з випадковим шумом та квадратичними обмеженнями на параметри. Пропонується визначення верхньої та нижньої мінімаксної оцінок. Сформульовані достатні умови для існування мінімаксних оцінок. Отримані результати ілюструються прикладами.

В работе предлагается подход к линейному минимаксному оцениванию обобщенных полиномов с неизвестными параметрами. Линейные минимаксные оценки строятся на основе измерений с случайным шумом. Вводятся определения верхних и нижних минимаксных оценок. Сформулированы достаточные условия для существования минимаксных оценок. Полученные результаты иллюстрируются примерами.

Поляков Б. Н. ***Эффективный метод ранжирования независимых переменных и отбрасывания несущественных параметров при многофакторном статистическом анализе*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 31-37.

УДК 519.23

Наводяться обґрунтування та пропонуються надійні критерії ранжирування незалежних змінних і відкидання несуттєвих параметрів при багатофакторному статистичному аналізі, ефективність яких ілюструється конкретним прикладом і підтверджується більш ніж 30-літньою практикою успішного проведення статистичних досліджень у машинобудуванні, металургії й медицині.

Приводятся обоснования и предлагаются надёжные критерии ранжирования независимых переменных и отбрасывания несущественных параметров при многофакторном статистическом анализе, эффективность которых иллюстрируется конкретным примером и подтверждается более чем 30-летней практикой успешного проведения статистических исследований в машиностроении, металлургии и медицине.

Анафиев А. С. ***Некоторые положения и задачи теории шаблонов*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 39-45.

УДК 519.7

У роботі розглядаються основні означення, положення та проблеми теорії шаблонів. Вводяться нові типи шаблонів, наводяться деякі властивості поліноміальних шаблонів. Виділяються нові задачі, які грають важливу роль при побудові якісних вирішальних правил.

В работе рассматриваются основные определения, положения и проблемы теории шаблонов. Вводятся новые типы шаблонов, приводятся некоторые свойства полиномиальных шаблонов. Выделяются новые задачи, которые играют важную роль при построении качественных решающих правил.

Блыщик В. Ф. ***Условно контролирующие стратегии и последовательный выбор решений в многошаговой игре с булевыми стратегиями*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 47-51.

УДК 519.83

У статті розглядаються багатово-крокові ігри з булевими стратегіями і послідовним вибором дій гравцями. Вводиться поняття умовно-контролюючої стратегії та на основі цього поняття пропонується алгоритм вибору рішення.

В статье рассматриваются многошаговые игры с булевыми стратегиями и последовательным выбором действий игроками. Вводится понятие условно-контролирующей стратегии и на основе этого понятия предлагается алгоритм выбора решения.

Жук С. М. ***Мінімаксні оцінки розв'язків лінійних операторних рівнянь з лінійним необмеженим оператором у гільбертовому просторі*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 53-60.

УДК 519.962.22

У статті вивчається задача гарантованого оцінювання для лінійного операторного рівняння з необмеженим щільно визначеним оператором у гільбертовому просторі. Розглядаються апріорні середньоквадратичні лінійні мінімаксні оцінки. Для випадку квадратичних обмежень на праві частини та шуми одержано критерій скінченності

мінімаксної похибки середньоквадратичного оцінювання, запропоновано представлення оцінки у вигляді лінійного перетворення від розв'язку системи лінійних операторних рівнянь.

В статье изучается задача гарантированного оценивания для линейного операторного уравнения с неограниченным плотно определенным оператором в гильбертовом пространстве. Рассматриваются априорные среденквадратические линейные минимаксные оценки. Для квадратичных ограничений на правые части и шумы получен критерий конечности минимаксной среднеквадратической ошибки оценивания, предложено представление оценки в виде линейного преобразования решения системы линейных операторных уравнений.

Ежова Е.О., Моттль В.В., Красоткина О.В. ***Обобщение информационного критерия Акаике для выбора значений непрерывных параметров в моделях данных*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 61-70.

УДК 004.9311

Застосування інформаційного критерія Акаике (AIC) для вибору класу моделі з упорядкованої множини вкладених класів моделей обмежено припущенням, що класи визначаються зростаючої розмірністю вектора параметрів. Ми поширили принцип максимуму інформації за Кульбаку, що лежить в основі класичного інформаційного критерію Акаике, на більш широкий клас моделей, в якому розмірність вектора параметрів фіксована, але свобода вибору його значень обмежена системою безперервних вкладених сімейств апріорних плотностей розподілу. Ми проілюстрували застосування узагальненого критерію Акаике на задачі аналізу нестаціонарного сигналу, регресійна модель якого змінюється за часом.

Применение информационного критерия Акаике (AIC) для выбора класса модели из упорядоченного множества вложенных классов моделей ограничено предположением, что классы определяются возрастающей размерностью вектора параметров. Мы распространили принцип максимума информации по Кульбаку, лежащий в основе классичечского инфомационного критерия Акаике, на более широкий класс моделей, в котором размерность вектора параметров фиксирована, но свобода выбора его значений ограничена системой непрерывно вложенных семейств априорных плотностей распределения. Мы проиллюстрировали применение обобщенного критерия Акаике на задаче анализа нестационарного сигнала, регрессионная модель которого меняется во времени.

Воронов А.В. ***К теоретико-методологическим основам искусственного интеллекта*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 71-75.

УДК 004.81:007.5:612.822

Обгрунтовується актуальність дослідження рефлексії об'єктивними методами. Вивчено реагування різних областей кори головного мозку на слова. Виявлені активності в областях F7, F8 і ін. Виявлена послідовність станів, відповідних за збільшення активності кіркових структур. Розвиток робіт по створенню штучного інтелекту зв'язується з виявленням нейрофізіологічних механізмів і побудовою теорії нейробіоуправленія.

Обосновывается актуальность исследования рефлексии объективными методами. Изучено реагирование различных областей коры головного мозга на слова. Обнаружены активности в областях F7, F8 и др. Выявлена последовательность состояний, соответствующих увеличению активности корковых структур. Развитие работ по созданию искусственного интеллекта связывается с выявлением нейрофизиологических механизмов и построением теории нейробиоуправления.

Тышкевич Д.Л. ***О строении остаточного подпространства полуунитарной дилатации линейного ограниченного оператора в банаховом пространстве с внутренним произведением*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 77-92.

УДК 517.983

У даній роботі приведено опис конструкції залишкового підпростору напівунітарної ділатації лінійного обмеженого оператора, який діє у банаховому просторі з індефінітним внутрішнім добутком.

В данной работе приведено описание конструкции остаточного подпространства полуунитарной дилатации линейного ограниченного оператора, действующего в банаховом пространстве с индефинитным внутренним произведением.

Горбатенко М.Ю. ***Оцінювання за зашумленними спостереженнями невідомих даних лінійних еліптичних рівнянь, що допускають змішане варіаційне формулювання*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 93-102.

УДК 519.8

Одержаний новий клас систем варіаційних рівнянь через розв'язки яких виражаються мінімаксні оцінки значень функционалів від невідомих правих частин лінійних еліптичних рівнянь 2-го порядку.

Получен новый класс систем вариационных уравнений через решения которых выражаются минимаксные оценки значений функционалов от неизвестных правых частей линейных эллиптических уравнений 2-го порядка.

Перцов А.С. ***Минимаксное оценивание неизвестных данных краевой задачи для бигармонического уравнения с граничными условиями типа Неймана*** // Таврический вестник информатики и математики. – 2009. – №1. – С. 103-112.

УДК 519.8

Знайдені мінімаксні оцінки функціоналів від невідомих детермінованих даних крайової задачі для бігармонічного рівняння з граничними умовами типа Неймана.

Найдены минимаксные оценки функционалов от неизвестных детерминистических данных краевой задачи для бигармонического уранения с граничными условиями типа Неймана

СПИСОК АВТОРОВ НОМЕРА

Анафиев Айдер Сератович	к. ф.-м. н., доцент кафедры информатики факультета математики и информатики Таврического национального университета им. В. И. Вернадского, секретарь редакции журнала ТВИМ, Украина *e-mail: aydera@mail.ru*
Блыщик Владимир Федорович	к. ф.-м. н., доцент кафедры информатики факультета математики и информатики Таврического национального университета им. В. И. Вернадского, Украина *e-mail: veb@land.ru*
Воронов Алексей Валерианович	к. ф.-м. н., с. н. с. Института автоматики и процессов управления ДВО РАН, ученый секретарь Дальневосточного отделения научного совета РАН по методологии искусственного интеллекта, Россия *e-mail: voronov.ia@mail.ru*
Горбатенко Николай Юрьевич	ассистент факультета прикладной математики Черновицкого национального университета им. Ю. Федьковича, Украина *e-mail: Mikola.Gorbatenko@gmail.com*
Демиденко Сергей Владимирович	аспирант факультета кибернетики Киевского национального университета им. Тараса Шевченко, Украина *e-mail: s.demidenko@gmail.com*
Ежова Елена Олеговна	студентка Московского физико-технического института, Россия *e-mail: lena-ezhova@rambler.ru*
Жук Сергей Николаевич	к. ф.-м. н. научный сотрудник сектора проблем системного анализа факультета кибернетики Киевского национального университета им. Тараса Шевченко, Украина *e-mail: serhiy.zhuk@gmail.com*
Красоткина Ольга Вячеславовна	к. ф.-м. н., доцент Тульского государственного университета, Россия *e-mail: krasotkina@uic.tula.ru*
Моттль Вадим Вячеславович	д. т. н., ВЦ РАН, Россия *e-mail: vmottl@yandex.ru*
Наконечный Александр Григорьевич	д. ф.-м. н., профессор, заведующий кафедрой системного анализа и теории принятия решений Киевского национального университета им. Тараса Шевченко, член редакционного совета журнала ТВИМ, Украина *e-mail: nakonechniy@unicyb.kiev.ua*

Перцов Андрей Сергеевич	ассистент факультета прикладной математики Черновицкого национального университета им. Ю. Федьковича, Украина *e-mail: pertsov@ukr.net*
Поляков Борис Николаевич	д. т. н., профессор, США *e-mail: bpoliakov@hotmail.com*
Тышкевич Дмитрий Леонидович	к. ф.-м. н., доцент кафедры алгебры и функционального анализа факультета математики и информатики Таврического национального университета им. В. И. Вернадского, Украина *e-mail: bpoliakov@hotmail.com*

ДО ВІДОМА АВТОРІВ

Загальні положення

Для опублікування в журналі "Таврійський вісник інформатики і математики" приймаються раніше не опубліковані наукові праці в галузі математики та теоретичної інформатики, згідно зі списком провідних тематичних розділів.
Автору(-ам) потрібно надавати такі документи:

1. Відомості про автора(-ів) (прізвище, ім'я, по батькові, учені ступені та звання, місце роботи та посада, адреси проживання та організації, телефон, факс, адреса електронної пошти тощо).
2. Рецензію сторонньої організації (бажано).
3. Статтю, надруковану на принтері.
4. Файл статті на дискеті 3,5" або надісланий електронною поштою за адресою редакції.

Вимоги до рукописів

1. Основні елементи статті розміщуються у такій послідовності: індекс УДК, ініціали та прізвище автора, назва статті, анотація (до 10 рядків) українською, російською та англійською мовами (анотація повинна містити конкретну інформацію про отримані результати), текст, список літератури.
2. Стаття може бути написана українською, російською або англійською мовою. Обсяг статті повинен не перевищувати 10 сторінок разом з малюнками, таблицями, графіками (не більше трьох) та бібліографією. Стаття повинна бути структурована (поділена на розділи із заголовками).
3. ***Відповідно до постанови Президії ВАК України від 15 січня 2003 року №7-05/1*** текст статті повинен бути викладений лаконічно, зрозуміло і відповідати такій структурній схемі.

   У вступі необхідно чітко виділити (курсивом) такі пункти:

   Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

   Аналіз останніх досліджень і публікацій, в яких започатковано розв'язання даної проблеми і на які спирається автор

   Невирішені раніше частини загальної проблеми, котрим присвячується зазначена стаття

   Формулювання цілей статті (постановка задачі)

   У висновку з данного дослідження необхідно чітко виділити (курсивом)

   результати дослідження *та перспективи подальших розвідок у цьому напрямку*
4. У статті необхідно дотримуватись термінології, прийнятої державним стандартом; використовуючи новий термін або абревіатуру, автор повинен розшифрувати та пояснити їх.
5. Використана література наводиться загальним списком наприкінці статті за порядком посилання на неї в тексті (в квадратних дужках) мовою оригіналу, відповідно до форми Ф23 бюлетеню ВАК України, 2000, №2.

6. Стаття має бути підготовлена за допомогою видавничої системи LATEX з використанням стильового пакету twim.sty, який можно отримати за адресою **www.twim.crimea.edu**.

Робота редакції з авторами

1. Матеріали необхідно надіслати електронною поштою, а також у вигляді "твердої" копії за адресою редакції: **Таврійський національний університет ім. В.І. Вернадського, пр-т Вернадського, 4, м.Симферопіль, Крим, Україна, 95007**, e-mail: **twim_taurida@mail.ru**
2. Редакція залишає за собою право внесення змін редакційного характеру без згоди з автором (-ами).
3. За необхідності автору (-ам) надсилається коректура статті.
4. Остаточне рішення про публікацію приймає редакційна колегія.
5. Рукопис, який надійшов до редакції з порушенням зазначених правил оформлення, не реєструється і не розглядається, а повертається автору (-ам) для доопрацювання.

ДО УВАГИ АВТОРІВ!

Про підвищення вимог до фахових видань, внесених до переліків ВАК України

ПОСТАНОВА
ПРЕЗИДІЇ ВИЩОЇ АТЕСТАЦІЙНОЇ КОМІСІЇ УКРАЇНИ
від 15.01.2003 р. №7-05/1

Необхідною передумовою для внесення видань до переліку наукових фахових видань України є їх відповідність вимогам пункту 7 постанови Президії ВАК України від 10.02.1999 р. №1-02/3 "Про публікації результатів дисертацій на здобуття наукових ступенів доктора і кандидата наук та їх апробацію".

... Редакційним колегіям організувати належне рецензування та ретельний відбір статей до друку. Зобов'язати їх приймати до друку у видання 2003 року та й у подальші роки лише наукові статті, які мають такі необхідні елементи: постановку проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями; аналіз останніх досліджень і публікацій, в яких започатковано розв'язання данної проблеми і на які спирається автор; виділення невирішених раніше частин загальної проблеми, котрим присвячується зазначена стаття; формулювання цілей статті (постановка задачі); виклад основного матеріалу дослідження с повним обгрунтуванням наукових результатів; висновки з даного дослідження і перспективи подальших розвідок у цьому напрямку.

Голова ВАК України В.В.Скопенко

Вчений секретар Л.М.Артюшин

Подписано к печати 15.01.2009. Формат 38х30/2. Бумага тип ОП. Объем 2.5 п.л. Тираж 500 экз. Заказ 335.
Издано в редакционном отделе КНЦ НАНУ
просп. Вернадского, 2, г. Симферополь, АРК, Украина, 95007