Таврический Вестник Информатики и Математики

№ 2 ' 2011

МЕЖДУНАРОДНОЕ НАУЧНО-ТЕОРЕТИЧЕСКОЕ ИЗДАНИЕ
КРЫМСКИЙ НАУЧНЫЙ ЦЕНТР НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК
И МИНИСТЕРСТВА ОБРАЗОВАНИЯ И НАУКИ, МОЛОДЕЖИ И СПОРТА 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 26.05.2010 р. № 1-05/4 журнал "Таврійський вісник информатики та математики" внесено до переліку фахових видань України, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата фізико-математичних наук (01.01 – математика, 01.05 – інформатика і кібернетика).

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

АДРЕС РЕДАКЦИИ:

Крымский научный центр Национальной Академии наук
и Министерства образования и науки, молодежи и спорта 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@tnu.crimea.ua
e-mail (для переписки):	article@tvim.info
сайт журнала:	www.tvim.info

**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

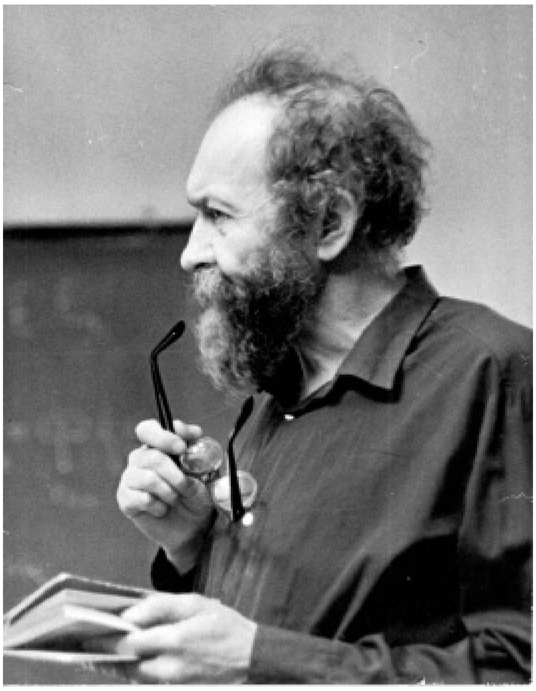
Математическая теория, алгоритмы и системы распознавания образов

**Печатается по решению научно технического Совета
КНЦ НАН и Министерства образования и науки,
молодежи и спорта Украины
Протокол № 8 от 22 декабря 2011 г.**

СОДЕРЖАНИЕ

ДВА ВЫДАЮЩИХСЯ МАТЕМАТИКА, ДВЕ СУДЬБЫ И ДВЕ ЖИЗНИ В НАУКЕ: А. А. ЛЯПУНОВ И А. П. ЕРШОВ

8 ноября 2011 г. исполнилось 100 лет со дня рождения великого математика, основателя Новосибирской физико-математической школы Алексея Андреевича Ляпунова



Выдающийся ученый-математик, обогатившй отечественную науку в области теории множеств, кибернетики и программирования, известный плодотворными приложениями математических методов в различных областях техники и естествознания Алексей Андреевич Ляпунов родился в Москве 8 октября 1911 г.

Свыше сорока лет своей жизни отдал А. А. Ляпунов служению отечественной науке. Лишь однажды был в ней перерыв, когда в годы Великой Отечественной войны он добровольцем ушел на фронт и в качестве офицера артиллерии прошел боевой путь от Крыма до Восточной Пруссии.

Круг научных интересов Алексея Андреевича Ляпунова был настолько широк, что его по праву можно назвать ученым-энциклопедистом. Он не только ориентировался в разных областях науки, но и плодотворно работал во многих из них. Основные труды А. А. Ляпунова относятся к чистой математике, но охватывают также ее прикладную и вычислительную части, приложения к естественным и гуманитарным наукам, философские проблемы естествознания и актуальные проблемы педагогики.

Отец Алексея Андреевича — Андрей Николаевич Ляпунов — математик, получивший образование сначала в Московском, а затем в Гайдельбергском университете, был человеком общительным и широкообразованным. Будучи знатоком и ценителем искусства, отец оказал большое влияние не только на формирование жизненных взглядов, научных и эстетических вкусов А. А. Ляпунова, но и на стиль его общения с людьми.

В 1928 году Алексей Андреевич поступил на физико-математический факультет Московского государственного университета. Учеба в университете не сложилась: отказавшись подписать письмо о сносе в Москве очередных церквей (такие кампании были тогда в моде), он вступил в конфликт с сокурсниками и перестал посещать занятия, за что и был отчислен в конце 1929 г.

Научные интересы Алексея Андреевича формировались под влиянием и непосредственным руководством академика Николая Николаевича Лузина. Заметив незаурядные способности юноши, Н. Н. Лузин приобщил его к работе в области теории множеств. В 1934 г. Алексей Андреевич становится младшим научным сотрудником Отдела теории функций действительного переменного Института математики им. В. А. Стеклова, где сближается со старшими учениками Н. Н. Лузина: Н. К. Бари, Л. В. Келдыш, А. Н. Колмогоровым, М. А. Лаврентьевым, Л. А. Люстерником, Д. Е. Меньшовым, П. С. Новиковым. В 1934-39 гг. Алексей Андреевич публикует ряд работ по дескриптивной теории множеств. Сдав экстерном экзамены по университетским курсам и кандидатские экзамены, он в 1939 г. защищает кандидатскую диссертацию на тему «Об униформизации аналитических дополнений».

После защиты кандидатской диссертации А. А. Ляпунов работает в области приложения теории вероятностей к естествознанию и технике, применения вероятностных методов в теории стрельбы. В 1939-40 гг. Алексей Андреевич по рекомендации академика А. Н. Колмогорова проводит статистическую обработку обширного экспериментального материала по расщеплению наследственных признаков у гибридов, полученного генетиками школы Н. И. Вавилова.

Жизнь в предвоенные годы складывалась нелегко. Сказывалось дворянское происхождение и жизненные убеждения. В 1937 г. Алексей Андреевич Ляпунов был уволен из Института математики «по сокращению штатов» в связи с расформированием отдела Н. Н. Лузина. Два следующих года, не имея постоянной работы, А. А. Ляпунов на временной договорной основе читал лекции, руководил семинаром по теории множеств при Научно-исследовательском институте математики МГУ, выполнял заказные переводы. В 1939 г. он восстанавливается в Институте математики им. В. А. Стеклова в должности старшего научного сотрудника и по совместительству занимает должность доцента в МГПИ им. К. Либкнехта. Здесь он читает лекции по математическому анализу и теории функций, руководит научной работой студентов вместе с В. И. Гливенко и П. С. Новиковым.

Начавшаяся в 1941 г. война застала Алексея Андреевича старшим научным сотрудником Института математики им. В. А. Стеклова. Осенью 1941 г. он участвовал в мероприятиях по противовоздушной обороне Москвы. Затем институт был эвакуирован в Казань. Об этом времени Алексей Андреевич вспоминал так: «Настроение было тяжелое. Научная работа не клеилась. Сотрудники Академии наук, имевшие ученую степень, подлежали бронированию, но три моих младших брата — Аскольд, Ярослав и Андрей были на фронте, и я от бронирования отказался. В марте 1942 г. я был направлен во Владимирское военное училище...». Далее — учеба и преподавание в училище, некоторое время пребывание в резервных формированиях, госпиталь в связи с тяжелым заболеванием сыпным тифом, едва не стоившим молодому лейтенанту жизни. С октября 1943 г. А. А. Ляпунов в качестве командира топографического разведывательного взвода на передовой линии фронта: *он участвует в боях в Крыму при взятии Перекопа и освобождении Керчи, затем — на Украине, в Прибалтике (участвует в боях за освобождение Шауляя) и заканчивает боевой путь в Восточной Пруссии. За участие в боях по освобождению Крыма Алексей Андреевич был награжден орденом Красной Звезды* (1944).

В апреле 1945 г. старшего лейтенанта А. А. Ляпунова отозвали с фронта и направили преподавателем в Артиллерийскую академию им. Ф. Э. Дзержинского в Москву. С академией у Алексея Андреевича Ляпунова связан весьма заметный и плодотворный период жизни, который длился около 5 лет. Вначале он был лаборантом кафедры артиллерийской инструментальной разведки и одновременно преподавателем, а после демобилизации в 1946 г. — старшим преподавателем кафедры математики. С самого начала преподавания в Академии Алексей Андреевич развернул интенсивную работу по перестройке курсов математики на основе последних достижений математической науки. В частности, им был создан новый курс теории стрельбы, основанный на теории вероятности и математической статистике. В эти годы Алексей Андреевич публикует ряд работ по теории стрельбы.

С 1946 г. Алексей Андреевич возобновляет исследования в области чистой математики. Он получает стипендию А. Н. Крылова и поступает в докторантуру Института математики им. В. А. Стеклова АН СССР. В эти годы он выполнил ряд работ по дескриптивной теории множеств, основные результаты которых вошли в его докторскую диссертацию «Об операциях, приводящих к измеримым множествам», которую он защитил в конце 1949 г. В этом же году А. А. Ляпунов начинает работать по совместительству в Институте геофизики АН СССР. Летом 1950 г. он — начальник Северо-Тяньшаньской экспедиции. Предметом геофизических исследований Алексея Андреевича в этот период были повторяемость землетрясений и интерпретация гравитационных наблюдений, а также глубинное сейсмическое зондирование.

В июне 1951 г. Алексей Андреевич возвращается на работу в Математический институт им. В. А. Стеклова АН СССР, а в 1953 г. по приглашению М. В. Келдыша он переходит во вновь созданное Отделение прикладной математики этого Института, *где организует отдел кибернетики. С этого времени кибернетика становится основным делом Алексея Андреевича до последнего дня жизни.* Одновременно с осени 1952 г. А. А. Ляпунов работает на механико-математическом факультете МГУ в качестве профессора кафедр математической логики и вычислительной математики. В 1953 г. он организует в МГУ семинар по программированию и специальный семинар для студентов младших курсов мехмата, в 1954 г. — семинар по исследованию проблем расширения областей применения вычислительных машин. Важным событием в научной жизни стал междисциплинарный семинар по кибернетике, организованный А. А. Ляпуновым в МГУ в 1956 г. Его участниками были математики, экономисты, инженеры, биологи, военные, лингвисты, философы. Этот семинар существовал до 1964 г. Он стал центром зарождения кибернетической мысли в СССР и сыграл большую роль в координации работ по кибернетике и формировании новых направлений исследований. Из числа регулярных участников семинара А. А. Ляпунова впоследствии вышли выдающиеся ученые в области теоретической и прикладной кибернетики: А. П. Ершов, Ю. И. Журавлев, Н. П. Бусленко, О. Б. Лупанов, С. В. Яблонский, М. Л. Цетлин и многие др. В 1956 г. А. А. Ляпунов организует издание серии сборников «Проблемы кибернетики». Сборники быстро получили мировую известность. Многие выпуски и отдельные статьи переведены на английский и немецкий языки. До 1973 г. под редакцией А. А. Ляпунова вышли 29 выпусков «Проблем кибернетики», издание их продолжали ученики Алексея Андреевича. Наряду с этим Алексей Андреевич заботился о переводе зарубежных работ. Многие из них изданы под его редакцией, с его предисловиями и комментариями в виде отдельных монографий, а также в основанной им серии «Кибернетический сборник», которую он редактировал вместе с О. Б. Лупановым.

В 1959 г. по инициативе А. А. Ляпунова при Президиуме АН создается Научный совет по комплексной проблеме «Кибернетика». По предложению Алексея Андреевича председателем Научного совета назначается академик А. И. Берг, а Алексей Андреевич становится его заместителем.

В 1960 г. академики М. А. Лаврентьев и С. Л. Соболев приглашают А. А. Ляпунова переехать в Новосибирск, где за несколько лет до этого было создано Сибирское отделение Академии наук СССР и на берегу Обского моря началось строительство Академгородка. Алексей Андреевич с энтузиазмом принял это предложение, сразу поняв, что это откроет большие возможности для развертывания работ по кибернетике и осуществления педагогических экспериментов на всех уровнях воспитания молодежи — от дошкольного до университетского. После переезда в Новосибирск в 1961 г. со всей присущей ему страстностью и энергией Алексей Андреевич включился в работу по созданию кибернетических научных коллективов в рамках Сибирского отделения АН СССР. Еще ранее по его инициативе в новосибирский Академгородок приехали многие из его учеников и последователей. Алексей Андреевич сыграл определяющую роль в создании отдела кибернетики в Институте математики СО АН СССР; он организовал в Новосибирском университете кафедру математического анализа, а позже — кафедру теоретической кибернетики.

В 1964 г. А. А. Ляпунов был избран членом-корреспондентом Академии наук СССР по отделению математики. Научные, педагогические и организационные заслуги А. А. Ляпунова отмечены правительственными наградами. Он был награжден орденом «Знак Почета» (1953), двумя орденами Красного Знамени (1956 и 1967) и орденом Ленина (1971).

А. А. Ляпунов скоропостижно скончался 23 июня 1973 г. в Москве, куда приехал на общее собрание Академии наук.

В научной деятельности А. А. Ляпунова, несмотря на ее разносторонний характер, можно выделить два этапа: первый, длившийся до начала пятидесятых годов, связан, главным образом, с теорией множеств; второй — с развитием кибернетики. Интерес к теории множеств Андрей Алексеевич пронес через всю жизнь и неоднократно возвращался к занятию ею и в период работы в области кибернетики. Более того, занимаясь кибернетическими проблемами, он зачастую применял теоретико-множественные методы и привлекал к ним внимание своих учеников.

Первый цикл работ А. А. Ляпунова связан с проблемой отделимости и униформизации множеств. Он показал, что для A-множеств имеет место первая теорема о кратной отделимости по отношению к операции предела счетной последовательности множеств (1936 г.), а для CA-множеств не имеет места первая теорема отделимости по отношению к операции верхнего предела. Далее А. А. Ляпунов детально изучает общие законы отделимости и неотделимости по отношению к A-операции. Опираясь на принцип сравнения индексов П. С. Новикова, А. А. Ляпунов доказал первую и вторую теоремы о кратной отделимости для класса A(M) по отношению к A-операции. Результаты, полученные в области B-, A-, C- и CA-множеств, позволили решить до конца некоторые вопросы, относящиеся к изучению природы основных объектов математического анализа.

Ряд существенных результатов был получен А. А. Ляпуновым в области униформизации множеств. А. А. Ляпуновым было предпринято исследование трансфинитных классов R-множеств, получающихся R-операциями нормального ряда (1949 г.). Он существенно продвинул вперед теорию R-множеств и вопрос о расширении теоретико-множественных операций, приводящих к измеримым множествам. Основные результаты его работы в этом направлении представлены в работе «Об операциях, приводящих к измеримым множествам» (1949 г.) и в монографии «R-множества» (1953 г.), представляющей собой систематическое изложение теории R-множеств.

Ряд работ А. А. Ляпунова относится к области метрической теории множеств и посвящен изучению вполне аддитивных вектор-функций множеств и законов распределения случайных величин. Теорема А. А. Ляпунова о множестве значений аддитивной вектор функции множеств, доказанная в 1940 г., получила широкий резонанс и развитие в работах многих исследователей. А. А. Ляпунов показал, что вполне аддитивная вектор-функция, лишенная скачков, определенная на системе подмножеств некоторого множества, инвариантной относительно счетных сумм и пересечений и взятия дополнений и принимающая значения n-мерного эвклидова пространства, имеет выпуклое множество значений. В 1946 г. им было показано, что это свойство теряется, если вместо конечномерного пространства взять бесконечномерное, хотя бы даже компактное пространство. Алексей Андреевич возвращается к анализу вполне аддитивных вектор-функций в 60-е годы. Он публикует две статьи в «Проблемах кибернетики», подчеркивая этим важность разрабатываемого им подхода для решения задач, смежных для кибернетики и математической экономики, в частности, для принятия решений о справедливых дележах.

Основная заслуга А. А. Ляпунова в области информатики состоит в создании им операторного метода программирования. Этот метод получил широкое распространение в реальном программировании и оказал огромное влияние на все последующее развитие теории программирования. Операторный метод был подробно изложен Ляпуновым в курсе лекций, прочитанном в 1952-1953 гг. для студентов кафедры вычислительной математики механико-математического факультета МГУ, и опубликован в работах 1957-1958 гг. В работе «О логических схемах программ» (1958) А.А.Ляпунов дал определение программирования как отдельного научного направления, отличного от классической теории алгоритмов, и первое описание операторного метода. Схема счета и схема программы могут рассматриваться как алгебраические объекты, записанные на некотором формальном языке. Над ними можно выполнять различные эквивалентные преобразования и, следовательно, ставить задачу приведения схемы счета или схемы программы к простейшему или достаточно простому виду. В статье «К алгебраической трактовке программирования» (1962) Алексей Андреевич предложил рассматривать логическую схему программы как класс программ. Конкретная программа получается из схемы, если в схеме некоторым способом интерпретировать символы операторов и предикатов. Две схемы называют эквивалентными, если при любой интерпретации входящих в них переменных (операторов и предикатов) получаются эквивалентные программы. А. А. Ляпуновым была поставлена задача отыскания алгоритма, распознающего эквивалентность схем программ и отыскания полной системы эквивалентных преобразований. Эти задачи были решены учеником А. А. Ляпунова Ю. И. Яновым. Алгебраическая теория программирования, основы которой были заложены в работах Алексея Андреевича и его учеников, получила бурное развитие в СССР и за рубежом и дала серьезные теоретические и прикладные результаты.

А. А. Ляпунову принадлежит идея автоматического программирования, т.е. создания программы, которая по сжатой, особым образом записанной информации о задаче строит программу для решения задачи. Сейчас такие, по терминологии А. А. Ляпунова «программирующие программы», принято называть трансляторами. Создание трансляторов, исследование их строения и принципов их работы — это основное направление в современном программировании. Основателем этого направления, безусловно, является А. А. Ляпунов.

Математическая лингвистика и машинный перевод (МП) были теми областями, где Алексей Андреевич видел широкие возможности практического применения развиваемых им методов кибернетического анализа и математической теории программирования. Он понимал, как значительна методологическая ценность исследований в этом направлении, поскольку задачи МП порождают принципиально новый класс кибернетических проблем. Лингвистика и МП привлекли его внимание уже в 1954 г. Алексей Андреевич рассматривал естественные языки, а также искусственные языки разных типов (например, языки программирования) как сложные и разветвленные системы кодирования информации. Разработка рациональных методов перевода текстов с одного языка на другой требует формализации и систематизации основных понятий лингвистики, что позволило бы применять для их анализа строгий математический аппарат. Вклад Алексея Андреевича в развитие указанных областей состоит не столько в получении конкретных результатов, сколько в определении стратегии всего направления, в постановках задач, для решения которых он привлекал лингвистов.

Педагогическая деятельность Алексея Андреевича достигает своей вершины в новосибирском Академгородке, где условия для экспериментирования и пропаганды новых идей были весьма благоприятными. Он был среди инициаторов создания в 1962 г. первой в нашей стране физматшколы-интерната (ФМШ) при Новосибирском университете. Будучи первым председателем Ученого совета ФМШ и активным ее лектором, он оказал большое влияние на становление и развитие этой школы нового типа. Он был также одним из организаторов сибирских математических олимпиад и летних физматшкол в Академгородке. Однако увлечение физматшколой не заслоняло Алексея Андреевича от проблем обычной школы. Он глубоко верил в то, что идеи современной науки — не удел какой-то элиты, и при правильном методическом осмыслении они могут и должны стать достоянием всех учащихся. Поэтому А. А. Ляпунов уделял постоянное внимание преподаванию в средней школе, а в 1972-73 учебном году, несмотря на колоссальную загруженность, начал вести регулярные занятия в 9 классе школы №130 Академгородка. Он намеревался продолжать занятия и в 10 классе; к сожалению, этот интересный эксперимент остался незавершенным. Конкретные соображения Алексея Андреевича о содержании естественно-математических предметов в школе и о методике их преподавания изложены в ряде статей, опубликованных в центральных журналах (в том числе «Математика в школе»), а также в тематических сборниках «Наука и просвещение», издававшихся Научным советом по проблемам образования при Президиуме Сибирского отделения АН. В общих чертах эти соображения созвучны идеям модернизации школьных программ, получившим в последние годы распространение в достаточно широких кругах научной и педагогической общественности. В частности, они касаются преподавания элементов дифференциального и интегрального исчисления на приемлемом интуитивном уровне без предварительной чрезмерной формализации учения о пределах, непрерывности и действительных числах. Алексей Андреевич настаивал также на расширении преподавания комбинаторики и введении на этой основе элементов теории вероятностей и статистики в программы старших классов школы.

19 апреля 2011 г. исполнилось 80 лет со дня рождения академика Андрея Петровича Ершова

Андрей Петрович, к сожалению, ушёл из жизни в самом расцвете творческих сил (8 декабря 1988 г.), успев, однако, при этом сделать очень многое. Он является пионером программирования и создателем Сибирской школы информатики. Его работы положили начало развитию языков программирования не только в СССР, но и во всём мире.

Еще студентом МГУ, под влиянием А. А. Ляпунова он увлекся программированием. Окончив университет, А. П. Ершов поступил на работу в Институт точной механики и вычислительной техники — организацию, в которой складывался один из первых советских коллективов программистов. В 1957 г. его назначают заведующим отделом автоматизации программирования во вновь созданном Вычислительном центре АН СССР. В связи с образованием Сибирского отделения АН СССР по просьбе директора Института математики СО АН СССР академика С.Л.Соболева он берет на себя обязанность организатора и фактического руководителя отдела программирования этого института, а затем переходит в Вычислительный центр СО РАН.

Фундаментальные исследования А. П. Ершова в области схем программ и теории компиляции оказали заметное влияние на его многочисленных учеников и последователей. Книга А. П. Ершова «Программирующая программа для электронной вычислительной машины БЭСМ» была одной из первых в мире монографий по автоматизации программирования. За существенный вклад в теорию смешанных вычислений А. П. Ершов был удостоен премии имени академика А. Н. Крылова.

Работы Ершова по технологии программирования заложили основы этого научного направления в СССР. Язык программирования АЛЬФА и оптимизирующий Альфа-транслятор, первая советская система разделения времени АИСТ-0, система учебной информатики «Школьница», система подготовки печатных изданий «Рубин», многопроцессорная рабочая станция «МРАМОР» — все эти проекты были инициированы А. П. Ершовым и выполнялись под его руководством.

Благодаря уникальным способностям научного предвидения А. П. Ершов одним первых осознал ключевую роль вычислительной техники в прогрессе науки и общества. Его блестящие идеи заложили основу для развития в СССР таких научных направлений, как параллельное программирование и искусственный интеллект. Более 35 лет тому назад он начал эксперименты по преподаванию программирования в средней школе, которые привели к введению курса информатики и вычислительной техники в средние школы страны и обогатили нас тезисом «программирование — вторая грамотность».

Трудно переоценить роль А. П. Ершова как организатора науки: он принимал самое активное участие в подготовке множества международных конференций и конгрессов, был редактором или членом редколлегий как русских журналов — «Микропроцессорные средства и системы», «Кибернетика», «Программирование», так и международных — Acta Informatica, Information Processing Letters, Theoretical Computer Science.

При подготовке юбилейной статьи использованы материалы публикаций:

1. Журавлев Ю. И. Краткий очерк научной, педагогической и общественной деятельности / Журавлев Ю. И., Бусленко Н. П., Гливенко Е. В., Ляпунова Н. А., Сосюра О. В., Трахтенброт Б. А. // Алексей Андреевич Ляпунов / Рос. Академия наук; сост. Р. И. Кузьменко и Н. А. Ляпунова. — М.: Наука, 1996. — (Материалы к биобиблиографии ученых. Сер. Математических наук; Вып. 19). — С. 8-15.
2. Андрей Петрович Ершов // Наука в Сибири (Еженедельная газета Сибирского отделения РАН), 21 апреля 2011 г.

УДК 517.97

ОБ ОДНОЙ ЗАДАЧЕ ИДЕНТИФИКАЦИИ ДЛЯ ОДНОМЕРНОГО НЕЛИНЕЙНОГО СТАЦИОНАРНОГО УРАВНЕНИЯ КВАЗИОПТИКИ

© Ибрагимов Н. С.

Бакинский Государственный Университет
Кафедра экономической информатики
ул. Академика Захид Халилова 23, AZ 1148, г. Баку, AZ-1073/1,
Азербайджанская Республика
e-mail: *ns.ibragimov@gmail.com*

Abstract. In this paper we consider the problem of identification problem for one-dimensional nonlinear stationary equation of quasi optics of determining the coefficients in the class of square-integrable functions, where the nonlinear part include purely imaginary coefficient. This proved the existence and uniqueness of solution of the identification problem. In addition, the theorem of existence and uniqueness of solution of the corresponding direct problem is proved.

Введение

Известно, что нелинейное стационарное уравнение квазиоптики возникает при изучении распространения световых волн в нелинейных средах, когда процесс распространения не зависит от времени [1]. При прохождении световых волн через нелинейную среду коэффициенты уравнения, характеризующие показатели преломления и поглощения среды, могут оказаться неизвестными функциями [1, стр.233]. Поэтому изучение подобных процессов делает актуальными исследования обратных задач или задач идентификации об определении коэффициентов нелинейного стационарного уравнения квазиоптики.

Данная работа посвящена изучению задачи идентификации для одномерного нелинейного стационарного уравнения квазиоптики об определении коэффициентов уравнения в классе квадратично суммируемых функций, когда нелинейная часть уравнения содержит чисто мнимый коэффициент. Следует отметить, что задачи идентификации об определении только начального фазового профиля распространения световых волн с известными показателями или только показателя преломления среды с известным начальным фазовым профилем ранее подробно изучены, например, в работах [1 – 11] и др.

1. Постановка задачи

Пусть $l > 0, L > 0$ – заданные числа, $0 \leq x \leq l$, $0 \leq z \leq L$, $\Omega_z = (0,l) \times (0,z)$, $\Omega = \Omega_L$; $L_p(0,l)$ – лебегово пространство измеримых функций на $(0,l)$, суммируемых со со степенью $p \geq 1$; $C^k([0,L],B)$ – банахово пространство, состоящее из всех определенных и $k \geq 0$ раз непрерывно дифференцируемых на $[0, L]$ функций со значениями в банаховом пространстве B; $W_p^k(0,l)$, $W_p^{k,m}(\Omega)$ – соболевы пространства функций с обобщенными производными порядка $k \geq 0$ по переменной x и $m \geq 0$ по переменной z, соответственно, которые суммируемы со степенью $p \geq 1$, $\overset{\circ}{W}{}_2^1(0,l)$ – подпространство пространства $W_2^1(0,l)$, элементы которого обращаются в нуль на концах отрезка $[0,l]$, $\overset{\circ}{W}{}_2^2(0,l) \equiv W_2^2(0,l) \cap \overset{\circ}{W}{}_2^1(0,l)$; символ $\overset{\circ}{\forall}$ означает, что данное свойство имеет место для почти всех значений переменной величины. Ниже всюду постоянные, не зависящие от оцениваемых величин, обозначим через $c_j, j = 0, 1, 2, ...$

Рассмотрим процесс, состояние которого описывается следующим одномерным нелинейным стационарным уравнением квазиоптики [1]:

$$i\frac{\partial \psi}{\partial z} + \frac{\partial}{\partial x}\left(a_0(x)\frac{\partial \psi}{\partial x}\right) - a(x)\psi + v_0(x)\psi + iv_1(x)\psi + ia_1|\psi|^2\psi = f(x,z), \ \ (x,z) \in \Omega, \quad (1)$$

где $i = \sqrt{-1}$ – мнимая единица, $\psi = \psi(x,z)$ – волновая функция, $a_1 > 0$ – заданное вещественное число, $a_0(x), a(x)$ – заданные вещественнозначные измеримые ограниченные функции, удовлетворяюшие условиям:

$$0 < \mu_0 \leq a_0(x) \leq \mu_1, \ \left|\frac{da_0(x)}{dx}\right| \leq \mu_2, \ \overset{\circ}{\forall}\, x \in (0,l), \mu_0, \mu_1, \mu_2 = const > 0, \quad (2)$$

$$0 \leq a(x) \leq \mu_3, \ \overset{\circ}{\forall}\, x \in (0,l), \mu_3 = const > 0, \quad (3)$$

$f(x,z)$ – заданная комплекснозначная измеримая функция, удовлетворяющая условию:

$$f \in W_2^{0,1}(\Omega), \quad (4)$$

$v_0(x), v_1(x)$ – неизвестные коэффициенты уравнения.

Пусть для уравнения (1) заданы следующие начальное и краевое условия:

$$\psi(x,0) = \varphi(x), \ x \in (0,l), \ \psi(0,z) = \psi(l,z) = 0, \ z \in (0,L), \quad (5)$$

где $\varphi(x)$ – заданная комплекснозначная измеримая функция, удовлетворяющая условию:

$$\varphi \in \overset{\circ}{W}{}_2^2(0,l). \quad (6)$$

Наша цель заключается в определении неизвестных коэффициентов $v_0(x), v_1(x)$ на основе дополнительной информации:

$$\psi(x,L) = y(x), \quad x \in (0,l), \quad (7)$$

где $y = y(x)$ – заданная комплекснозначная функция из $L_2(0,l)$.

Пусть $v = (v_0, v_1)$, где $v_0 = v_0(x), v_1 = v_1(x)$. Элемент $v = v(x)$ будет найден на множестве:

$$V \equiv \left\{ v = (v_0, v_1) : v_m \in L_2(0,l), \|v_m\|_{L_2(0,l)} \le b_m, m = 0, 1, v_1(x) \ge b_2, \overset{\circ}{\forall}\, x \in (0,l) \right\},$$

где $b_m, m = 0, 1, 2$ – заданные положительные числа. Множество V будем называть множеством допустимых элементов. Определение $v = (v_0, v_1)$ из множества V при условиях (1), (5), (7) является задачей идентификации по финальному наблюдению для уравнения квазиоптики вида (1). Вариационная формулировка этой задачи заключается в минимизации функционала:

$$J_\alpha(v) = \|\psi(.,L) - y\|^2_{L_2(0,l)} + \alpha\|v - \omega\|^2_H \tag{8}$$

на множестве V при условиях (1), (5), где $\alpha \ge 0$ – заданное число, $\omega \in H$ – заданный элемент, $H \equiv L_2(0,l) \times L_2(0,l)$. Эту задачу в дальнейшем будем называть задачей идентификации (1), (5), (7).

Задачу об определении функции $\psi = \psi(x,z) \equiv \psi(x,z;v)$ из условий (1), (5) при каждом $v \in V$ будем называть прямой задачей. Под решением этой задачи будем понимать функцию $\psi = \psi(x,z) \equiv \psi(x,z;v)$ из пространства $B_0 \equiv C^0([0,L], \overset{\circ}{W}{}^2_2(0,l)) \cap C^1([0,L], L_2(0,l))$, удовлетворяющую уравнению (1) для любого $z \in [0,L]$ и почти всех $x \in (0,l)$, а условиям (2) для почти всех $x \in (0,l)$ и $z \in (0,L)$, соответственно.

2. Разрешимость прямой задачи

Прямую задачу для одномерного нелинейного стационарного уравнения квазиоптики вида (1) можно рассмотреть как прямую задачу для одномерного нелинейного нестационарного уравнения Шредингера с комплекснозначным квантовомеханическим потенциалом. Известно, что прямая задача, то есть начально-краевая задача для одномерного нелинейного нестационарного уравенеия Шредингера с вещественнозначным квантовомеханическим потенциалом, зависящим от переменной x, ранее изучена в работах [4–6, 8, 10–12] и др., когда потенциал является функцией из класса измеримых ограниченных или квадратично суммируемых функций, имеющих обобщенные производные первого порядка. В настоящей работе потенциал является комплекснозначным и принадлежит классу квадратично суммируемых функций. Поэтому результаты, полученные в этих работах относительно разрешимости начально-краевой задачи для одномерного нелинейного уравнения Шредингера, не достаточны для исследования задачи идентификации (1), (5), (8) по определению комплекснозначных потенциалов и возникает необходимость изучения разрешимости прямой задачи (1), (5) при $v \in V$.

Теорема 1. *Пусть функции $a_0(x), a(x), f(x,z), \varphi(x)$ удовлетворяют условиям (2)–(4), (6). Тогда при каждом $v \in V$ прямая задача (1), (5) имеет единственное решение из пространства B_0 и для этого решения справедлива оценка:*

$$\|\psi(.,z)\|_{\overset{\circ}{W}{}_2^2(0,l)} + \left\|\frac{\partial\psi(.,z)}{\partial z}\right\|_{L_2(0,l)} \le c_0\left(\|\varphi\|_{\overset{\circ}{W}{}_2^2(0,l)} + \|f\|_{W_2^{0,1}(\Omega)} + \|\varphi\|^3_{\overset{\circ}{W}{}_2^1(0,l)} + (\tilde{c}_0)^{3/2}\right) \quad (9)$$

для $\forall z \in [0,L]$, где $c_0 > 0$ – некоторая постоянная, а постоянная $\tilde{c}_0$ определяется формулой:

$$\tilde{c}_0 = \left(\|\varphi\|^2_{\overset{\circ}{W}{}_2^2(0,l)} + \|f\|^2_{W_2^{0,1}(\Omega)} + \|\varphi\|^6_{\overset{\circ}{W}{}_2^1(0,l)}\right).$$

Доказательство. Для доказательства будем использовать метод Галеркина. Возьмем какую-либо фундаментальную в $\overset{\circ}{W}{}_2^2(0,l)$ и ортонормированную в $L_2(0,l)$ систему функций $u_k = u_k(x), k = 1,2,...$, например, систему собственных функций следующей спектральной задачи:

$$LX(x) = \lambda X(x), \quad x \in (0,l), \quad X(0) = X(l) = 0, \quad (10)$$

при $\lambda = \lambda_k, k = 1,2,...$, где оператор L определяется формулой:

$$L = -\frac{d}{dx}\left(a_0(x)\frac{d}{dx}\right) + a(x). \quad (11)$$

Известно, что задача (10) есть спектральная задача, изученная, например, в [13]. Она имеет нетривиальные решения $u_k(x), k = 1,2,...$ при $\lambda = \lambda_k, k = 1,2,...$, образующих спектр задачи (10) и эти решения образуют базис в пространстве $\overset{\circ}{W}{}_2^2(0,l)$ и ради удобства предположим, что эти функции ортонормированы в $L_2(0,l)$:

$$(u_k, u_m)_{L_2(0,l)} = \int_0^l u_k(x)u_m(x)dx = \delta_k^m, \quad k,m = 1,2,..., \quad (12)$$

где δ_k^m символы Кронекера:

$$\delta_k^m = \begin{cases} 1, & k = m, \\ 0, & k \neq m, \end{cases} \quad k,m = 1,2,\ldots.$$

Ясно, что функции $u_k = u_k(x), k = 1,2,...$ ортогональны и в следующем смысле:

$$[u_k, u_m] = L(u_k, u_m) = (u_k, u_m)_{W_2^1(0,l)} =$$

$$= \int_0^l \left(a_0(x)\frac{du_k(x)}{dx}\frac{du_m(x)}{dx} + a(x)u_k(x)u_m(x)\right)dx = \lambda_k\delta_k^m, \ \ k,m = 1,2,..., \quad (13)$$

$$\{u_k, u_m\} = (Lu_k, Lu_m)_{L_2(0,l)} = (u_k, u_m)_{W_2^2(0,l)} = \lambda_k^2\delta_k^m, \ \ k,m = 1,2,\ldots. \quad (14)$$

В силу предположения $a(x) \geq 0$ все собственные значения $\lambda_k, k = 1, 2, ...$ вещественны, положительны и расположены в порядке возрастания и $\lambda_k \to +\infty$ при $k \to \infty$. Предположим также, что

$$\|u_k\|_{\overset{\circ}{W}{}_2^2(0,l)} \leq d_k, \quad k = 1, 2, ..., \tag{15}$$

где $d_k, k = 1, 2, ...,$ – положительные постоянные.

Приближенное решение прямой задачи (1), (2) будем искать в виде:

$$\psi^N(x, z) = \sum_{k=1}^{N} c_k^N(z) u_k(x), \tag{16}$$

где $c_k^N(z) = (\psi^N(., z), u_k)_{L_2(0,l)}, k = \overline{1, N}$, определяются из условий:

$$i \frac{d}{dz}(\psi^N(., z), u_k)_{L_2(0,l)} - (L\psi^N(., z), u_k)_{L_2(0,l)} + (v_0 \psi^N(., z), u_k)_{L_2(0,l)} +$$

$$+ (iv_1 \psi^N(., z), u_k)_{L_2(0,l)} + (ia_1 \left|\psi^N(., z)\right|^2 \psi^N(., z), u_k)_{L_2(0,l)} = f_k(z), \quad k = \overline{1, N}, \tag{17}$$

$$c_k^N(0) = (\psi^N(., 0), u_k)_{L_2(0,l)} = (\varphi^N, u_k)_{L_2(0,l)} = \varphi_k^N, \quad k = \overline{1, N}, \tag{18}$$

$f_k(z) = (f(., z), u_k)_{L_2(0,l)}, \ \ k = \overline{1, N}$.

Система (17) есть не что иное, как система N нелинейных обыкновенных дифференциальных уравнений. При $v \in V$ из предположений (2) – (4) и из свойства функций $u_k(x), k = 1, 2, ...,$ следует, что второе-пятое слагаемые левой части, а также правая часть являются непрерывными на каждом множестве $\{z \in [0, L], |c_k^N| \leq const\}$ функции $z, c_k^N, k = \overline{1, N}$. Поэтому для существования по крайней мере одного решения задачи Коши (17), (18) на всем отрезке $[0, L]$ достаточно знать, что все ее возможные решения равномерно ограничены на $[0, L]$. Такая ограниченность следует из следующей леммы:

Лемма 1. *Для любого $v \in V$ галеркинские приближения удовлетворяют следующей оценке:*

$$\sum_{k=1}^{N} \left|c_k^N(z)\right|^2 + \sum_{k=1}^{N} \left|\frac{dc_k^N(z)}{dz}\right|^2 \leq \left\|\psi^N(., z)\right\|^2_{\overset{\circ}{W}{}_2^2(0,l)} + \left\|\frac{\partial \psi^N(., z)}{\partial z}\right\|^2_{L_2(0,l)} \leq$$

$$\leq c_1 \left(\|\varphi\|^2_{\overset{\circ}{W}{}_2^2(0,l)} + \|f\|^2_{W_2^{0,1}(\Omega)} + \|\varphi\|^6_{\overset{\circ}{W}{}_2^1(0,l)} + (\tilde{c}_0)^3 \right) \tag{19}$$

для $\forall z \in [0, L]$.

Для доказательства этой леммы была использована методика работы [11].

Рассмотрим функции:

$$l_{N,k}(z) = (\psi^N(., z), u_k)_{L_2(0,l)}, \quad N, k = 1, 2, \dots . \tag{20}$$

Из равенства (20) и оценки (19), а также из ортонормированности функций $u_k(x), k = 1, 2, ...$, следует справедливость неравенств:

$$|l_{N,k}(z)| \leq c_2, \quad \left|\frac{dl_{N,k}(z)}{dz}\right| \leq c_3, \quad \forall z \in [0, L], \quad N, k = 1, 2, \ldots . \tag{21}$$

Используя систему (17), а также предположение (15), можем установить справедливость соотношений:

$$|l_{N,k}(z + \Delta z) - l_{N,k}(z)| \leq c_4 d_k |\Delta z|, \tag{22}$$

$$\left|\frac{dl_{N,k}(z + \Delta z)}{dz} - \frac{dl_{N,k}(z)}{dz}\right| \leq c_5 d_k |\Delta z|^{1/2} \tag{23}$$

для $\forall z \in [0, L], N, k = 1, 2, \ldots$. Следуя неравенствам (21) – (23) заключаем, что семейство функций $l_{N,k}(z), N, k = 1, 2, ...$ и их производных $\dfrac{dl_{N,k}(z)}{dz}, N, k = 1, 2, ...$ равномерно ограничены на отрезке $[0, L]$ и равностепенно непрерывны при фиксированном k и при произвольном $N \geq k$ на этом отрезке. Тогда можем выбрать подпоследовательность $N_m, m = 1, 2, ...$ по которой функции $l_{N_m,k}(z), m = 1, 2, ...$ и их производные $\dfrac{dl_{N_m,k}(z)}{dz}, N_m, k = 1, 2, ...$ сходятся равномерно на отрезке $[0, L]$ к непрерывным функциям $l_k(z)$ и $\dfrac{dl_k(z)}{dz}$, соответственно, для каждого $k = 1, 2, \ldots$. Функции $l_k(z), k = 1, 2, ...$ и их производные определяют функции:

$$\psi(x, z) = \sum_{k=1}^{\infty} l_k(z) u_k(x), \tag{24}$$

$$\frac{\partial \psi(x, z)}{\partial z} = \sum_{k=1}^{\infty} \frac{dl_k(z)}{dz} u_k(x). \tag{25}$$

Действуя аналогично работам [9,10], доказываем, что подпоследовательности $\{\psi^{N_m}(x, z)\}$, $\left\{\dfrac{\partial \psi^{N_m}(x, z)}{\partial z}\right\}$ сходятся слабо в $\overset{\circ}{W}{}_2^2(0, l)$ и $L_2(0, l)$ к функциям $\psi(x, z)$, $\dfrac{\partial \psi(x, z)}{\partial z}$, соответственно, равномерно относительно $z \in [0, L]$. Нетрудно установить, что $\{\psi^{N_m}(x, z)\}$, принадлежит пространству B_0. Тогда можем утверждать, что предельная функция $\psi(x, z)$, определенная формулой (24), также принадлежит пространству B_0 и для этой функции справедлива оценка:

$$\|\psi(., z)\|^2_{\overset{\circ}{W}{}_2^2(0,l)} + \left\|\frac{\partial \psi(., z)}{\partial z}\right\|^2_{L_2(0,l)} \leq$$

$$\leq c_1 \left(\|\varphi\|^2_{\overset{\circ}{W}{}_2^2(0,l)} + \|f\|^2_{W_2^{0,1}(\Omega)} + \|\varphi\|^6_{\overset{\circ}{W}{}_2^1(0,l)} + (\tilde{c}_0)^3 \right), \ \forall z \in [0, L], \tag{26}$$

которая следует непосредственно из оценки (19) с переходом к нижнему пределу по подпоследовательности $N = N_m, m = 1, 2, \ldots$. Из этой же оценки следует оценка (9).

Далее действуя, как и в работе [11], доказываем, что предельная функция $\psi(x,z)$ из B_0 является решением прямой задачи (1), (5) при каждом $v \in V$. Кроме того, используя оценку (9) и методику доказательства единственности решения начально-краевых задач, как и в работах [6, 11], устанавливаем единственность решения прямой задачи (1), (5). Теорема 1 доказана. □

3. Существование и единственность решения задачи идентификации

Теперь изучим вопрос существования и единственности решения задачи идентификации (1), (5), (8).

Теорема 2. *Пусть выполнены условия теоремы 1 и $y \in L_2(0,l)$. Тогда существует плотное подмножество G пространства H такое, что для любого $\omega \in G$ и при $\alpha > 0$ задача идентификации (1), (5), (8) имеет единственное решение.*

Доказательство. Сперва докажем непрерывность функционала:

$$J_0(v) = \|\psi(.,L) - y\|^2_{L_2(0,l)} \tag{27}$$

на множестве V. Пусть $\delta v \in H$ – приращение любого элемента $v \in V$ такое, что $v+\delta v \in V$ и $\delta\psi = \delta\psi(x,z) \equiv \psi(x,z;v+\delta v) - \psi(x,z;v)$, где $\psi(x,z;v)$ – решение прямой задачи (1), (5) при $v \in V$. Из условий (1), (5) следует, что функция $\delta\psi = \delta\psi(x,z)$ является решением следующей начально-краевой задачи:

$$\begin{gathered} i\frac{\partial\delta\psi}{\partial z} + \frac{\partial}{\partial x}\left(a_0(x)\frac{\partial\delta\psi}{\partial x}\right) - a(x)\delta\psi + \\ +(v_0(x)+\delta v_0(x))\delta\psi + i(v_1(x)+\delta v_1(x))\delta\psi + ia_1(|\psi_\delta|^2 + |\psi|^2)\delta\psi + \\ +ia_1\psi_\delta\psi\delta\bar{\psi} = -\delta v_0(x)\psi - i\delta v_1(x)\psi, \quad (x,z) \in \Omega, \end{gathered} \tag{28}$$

$$\delta\psi(x,0) = 0,\ x \in (0,l), \quad \delta\psi(0,z) = \delta\psi(l,z) = 0,\ z \in (0,L), \tag{29}$$

где $\psi_\delta = \psi_\delta(x,z) \equiv \psi(x,z;v+\delta v)$.

Теперь оценим решение этой начально-краевой задачи. С этой целью обе части уравнения (28) умножим на функцию $\delta\bar{\psi}(x,z)$ и полученное равенство проинтегрируем по области Ω_z. Тогда, из полученного равенства вычитая его комплексное сопряжение и используя начальное условие из (29), имеем:

$$\|\delta\psi(.,z)\|^2_{L_2(D)} + 2\int\limits_{\Omega_z}(v_1(x)+\delta v_1(x))|\delta\psi|^2 dx d\tau = -2\int\limits_{\Omega_z}\mathrm{Im}(\delta v_0(x)\psi\delta\bar{\psi})dx d\tau -$$

$$-2\int\limits_{\Omega_z}\mathrm{Re}(\delta v_1(x)\psi\delta\bar{\psi})dxd\tau-2a_1\int\limits_{\Omega_z}\mathrm{Re}\left[(|\psi_\delta|^2\psi_\delta-|\psi|^2\psi)\delta\bar{\psi}\right]dxd\tau$$

для $\forall z\in[0,L]$. Отсюда в силу условия $v+\delta v\in V$ получим справедливость неравенства:

$$\|\delta\psi(.,z)\|^2_{L_2(D)}+2b_2\int\limits_{\Omega_z}|\delta\psi|^2dxd\tau\le 2\int\limits_{\Omega_z}|\delta v_0(x)||\psi||\delta\psi|dxd\tau+2\int\limits_{\Omega_z}|\delta v_1(x)||\psi||\delta\psi|dxd\tau+$$

$$+2a_1\int\limits_{\Omega_z}\left(|\psi_\delta|^2+|\psi|^2+|\psi_\delta\psi|\right)|\delta\psi|^2dxd\tau,\quad\forall z\in[0,L]. \tag{30}$$

В силу теоремы вложения пространство $C^0([0,L],\overset{\circ}{W}{}^1_2(0,l))$ вложено в пространство $C(\bar{\Omega})$. Поэтому можем написать следующие неравенства:

$$\max_{(x,z)\in\bar{\Omega}}|\psi(x,z)|^2\le c_6\|\psi\|^2_{C^0([0,L],\overset{\circ}{W}{}^1_2(0,l))},\quad \max_{(x,z)\in\bar{\Omega}}|\psi_\delta(x,z)|^2\le c_6\|\psi_\delta\|^2_{C^0([0,L],\overset{\circ}{W}{}^1_2(0,l))}. \tag{31}$$

Учитывая эти неравенства и оценку (9), из неравенства (30) с помощью неравенства Коши-Буняковского и леммы Гроноулла имеем:

$$\|\delta\psi(.,z)\|^2_{L_2(0,l)}\le c_7\left(\|\delta v_0\|^2_{L_2(0,l)}+\|\delta v_1\|^2_{L_2(0,l)}\right),\quad\forall z\in[0,L]. \tag{32}$$

Теперь рассмотрим приращение функционала $J_0(v)$ на любом элементе $v\in V$. В силу формулы (27) имеем:

$$\delta J_0(v)=J_0(v+\delta v)-J_0(v)=2\int\limits_0^l\mathrm{Re}[(\psi(x,L)-y(x))\delta\bar{\psi}(x,L)]dx+\|\delta\psi(.,L)\|^2_{L_2(0,l)}. \tag{33}$$

Применяя неравенство Коши-Буняковского и используя оценки (9), (32), а также условие $y\in L_2(0,l)$, получим справедливость оценки:

$$|\delta J_0(v)|\le c_8\left(\|\delta v_0\|_{L_2(0,l)}+\|\delta v_1\|_{L_2(0,l)}+\|\delta v_0\|^2_{L_2(0,l)}+\|\delta v_1\|^2_{L_2(0,l)}\right). \tag{34}$$

Из оценки (34) следует соотношение:

$$\delta J_0(v)\to 0\quad\text{при}\quad\|\delta v\|_H\to 0 \tag{35}$$

для $\forall v\in V$. Отсюда получаем, что функционал $J_0(v)$ непрерывен на множестве V. Ввиду $J_0(v)\ge 0,\forall v\in V$ функционал снизу ограничен на множестве V. Кроме того, нетрудно установить, что множество V является замкнутым, ограниченным и выпуклым множеством пространства H. А пространство H является равномерно выпуклым пространством [14], поскольку H является гильбертовым. Следовательно, выполняются все условия теоремы работы [15]. Тогда в силу этой же теоремы заключаем, что существует плотное подмножество G из пространства H такое, что для любого

$\omega \in G$ при $\alpha > 0$ задача идентификации (1), (5), (8) имеет единственное решение. Теорема 2 доказана. $\square$

Теперь покажем, что задача идентификации (1), (5), (8) при $\alpha > 0$ и $\forall \omega \in H$ имеет хотя бы одно решение.

Теорема 3. *Пусть выполнены условия теоремы 2. Тогда задача идентификации (1), (5), (8) имеет хотя бы одно решение при $\alpha \geq 0$ и $\forall \omega \in H$.*

Доказательство. Возьмем любую минимизирующую последовательность $\{v^k\} \in V$:

$$\lim_{k\to\infty} J_\alpha(v^k) = J_{\alpha^*} = \inf_{v\in V} J_\alpha(v). \tag{36}$$

Положим $\psi_k = \psi_k(x,z) \equiv \psi(x,z;v^k), k = 1,2,\ldots$. Ввиду того, что элемент v^k при каждом k принадлежит множеству V, в силу теоремы 1 можем утверждать, что прямая задача (1), (2) при каждом k имеет единственное решение из B_0 и справедлива оценка:

$$\|\psi_k(.,z)\|_{\overset{\circ}{W}{}_2^2(0,l)} + \left\|\frac{\partial\psi_k(.,z)}{\partial z}\right\|_{L_2(0,l)} \leq$$

$$\leq c_0\left(\|\varphi\|_{\overset{\circ}{W}{}_2^2(0,l)} + \|f\|_{W_2^{0,1}(\Omega)} + \|\varphi\|^3_{\overset{\circ}{W}{}_2^1(0,l)} + (\tilde{c}_0)^{3/2}\right),\ \ k = 1,2,... \tag{37}$$

для $\forall z \in [0,L]$, где правая часть оценки не зависит от k.

Поскольку множество V есть замкнутое ограниченное и выпуклое множество рефлексивного пространства H, то из последовательности $\{v^k\}$ можно извлечь подпоследовательность, которую для простоты изложения снова обозначим через $\{v^k\}$, что

$$v_m^k \to v_m \quad \text{слабо в} \quad L_2(0,L), \tag{38}$$

$m = 0,1$ при $k \to \infty$. Кроме того, в силу структуры множества V нетрудно доказать, что оно является слабо замкнутым множеством, то есть $v \in V$. Поэтому можем написать следующее предельное соотношение:

$$\int_0^l v_m^k(x)q(x)dx \to \int_0^l v_m(x)q(x)dx,\ \ m = 0,1 \tag{39}$$

для любой функции $q \in L_2(0,l)$ при $k \to \infty$.

Из оценки (37) следует, что последовательность $\{\psi_k(x,z)\}$ равномерно ограничена в норме пространства B_0. Тогда из этой последовательности можно извлечь подпоследовательность, которую для простоты изложения снова обозначим через $\{\psi_k(x,z)\}$, что

$$\psi_m^k(.,z) \to \psi_m(.,z) \text{ слабо в } W_2^2(0,l), \tag{40}$$

$$\frac{\partial\psi_k(.,z)}{\partial z}\to\frac{\partial\psi(.,z)}{\partial z}\text{ слабо в }\ L_2(0,l) \tag{41}$$

для каждого $z\in[0,L]$ при $k\to\infty$.

Ясно, что каждый элемент подпоследовательности $\{\psi_k(x,z)\}\in B_0$ удовлетворяет тождеству:

$$\int_0^l\left(i\frac{\partial\psi_k(x,z)}{\partial z}+\frac{\partial}{\partial x}\left(a_0(x)\frac{\partial\psi_k(x,z)}{\partial x}\right)-a(x)\psi_k(x,z)+v_0^k(x)\psi_k(x,z)+iv_1^k(x)\psi_k(x,z)+\right.$$

$$\left.+ia_1|\psi_k(x,z)|^2\psi_k(x,z)-f(x,z)\right)\bar\eta(x)dx=0,\quad k=1,2,... \tag{42}$$

для $\forall z\in[0,L]$ и для любой функции $\eta\in L_2(0,l)$, начальному условию:

$$\psi_k(x,0)=\varphi(x),\quad \overset{\circ}{\forall}\,x\in(0,l),k=1,2,... \tag{43}$$

и краевому условию:

$$\psi_k(0,z)=\psi_k(l,z)=0,\quad \overset{\circ}{\forall}\,z\in(0,L),k=1,2,\dots. \tag{44}$$

Используя предельные соотношения (40), (41) , имеем:

$$\int_0^l\left(i\frac{\partial\psi_k(x,z)}{\partial z}+\frac{\partial}{\partial x}\left(a_0(x)\frac{\partial\psi_k(x,z)}{\partial x}\right)-a(x)\psi_k(x,z)\right)\bar\eta(x)dx\to$$

$$\to\int_0^l\left(i\frac{\partial\psi(x,z)}{\partial z}+\frac{\partial}{\partial x}\left(a_0(x)\frac{\partial\psi(x,z)}{\partial x}\right)-a(x)\psi(x,z)\right)\bar\eta(x)dx \tag{45}$$

для каждого $z\in[0,L]$ и для любой функции $\eta\in L_2(0,l)$ при $k\to\infty$.

Теперь докажем, что имеют место соотношения:

$$\int_0^l v_m^k(x)\psi_k(x,z)\bar\eta(x)dx\to\int_0^l v_m(x)\psi(x,z)\bar\eta(x)dx,\ \ m=0,1, \tag{46}$$

$$\int_0^l a_1|\psi_k(x,z)|^2\psi_k(x,z)\bar\eta(x)dx\to\int_0^l a_1|\psi(x,z)|^2\psi(x,z)\bar\eta(x)dx \tag{47}$$

для каждого $z\in[0,L]$ и для любой функции $\eta\in L_2(0,l)$ при $k\to\infty$.

Ясно, что имеют место равенства:

$$\int_0^l v_m^k(x)\psi_k(x,z)\bar\eta(x)dx=\int_0^l(v_m^k(x)-v_m(x))\psi(x,z)\bar\eta(x)dx+$$

$$+\int_0^l v_m^k(x)(\psi_k(x,z)-\psi(x,z))\bar{\eta}(x)dx+\int_0^l v_m(x)\psi(x,z)\bar{\eta}(x)dx, m=0,1, k=1,2,... \quad (48)$$

для каждого $z \in [0,L]$ и для любой функции $\eta \in L_2(0,l)$. Ввиду того, что для каждого $z \in [0,L]$ функция $\psi(x,z)$ принадлежит пространству $W_2^1(0,l)$, а пространство $W_2^1(0,l)$ вложено в $L_\infty(0,l)$, для функции $\eta \in L_2(0,l)$ имеем:

$$q(.,z)=\psi(.,z)\bar{\eta} \in L_2(0,l), \quad z \in [0,L].$$

С учетом этого и предельного соотношения (39) получим:

$$\int_0^l v_m^k(x)\psi(x,z)\bar{\eta}(x)dx \to \int_0^l v_m(x)\psi(x,z)\bar{\eta}(x)dx, m=0,1 \quad (49)$$

для каждого $z \in [0,L]$ и для любой функции $\eta \in L_2(0,l)$ при $k \to \infty$. В силу теоремы вложения пространство $W_2^1(0,l)$ компактно вложено в пространство $L_\infty(0,l)$. Тогда для слабо сходящейся подпоследовательности $\{\psi_k(x,z)\}$ из $C^0([0,L],W_2^1(0,l))$ справедливо предельное соотношение:

$$\psi_k(.,z) \to \psi(.,z) \text{ сильно в } L_\infty(0,l) \quad (50)$$

для каждого $z \in [0,L]$. Используя (50) и неравенство:

$$\left|\int_0^l v_m^k(x)(\psi_k(x,z)-\psi(x,z))\bar{\eta}(x)dx\right| \le$$

$$\le \|v_m^k\|_{L_2(0,l)}\|\eta\|_{L_2(0,l)}\|\psi_k(.,z)-\psi(.,z)\|_{L_\infty(0,l)}, \ m=0,1 \quad (51)$$

для каждого $z \in [0,L]$, а также условие:

$$\|v_m^k\|_{L_2(0,l)} \le b_m, \quad m=0,1, \quad (52)$$

получим справедливость предельных соотношений:

$$\int_0^l v_m^k(x)(\psi_k(x,z)-\psi(x,z))\bar{\eta}(x)dx \to 0, \quad m=0,1, \quad (53)$$

для каждого $z \in [0,L]$ и для любой функции $\eta \in L_2(0,l)$ при $k \to \infty$. Тогда, используя предельные соотношения (49), (53), если переходить к пределу в равенствах (48), то при $k \to \infty$ получим справедливость предельных соотношений (46). Наряду с этим, используя предельное соотношение (50), нетрудно установить справедливость соотношения (47). Таким образом, с учетом предельных соотношений (45)–(47), если переходить к пределу в интегральном тождестве (42), то при $k \to \infty$ получим справедливость следующего интегрального тождества:

$$\int_0^l \left(i\frac{\partial \psi(x,z)}{\partial z} + \frac{\partial}{\partial x}\left(a_0(x)\frac{\partial \psi(x,z)}{\partial x}\right) - a(x)\psi(x,z) + v_0(x)\psi(x,z) + iv_1(x)\psi(x,z) + \right.$$

$$\left. + ia_1|\psi(x,z)|^2\psi(x,z) - f(x,z)\right)\bar{\eta}(x)dx = 0 \tag{54}$$

для каждого $z \in [0, L]$ и для любой функции $\eta \in L_2(0, l)$. Отсюда следует, что предельная функция $\psi(x, z)$ для каждого $z \in [0, L]$ и для почти всех $x \in (0, l)$ удовлетворяет уравнению (1).

Действуя аналогично работам [9, 11], можем установить, что предельная функция $\psi(x, z)$ удовлетворяет начальному и краевому условиям (5). Кроме того, для предельной функции $\psi(x, z)$ справедлива оценка (9), которая следует из оценки (37) с переходом к нижнему пределу по слабо сходящейся подпоследовательности $\{\psi_k(x, z)\}$ к функции $\psi(x, z)$ для каждого $z \in [0, L]$. Также установили, что $\psi = \psi(x, z) \equiv \psi(x, z; v)$ является решением прямой задачи (1), (5) при $v \in V$. Тогда в силу теоремы 1 предельная функция $\psi(x, z)$ принадлежит пространству B_0.

В силу теоремы вложения пространство B_0 компактно вложено в пространство $C^0([0, L], L_2(0, l))$ (см. [16], [17]). Тогда для слабо сходящейся подпоследовательности $\{\psi_k(x, z)\}$ из B_0 к функции $\psi(x, z)$ имеет место соотношение:

$$\psi_k(., L) \to \psi(., L) \quad \text{сильно в} \quad L_2(0, l) \tag{55}$$

при $k \to \infty$. Используя (55) и слабую полунепрерывность снизу норм в пространствах $L_2(0, l)$ и H при $\alpha \geq 0$ и для любого $\omega \in H$ имеем:

$$J_{\alpha^*} \leq J_\alpha(v) \leq \varliminf_{k\to\infty} J_\alpha(v^k) = \inf_{v\in V} J_\alpha(v) = J_{\alpha^*}.$$

Отсюда следует, что $v = v(x)$ из множества V предоставляет минимум функционалу (8) при условиях (1), (5), то есть $v \in V$ является решением задачи идентификации (1), (5), (8) при $\alpha \geq 0$ и при любом $\omega \in H$. Теорема 3 доказана. $\square$

Список литературы

1. Воронцов М.А. Принципы адаптивной оптики / М.А. Воронцов, В.И. Шмальгаузен – М.: Наука, 1985. — 336 с.
2. Ягубов Г.Я. Задача оптимального управления для уравнения типа Шредингера / Г.Я. Ягубов // В сб.: «Численные методы и матем. обеспечение ЭВМ». — Баку: изд-во АГУ, 1984. — С. 116–125.

3. Шамеева Т.Ю. Об оптимизации в задаче о распространении светового пучка в неоднородной среде / Т.Ю. Шамеева // Вестн. Московск. ун-та. Сер. вычисл. матем. и киберн. — 1985. — №1. — С. 12–19.
4. Ягубов Г.Я. Вариационный метод решения обратной задачи об определении квантовомеханического потенциала / А.Д. Искендеров., Г.Я. Ягубов // ДАН СССР. — 1988. — Т.303, № 5. — С. 1044–1048.
5. Ягубов Г.Я. Оптимальное управление нелинейными квантовомеханическими системами / А.Д. Искендеров., Г.Я. Ягубов // Автоматика и телемехан. — 1989. — № 12. — С. 27–38.
6. Ягубов Г.Я. Оптимальное управление коэффициентом квазилинейного уравнения Шредингера / Г.Я. Ягубов // докт. дисс. — Баку, 1993. — 318 с.;
7. Ягубов Г.Я. О вариационном методе решения многомерной обратной задачи для нелинейного нестационарного уравнения Шредингера / Г.Я. Ягубов, М.А. Мусаева // Изв. АН. Азерб. Сер. физ.-техн.-матем. наук. — 1994. — Т. XV, № 5-6. — С. 56–61.
8. Ягубов Г.Я. Об одной задаче идентификации для нелинейного уравнения Шредингера / Г.Я. Ягубов, М.А. Мусаева // Дифференц. уравнения. — 1997. – Т. 33, № 12. — С. 1691–1698.
9. Искендеров А.Д. Определение потенциала в нестационарном уравнении Шредингера / А.Д. Искендеров // В сб.: «Проблемы матем. моделирования и опт.управления». — Баку, 2001. – С. 6 — 36.
10. Iskenderov A.D. Identification problem for time dependent Schrodinger type equation / Iskenderov A.D. // Proceedings of the Lankaran State University. — Lankaran, 2005. — P. 31–53.
11. Ягубов Г.Я. Оптимальное управление неограниченным потенциалом в многомерном нелинейном и нестационарном уравнении Шредингера / А.Д. Искендеров., Г.Я. Ягубов // Вестник Ленкоранского государственного университета. Серия естественных наук. — Ленкорань, 2007. — С. 3–56.
12. Mahmudov N.M. Solvability of boundary value problems for a Schrodinger equation with pura imaginary coefficient in the nonlinear part of this equation / Mahmudov N.M. // Proc. of IMM of NAS of Azerb. — 2007. — Vol. XXVII. — P. 25–37.
13. Ладыженская О.А. Краевые задачи математической физики / О.А. Ладыженская — М.: Наука, 1973. — 408 с.
14. Иосида К. Функциональный анализ / К. Иосида. — М.: Мир, 1967. — 624 с.
15. Goebel M. On existence of optimal control / M. Goebel // Math. Nachr. — 1978. — Vol. 93. — P. 67–73.
16. Simon J. Compact sets in the space / Simon J. // Ann. Mat. Pura Appl. — 1987. — 146 (4) — P. 65–96.
17. Baudoin L. Regularity for Schrodinger equation with singular potentials and application to bilinear optimal control / Baudoin L., Kavian O., Puel J.-P. // J. Differential Equations. — 2005. – 216. — P. 188–222.

Статья поступила в редакцию 06.06.2011

УДК 519.95

ЭМПИРИЧЕСКОЕ ОБОБЩЕНИЕ И РАСПОЗНАВАНИЕ: КЛАССЫ ЗАДАЧ, КЛАССЫ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ И ПРИМЕНИМОСТЬ ТЕОРИЙ. ЧАСТЬ II

© Донской В. И.

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

Abstract. Classification of pattern recognition problems is offered. This classification is founded on the basic properties of pattern recognition problems. It is shown, a choice of methods of decisions must be coordinated with features of classes of pattern recognition problems.

1. Корректные алгоритмы и задачи распознавания из класса $D/-/R_1T/AS \vee LS/-$

В первой части статьи [1] рассматривались задачи обучения распознаванию в детерминистской постановке и предполагалось, что выбор решающего правила $f : \mathbf{X} \longrightarrow \{0,1\}$ осуществляется из некоторого семейства S и отождествляется с нахождением точного решения системы

$$\begin{cases} f(X_1) = \alpha_1; \\ f(X_2) = \alpha_2; \\ \dots\dots\dots\dots\dots\dots \\ f(X_l) = \alpha_l; \\ f \in S, \end{cases} \qquad (1)$$

Здесь $\omega : \mathbf{X} \longrightarrow \{0,1\}$ — отображение объектов генеральной совокупности $\mathbf{X}$ в значения их основного свойства: $\alpha_q = \omega(X_q)$; $X_q \in \mathbf{X}$; $\alpha_q \in \{0,1\}$; Отображение ω порождает столбец значений этого основного свойства в правой части системы (1) и отражает в таблице обучения $\{(X_q, \tilde{\omega}(X_q)), q = 1, ..., l\}$ закономерность, в соответствии с которой объекты из генеральной совокупности обладают или не обладают основным свойством. Следует подчеркнуть, что в теории машинного обучения и распознавания представляют интерес именно те задачи, в которых закономерность ω существует и отличается от случайной функции, равномерно распределяющей значения $\{0;1\}$ на множестве $\mathbf{X}$. В этом смысле закономерность — это неслучайность, и в правой части системы (1) должны содержаться не какие угодно столбцы, а именно те, которые связаны с объектами выборки некоторой закономерностью.

Если извлеченная из генеральной совокупности в соответствии с моделью R_1T выборка является безошибочной, то выбор в процессе обучения правила распознавания, которое допускает ошибки на этой выборке, представляется бессмысленным. Поэтому в этом случае требуется найти точное решение f^* функциональной системы (1) в некотором классе решений S. Это решение f^* называется *точной настройкой на выборку*. Если система имеет более чем одно решение $f^* \in S$, то при достаточно широком семействе S выбранное правило может давать большую ошибку

$$Er(f^*) = \int\limits_{X \in \mathbf{X}} \big|f^*(X) - f_{\text{o}}(X)\big| dP(X),$$

определяемую отличием выбранного решения f^* от существующего истинного правила классификации f_{o} по вероятностной мере $P(\cdot)$ на генеральной совокупности $\mathbf{X}$. Мера $P(\cdot)$, если она существует, оценивает вероятность появления объектов $X \in \mathbf{X}$. Если такая мера не существует, не имеет смысла для некоторых задач, то ошибка точного на выборке решения f^* в некоторых случаях может иметь место почти всюду на $\mathbf{X}$. Поэтому для того, чтобы выполнялось равенство $f^* = f_{\text{o}}$ при условии, что выборка — безошибочная, решение системы (1) должно быть единственным. Действительно, если существуют два решения этой системы f_1^* и f_2^* — функции, совпадающие на выборке $\tilde{X}_l$, — то их продолжения на $\mathbf{X}$ могут различаться почти всюду (например: $f_2^*(X) = f_1^*(X)$, если $X \in \tilde{X}_l$, и $f_2^*(X) = \bar{f}_1^*(X)$, если $X \in \mathbf{X} \setminus \tilde{X}_l$; $\bar{y}$ — инверсия значения $y \in \{0, 1\}$).

Здесь и в дальнейшем *продолжением* выборки будем называть любую последовательность объектов из $\mathbf{X}$, не содержащую элементов этой выборки.

Представим теперь, что в обучающей информации появились ошибки, которые привели к изменению только столбца $\tilde{\alpha} = (\alpha_1, ..., \alpha_l)^T$ в системе (1) и превращению его в столбец с ошибками $\tilde{\alpha}^E$. Пусть система

$$\begin{cases} f(X_1) = \alpha_1^E; \\ f(X_2) = \alpha_2^E; \\ \dots\dots\dots\dots\dots\dots \\ f(X_l) = \alpha_l^E; \\ f \in S, \end{cases} \tag{2}$$

также имеет решение f^E. Но тогда $f^E \neq f_{\text{o}}$, и в таком случае представляется абсурдной точная настройка алгоритма обучения на выборку. Это приводит к следующему выводу: для любой выборки $\tilde{X}_l$, в которой отражена некоторая закономерность, должны существовать такие двоичные столбецы $\tilde{\alpha}^E$, что по таблице обучения $(\tilde{X}_l, \tilde{\alpha}^E)$

точная настройка является невозможной. В связи с этим представляется важной теорема 2 из первой части статьи, которую можно усилить до необходимого и достаточного условия существования таких столбцов:

Теорема 1. *Пусть в задаче обучения распознаванию решающее правило выбирается из семейства S, и обучающая информация представляется в виде $(\tilde{X}_l, \tilde{\alpha})$. Тогда для любой выборки $\tilde{X}_l$ булевский набор $\tilde{\alpha}$ такой, что точная настройка невозможна, найдется если и только если $VCD(S) < l$.*

Как и в первой части этой статьи [1], ёмкость класса S решающих функций (размерность Вапника-Червоненкиса [2]) обозначается $VCD(S)$.

Будем говорить, что найдено точное решение задачи, если реализован выбор $f^* = f_o$. В этом случае решение f^* системы часто называют *корректным алгоритмом*. Напомним, что в работах Ю. И. Журавлева корректными называются алгоритмы, безошибочно распознающие элементы любой заданной контрольной выборки [3]. Целесообразно уточнить используемые ниже понятия, связанные с корректностью алгоритмов.

Определение 1. Решающее правило (алгоритм) называется *корректным* (на выборке), если оно является точным решением системы (1). Решающее правило (алгоритм) называется *абсолютно точным решением* задачи обучения распознаванию, если позволяет безошибочно определить основное свойство $\alpha = f_o(X)$ для любого объекта X генеральной совокупности $\mathbf{X}$.

Следствие 1. *Корректный на выборке длины l алгоритм, выбранный из семейства S, имеющего ёмкость $VCD(S) \geq l$, может давать ошибку почти всюду на генеральной совокупности объектов $\mathbf{X}$.*

Нужно подчеркнуть, что алгоритмы всюду в этой статье рассматриваются с точностью до классов функциональной эквивалентности. Иначе говоря, одним и тем же считаются все алгоритмы (машины Тьюринга), которые для одной и той же начальной информации (слова на ленте) всегда выдают один и тот же результат.

Теорема 2. *Для того, чтобы выбор корректного на выборке алгоритма из заданного семейства S в задаче обучения распознаванию из класса $D/-/R_1T \vee ST/-/-$ всегда обеспечивал получение абсолютно точного решения, необходимо и достаточно, чтобы в семействе S для любой выборки существовал единственный с точностью до функциональной эквивалентности корректный на этой выборке алгоритм.*

Доказательство. Необходимость. Действительно, если для какой-нибудь выборки в семействе S не существует корректного на ней алгоритма, то некоторое число объектов этой выборки всегда классифицируется неверно. Если же корректность на выборках достигается всегда, но хотя бы для одной выборки — не единственным алгоритмом из S, а хотя бы двумя неэквивалентными алгоритмами, то их продолжения на множестве последовательностей из $\mathbf{X}$ не будут совпадать. Тогда хотя-бы один из них будет давать ошибки на своем продолжении.

Достаточность. Если выбор корректного на выборке алгоритма из заданного семейства S не всегда обеспечивает получение абсолютно точного решения, то для некоторой выборки существует корректный алгоритм, не являющийся абсолютно точным решением. Зафиксируем эту выборку. Постановка задачи предполагает существование абсолютно точного решения. Это точное решение — некоторое правило f_o — также будет корректным на зафиксированной (безошибочной в соответствие с рассматриваемой моделью) выборке. Тогда корректное на ней решение — не единственно. □

В задачах обучения распознаванию предполагается использование только конечных обучающих выборок любой длины. Очевидно, что если в семействе S для любой конечной выборки длины l существует единственный с точностью до функциональной эквивалентности корректный на этой выборке алгоритм, то $VCD(S) < l$.

Следствие 2. *Для того, чтобы выбор корректного на выборке длины l алгоритма из заданного семейства S в задаче обучения распознаванию из класса $D/-/R_1T \vee ST/-/-$ всегда обеспечивал получение абсолютно точного решения, необходимо выполнение условия $VCD(S) < l$.*

Но условие $VCD(S) < l$ не является достаточным. Это сразу же видно из случая, когда при его выполнении семейство S не содержит в себе истинного решающего правила f_o. Очевидно также, что условие $f_o \in S$ является необходимым для осуществления возможности нахождения абсолютно точного решения.

2. Обучение или настройка?

Детерминистская постановка задач обучения распознаванию предполагает существование точного решения f_{o} (истинного решающего правила), согласно которому каждой выборке $\tilde{X}_l$ должен сопоставляться единственный булевый вектор $\tilde{\alpha}^*$, в котором $\alpha_i^* = f_{\text{o}}(X_i), i = 1, \ldots, l$. Пары $< \tilde{X}_l, \tilde{\alpha}^* >$ для каждой выборки $\tilde{X}_l$ определяют закономерность (регулярность), выделяющую вектор $\tilde{\alpha}^*$ из всех остальных $2^l - 1$ соответствий $< \tilde{X}_l, \tilde{\alpha} >$. Выше установлено, что для устранения неоднозначности и обеспечения возможности получить точное решение задачи обучения распознаванию необходимо и достаточно чтобы существовала возможность нахождения точного решения системы (1), и это решение должно быть единственным. При этом обязательно должно выполняться емкостное ограничение $VCD(S) < l$, поскольку в противном случае условие единственности корректного алгоритма для любой выборки длины l немедленно нарушается.

Теперь можно рассмотреть вопрос об *отличии обучения от настройки*. Этот вопрос представляется важнейшим в теории обучения распознаванию.

Если не ограничивать емкость используемого для решения задачи обучения распознаванию семейства S, то всегда можно добиться точной настройки на непротиворечивую начальную информацию (достаточно представить абсурдно-тривиальный пример использования решающего правила как суммы характеристических точек выборки любой конечной длины). В таком случае ни о каком обучении говорить не приходится.

Будем называть *обучением ≪снизу-вверх≫* такой последовательный процесс построения по заданной обучающей выборке решающего правила $\hat{f}_{\text{o}}$, на каждом этапе которого происходит минимальное необходимое усложнение $\hat{f}_{\text{o}}$, обеспечивающее уменьшение ошибок решающего правила на обучающей выборке. При обучении происходит поэтапное усложнение решающего правила и, соответственно, расширение сейства S, которому оно принадлежит. Выбор начального приближения и способы поэтапного усложнения решающего правила определяют алгоритм обучения.

Для обоснования алгоритма обучения ≪снизу-вверх≫ целесообразно приводить теорему о возможности расширения (в процессе выполнения именно этого алгоритма) семейства правил, которому принадлежит вычисляемое решающее правило, до некоторого семейства S_o, содержащего истинное решающее правило f_o и имеющего ёмкость $VCD(S_o) < l$, где l — длина обучающей выборки.

Обучением ≪сверху-вниз≫ будем называть последовательный процесс нахождения решающего правила $\hat{f}_{\text{o}}$, принадлежащего некоторому подклассу минимальной сложности S' из выбранного изначально некоторым способом семейства S, направленный на достижение наибольшей точности правила $\hat{f}_{\text{o}}$ на заданной обучающей выборке.

Для обоснования алгоритма обучения по методу ≪сверху-вниз≫ целесообразно приводить теорему об адекватности изначально заданного семейства — наличии истинного решающего правила в начальном семействе S — и сохранении этого свойства при поэтапном сужении начального семейства.

Комбинированным обучением будем называть процесс построения решающего правила, сочетающий оба метода обучения — $\ll$ снизу-вверх$\gg$ и $\ll$ сверху-вниз $\gg$. Такой процесс аналогичен поиску с возвратом.

Процесс нахождение решения системы (1), отличающийся от обучения, будем называть *настройкой*.

3. Особенности класса $D/-/R_2F \vee SF/-/-$

Если выборка содержит ошибки, то можно считать, что их появление связано с «искажением» правильной выборки, или, говоря иначе, с переходом от правильной — к ошибочной выборке. Будем обозначать такой переход следующим образом: $<\tilde{X}_l, \tilde{\alpha}> \xrightarrow{\text{Er}} <\tilde{X}_l^*, \tilde{\alpha}^{Er}>$. Переход $\tilde{X}_l \xrightarrow{\text{Er}} \tilde{X}_l^*$ можно рассматривать как изменение набора точек выборочного пространства в пределах допустимого множества и говорить о безошибочной паре $<\tilde{X}_l^*, \tilde{\alpha}^*>$, переходящей в пару $<\tilde{X}_l^*, \tilde{\alpha}^{Er}>$. Обозначим $\Delta(Er) = \|\tilde{\alpha}^* - \tilde{\alpha}^{Er}\|$ — число ошибок в векторе $\tilde{\alpha}^{Er}$. Если условие теоремы 1 не выполняется, то из семейства S может быть выбрано решающее правило f^{Er}, точно настроенное на выборку $(\tilde{X}_l^*, \tilde{\alpha}^{Er})$ (эмпирическая ошибка при этом — нулевая: $\hat{\varepsilon^*} = 0$), но истинная ошибка этого правила $\varepsilon = \varepsilon(f^{ER}) \geq \Delta(Er)$.

Если выполняется условие теоремы 2, — точная настройка на ошибочную выборку (пару $<\tilde{X}_l^*, \tilde{\alpha}^{Er}>$) будет невозможна. Настроиться точно можно будет только на безошибочную выборку.

Пусть $\mathfrak{X}$ — такое подмножество элементов обучающей выборки, что их удаление из этой выборки позволяет осуществить точную настройку, но удаление никакого собственного подмножества $\mathfrak{X}' \subset \mathfrak{X}$ из выборки уже не позволяет настроиться точно. Будем называть такой набор детерминированной помехой. Из её определения усматривается переборный *алгоритм фильтрации* (удаления) $\mathfrak{X}$ из обучающей выборки по критерию возможности точной настройки.

Пусть Δ_{max} — наибольшее число ошибок $\Delta(Er)$, порождаемых переходом $<\tilde{X}_l, \tilde{\alpha}> \xrightarrow{\text{Er}} <\tilde{X}_l^*, \tilde{\alpha}^{Er}>$, является изначально заданным параметром задачи. Процесс обучения может состоять из следующих последовательно решаемых подзадач:

1^o Выбор адекватного начального семейства S, которое должно содержать истинное решающее правило f_o, для реализации алгоритма фильтрации.

2^o Выполнение алгоритма фильтрации (глубина перебора определяется параметром Δ_{max}).

3^o Собственно обучение одним из описанных выше методов.

Таким образом, для решения задач из класса $D/-/R_2F \vee SF/-/-$ целесообразно применять корректные алгоритмы.

4. Особенности класса $ND/D_k/-/-/-$

Класс недетерминированных задач обучения распознаванию характеризуется тем, что не имеется никакой информации о законах, определяющих существование и появление той или иной выборки $\tilde{X}_l$ — последовательности элементов из $\mathbf{X}$. Более того, неизвестно: существует ли точное решение f_o или нет.

Полагается, что недетерминированные и стохастические задачи обучения распознаванию принципиально различаются. Информация о существовании вероятностных распределений или более — об их типах — в стохастических задачах в некоторых случаях может дать возможность в явном виде выписать статистически оптимальное решающее правило. В стохастических задачах речь идет о решениях, получаемых с точностью, определяемой заданием вероятностных мер.

Рассмотрим следующую задачу. Пусть $\mathbf{X}^{MT}$ — множество шифров машин Тьюринга. $X(M) \in \mathbf{X}^{MT}$ — шифр машины M — является натуральным числом, которое, в частности,может быть представлено двоичной строкой конечной длины. Машина Тьюринга M называется самоприменимой, если, начав работу над словом $p = X(M)$, являющимся шифром этой машины M, она остановится, выполнив конечное число шагов.

Обозначим решающую функцию, которую должен найти алгоритм обучения, следующим образом:

$$f_o^{sa}(X) = \begin{cases} 1, & \text{если машина } M \text{ самоприменима;} \\ 0, & \text{если машина } M \text{ несамоприменима.} \end{cases}$$

Известно, что данная функция f_o^{sa} не является вычислимой — не существует алгоритма (строго определенного тезисом Черча-Тьюринга), правильно вычисляющего для любого входа $X(M)$ значение $f_o^{sa}(X(M)$. Тем не менее, можно сконструировать обучающую выборку $\mathfrak{M}^{sa} = \mathfrak{M}_1 \cup \mathfrak{M}_o$, состоящую из $|\mathfrak{M}_1| = m_1$ примеров шифров самоприменимых машин Тьюринга и $|\mathfrak{M}_o| = m_o$ примеров несамоприменимых машин. Что же будет, если выборку $\mathfrak{M}^{sa}$ предоставить как начальную информацию — таблицу обучения — для построения алгоритма распознавания свойства самоприменимости? Такого алгоритма в принципе не существует. Тем не менее, алгоритм обучения, выбранный из подходящего для данной задачи семейства и имеющий достаточную ёмкость, может дать в качестве решения частичную функцию $\hat{f}_o^{sa}$, которая безошибочно классифицирует все примеры выборки $\mathfrak{M}^{sa}$.

Приведенный пример принадлежит классу недетерминированных задач: неизвестно, существует ли вообще правильное решение (в данном примере — не существут алгоритмического правильного решения), и неизвестно, существует ли какой-нибудь закон появления объектов генеральной совокупности. Но предикат, определяющий основное свойство (здесь — свойство самоприменимости) и отражающий соответствующую закономерность — существует. Данный пример доказывает следующую теорему.

Теорема 3. *Существуют недетермированные задачи обучения распознаванию, для которых абсолютно точное решающее правило не является вычислимым.*

Для решения задач из класса $ND/D_k/-/-/-$ целесообразно использовать алгоритмы, извлекающие закономерность, которая имеет как можно меньшую колмогоровскую сложность. Действительно, недетерминированность предполагает полное отсутствие сведений о распределении объектов генеральной совокупности и вследствие этого допускает подход к выбору решения, которое можно обосновать как неслучайное.

Для задач из класса $ND/D_k/R_1T \vee ST/-/-$ целесообразно применение алгоритмов наименьшей колмогоровской сложности [5].

5. Классификация длин выборок

Практика применения машинного обучения показала, что одна и та же длина обучающей выборки может в некоторых случаях оказаться достаточной для получения требуемой точности распознавания, а в других случаях — быть слишком короткой. Так, взяв тривиальный пример детерминированной задачи с заведомо линейным но неизвестным решающим правилом-предикатом $f_o(x_1, x_2) =\ll ax_1 + bx_2 = c \gg$ для случая двух переменных-признаков и точную выборку, в которой две точки будут принадлежать классу "лежащих на прямой можно абсолютно точно решить задачу восстановления линейного предиката.

Для обучения многослойных нейронных сетей требуются большие выборки, поскольку нейросетевые семейства решающих правил имеют большую емкость.

Классы длин выборок для рассмотренных выше задач обучения распознаванию должны определяться ситуативно, в зависимости от емкости семейств, из которых в процессе обучения извлекается решающее правило.

Ниже приведена таблица, согласно которой следует определять значение параметра $SLen$ в стандартных кодах $STD/VAR/SFM/Slen/ADI$ задач обучения распознаванию; l — длина обучающей выборки.

Параметр SLen	Значение параметра	Определяющее условие
SS	Малая выборка	$l < VCD(S)$
AS	Средняя выборка	$VCD(S) \leq l < 1.5VCD(S)$
LS	Большая выборка	$l \geq 1.5VCD(S)$

6. Класс $S_k/C/-/-/SI$

Параметр $ADI = SI$ (специальная информация) в стохастических параметрических задачах чаще всего определяет типы используемых вероятностных распределений и, возможно, специфические характеристики параметров (в приведенном ниже примере — равенство ковариационных матриц классов).

Рассмотрим пример k-параметрической стохастической задачи обучения распознаванию объектов двух классов, которая хорошо изучена в теории статистических решений.

Пусть согласно дополнительной информации условные вероятности появления в выборке объектов каждого из двух классов $J \in \{0, 1\}$ имеют многомерное нормальное распределение

$$p(X|J) = \frac{1}{(2\pi)^{n/2}|\Sigma_J|^{1/2}} \exp\{-\frac{1}{2}(X - M_J)^T \Sigma_J^{-1}(X - M_J)\},$$

где $X \in \mathbb{R}^n$; M_J и $\sum_J$ — математическое ожидание и ковариационные матрицы двух классов $J = 1, 2$. Пусть также известны так называемые априорные вероятности появления объектов каждого из классов: p_0 и p_1. Известно, что оптимальная (минимизирующая средний риск ошибки) дискриминантная функция в случае равных ковариационных матриц $\sum_0 = \sum_1 = \sum$ является линейной и имеет вид [4]:

$$g(X) = X^T \sum\nolimits^{-1}(M_0 - M_1) - \frac{1}{2}M_0^T \sum\nolimits^{-1} M_0 + \frac{1}{2}M_1^T \sum\nolimits^{-1} M_1 + \ln(\frac{p_0}{p_1}). \qquad (3)$$

Соответствующая решающая функция f_g имеет вид

$$f_g(X) = \begin{cases} 0, & g(X) < 0; \\ 1, & g(X) \geq 0. \end{cases}$$

Решение приведенной задачи, когда задана только обучающая выборка $(\tilde{X}_l, \tilde{\alpha})$длины l, состоит в нахождении по этой выборке статистических оценок $\hat{p}_0, \hat{p}_1, \hat{M}_0, \hat{M}_1, \hat{\Sigma}$ и вычислении $\hat{g}(X)$ по формуле (3).

В рассматриваемой задаче, очевидно, не существует абсолютно точного решения f_o, но при точно заданных векторах математических ожиданий, априорных вероятностей и ковариационной матрице — соответствующая статистическая задача принятия решений имеет точное вероятностное решение f_g (*с точностью до заданной вероятностной меры — распределений и априорных вероятностей*). В постановке обучения распознаванию наилучшее решение $f_{\hat{g}}$ является известным и требует только вычисления статистических оценок $\hat{p}_0, \hat{p}_1, \hat{M}_0, \hat{M}_1, \hat{\Sigma}$.

Никакой корректный алгоритм для решения приведенной статистической задачи, разумеется, не подходит. Действительно, разделяющая поверхность $\hat{g}(X)$, соответствующая наилучшему решающему правилу, является линейной, в то же время классы пересекаются, и выборка, вообще говоря, может оказаться не разделимой линейно. Тогда корректный алгоритм построит нелинейное правило распознавания, заведомо худшее, чем $f_{\hat{g}}$.

Обобщим этот вывод на случай произвольной стохастической задачи обучения распознаванию с двумя пересекающимися классами. Среди всевозможных решающих правил для такой задачи обязательно существует правило, минимизирующее вероятность ошибки или заданную функцию потерь (взвешенную функцию ошибки). Будем обозначать такое наилучшее правило ${f_o}^S$, а соответствующую ему дискриминантную функцию обозначим ${g_o}^S$. Очевидно, что любой корректный алгоритм, примененный к рассматриваемой задаче, определит решающее правило, вообще говоря, отличающееся от ${f_o}^S$, поскольку точки обучающей выборки могут быть расположены «по разные стороны» дискриминантной функции ${g_o}^S$ произвольным образом. Следовательно, корректные алгоритмы для решения таких задач не подходят.

Для стохастических параметрических задач распознавания ёмкость класса, которому принадлежит дискриминантная функция, вообще говоря, не имеет значения; важно лишь то, чтобы эта функция минимизировала средний риск ошибки. Эта функция уже определена стохастическими параметрами задачи, и её не требуется *отыскивать* ни в каком классе.

7. Стохастические непараметрические задачи обучения распознаванию $(STD = S)$

С непараметрическими задачами распознавания дело обстоит иначе. Вероятностные распределения предполагаются существующими, но они неизвестны; их восстановление по обучающей выборке как правило приводит к не менее сложным задачам, чем задача обучения распознаванию в классической постановке.

Для задач рассматриваемого класса всегда можно полагать существование некоторой решающей функции ${f_o}^S$ (дискриминантной функции ${g_o}^S$) наилучшей в статистическом смысле. Эта функция является неизвестной, и алгоритм обучения, конечно, должен находить её наилучшее приближение. Понятно, что такой алгоритм вовсе не обязан быть корректным на выборке. Но должен ли он давать на этой выборке минимальную эмпирическую ошибку, т.е. иметь на ней как можно более близкую к точной настройку?

Учитывая результаты рассмотрения параметрических стохастических задач, можно предположить, что для рассматриваемого класса задач обучения перенастройка (выбор корректного или с очень малой эмпирической ошибкой алгоритма) может привести к большим ошибкам классификации объектов, не принадлежащих обучающей выборке. По-видимому, это связано с тем, что неизвестная дискриминантная функция ${g_o}^S$ (если она бейесовская, минимизирующая средний риск, т.е. статистически оптимальная) должна быть полиномом невысокой степени для неизвестных, но существующих многоэкстремальных (и, тем более, одноэкстремальных) вероятностных распределений.

Представляется целесообразным пытаться искать решающее правило как можно более близкое к бейесовскому классификатору — по максимуму апостериорной условной вероятности класса. Для некоторых семейств моделей распознавания доказаны теоремы о качестве приближения отыскиваемых решающих правил к бейесовскому. Именно такие модели наиболее пригодны для работы с непараметрическими стохастическими задачами обучения распознаванию.

Нужно обратить внимание на то, что в ряде случаев, оценивая эмпирический риск, на самом деле осуществляют оценку наихудшего правила из используемого семейства S. Подчеркнём — в определённом выбранном для решения задачи семействе правил S. Так в работе [2] эмпирический риск оценивается как равномерное по всему семейству S уклонение частоты ошибки решающего правила A на выборке от вероятности ошибки *этого же самого правила* A:

$$\sup_{A \in S} |\nu(A) - P(A)|.$$

А на самом деле необходим выбор правила A, наиболее близкого к неизвестному наилучшему правилу ${f_o}^S$, которое для задач рассматриваемого класса является статистически оптимальным.

Заключение

В статье предложено выделить среди задач обучения распознаванию по эмпирической информации такие классы однотипных задач, для решения которых можно осуществить обоснованный выбор метода решения. Целесообразность такого подхода объясняется тем, что достаточно богатый набор моделей и методов решения задач распознавания образов, созданный более чем за полвека, не всегда правильно используется.

Построенная классификация задач дополнена указаниями методов обучения распознаванию, наиболее подходящих для каждого конкретного класса задач.

Направления дальнейшей работы связаны с уточнением классификатора задач и дальнейшим углублённым исследованием подходов к обоснованию применимости или неприменимости различных математических моделей обучения к задачам из введенных в работе типовых классов.

Список литературы

1. *Донской В. И.* Эмпирическое обобщение и распознавание: классы задач, классы математических моделей и применимость теорий. Часть I / В. И. Донской // Таврический вестник информатики и математики. — 2010. — №1. — С.15 — 23.
2. *Вапник В. Н.* Теория распознавания образов / В. Н. Вапник, А. Я. Червоненкис. — М.: Наука, 1974. — 416 с.
3. *Журавлев Ю.И.* Об алгебраическом подходе к решению задач распознавания или классификации / Ю. И. Журавлев // Проблемы кибернетики. — Вып.33.— М.: Наука, 1978.— С. 5—68.
4. *Нильсон Н.* А Обучающиеся машины / Н. Нильсон. — М.:Мир, 1967. — 180 с.
5. *Donskoy V. I.* The Estimations Based on the Kolmogorov Complexity and Machine Learning from Examples / V. I. Donskoy // Proceedings of the Fifth International Conference "Neural Networks and Artificial Intelligence" (ICNNAI'2008). — Minsk: INNS. — 2008. — P. 292 — 297.

Статья поступила в редакцию 10.12.2011

УДК 519.872

ИССЛЕДОВАНИЕ РАБОТЫ ПВО ИЗ ДВУХ ЗРК

© Коваленко А. И., Марянин Б. Д., Смолич В. П.

Таврический национальный университет им. В. И. Вернадского
факультет математики и информатики
пр-т Вернадского,4, г.Симферополь, Крым, Украина, 95007
E-mail: *svp54@mail.ru*

Abstract. A queueing system of type M/G/2/0 is considered. An only customer in service engages both servers. Two customers are served separately. There is a limit on service time. The equilibrium probabilities of the system are obtained in the article.

Введение

В последнее время для исследования надёжностных характеристик систем массового обслуживания чаще привлекаются асимптотические и статистические методы исследования, т.к. аналитическое моделирование приводит к достаточно сложным системам интегро-дифференциальных уравнений, к решению которых приходится применять либо приближённые либо численные методы. Авторы предлагают задачу, математическое моделирование которой не является громоздким и позволяет получать аналитическое решение и вывод точных вероятностных характеристик функционирования соответствующей системы в стационарном режиме.

1. Постановка задачи

Рассмотрим военный объект, безопасность которого обеспечивается двумя различными ЗРК (зенитно-ракетными комплексами). Бомбардировщики, атакующие объект, образуют простейший поток с интенсивностью λ. Время нахождения самолёта над зоной обслуживания — случайная величина γ, распределённая по экспоненциальному закону с параметром μ. Время «обслуживания», самолёта i-й линией (ЗРК) — непрерывная случайная величина ω_i с интенсивностью $\mu_i(x)$, $i = 1, 2$. Ракетные установки (линии обслуживания) могут выходить из строя в моменты обслуживания заявок (подвергаться бомбовым ударам). Потоки поломок линий простейшие с интенсивностями для каждой из линий равными α, если в зоне обслуживания одна заявка и 2α — если две. Время восстановления (ремонта) i-й линии — непрерывная случайная величина θ_i с интенсивностью $\beta_i(x)$.

Правила обслуживания. Поступившая заявка (бомбардировщик) начинает обслуживаться (обстреливаться) немедленно двумя линиями. Если в момент поступления очередной заявки в системе на обслуживании находится одна заявка, то одна из линий переключается на обслуживании новой заявки, при этом, если времена обслуживания текущей заявки были одинаковы ($\omega_1 = \omega_2$), то переключается первая линия, если же времена обслуживания заявки были разными ($\omega_1 \neq \omega_2$), то переключается линия, имевшая большее время обслуживания. Заявка, поступившая в момент, когда в системе обслуживаются две заявки, теряется (самолёт беспрепятственно пролетает зону обслуживания). Если одна из двух обслуживаемых заявок уходит из системы либо в результате окончания обслуживания (самолёт сбит), либо в результате истечения времени γ её пребывания в зоне обслуживания (самолёт вылетел невредимым из зоны обстрела), то ведущая её линия переключается на помощь для обслуживания оставшейся в системе заявки. Заявка, поступившая в момент обслуживания одной из линий предыдущей заявки при неисправной другой линии (или при неисправных обеих линиях), теряется. Восстановленная линия немедленно включается на помощь линии, ведущей обслуживание.

2. Математическое моделирование

Пронумеруем состояния системы двумя индексами: первый индекс несёт информацию о количестве исправных линий, второй – о количестве заявок в системе.

- $(2,0)$ — обе линии исправны и в системе нет заявок
- $(2,1_{eq})$ — обе линии ведут обслуживание одной заявки с равными временами обслуживания ($\omega_1 = \omega_2$)
- $(2,1_{neq})$ — обе линии обслуживают одну заявку с неравными временами обслуживания ($\omega_1 \neq \omega_2$)
- $(2,2)$ — обе линии обслуживают две заявки (каждая свою)
- $(1_1,1)$ — исправна только первая линия и она обслуживает одну заявку (вторая линия восстанавливается)
- $(1_2,1)$ — исправна только вторая линия и она занята обслуживанием (первая линия на ремонте)
- $(1_1,0)$ — в системе нет заявок, первая линия исправна, вторая на ремонте
- $(1_2,0)$ — в системе нет заявок, вторая линия исправна, первая восстанавливается
- $(0,0)$ — обе линии ремонтируются, в системе нет заявок

Пусть $\xi(t)$ – случайный процесс, описывающий эволюцию системы, фазовое пространство которого состоит из девяти состояний системы (i,j).

Введём функции:

$$p_{i,j}(t) := \mathbb{P}\{\xi(t) = (i,j)\},$$

$$Q_{2,1_{eq}}(t,x) := \mathbb{P}\{\xi(t) = (2,1_{eq}),\ \omega_1 < x, \omega_1 = \omega_2\}, \qquad q_{2,1_{eq}}(t,x) := \frac{\partial Q_{2,1_{eq}}(t,x)}{\partial x}$$

$$Q_{2,1_{neq}}(t,x,y) := \mathbb{P}\{\xi(t) = (2,1_{neq}),\ \omega_1 < x,\ \omega_2 < y\}, \quad q_{2,1_{neq}}(t,x,y) := \frac{\partial^2 Q_{2,1_{neq}}(t,x,y)}{\partial x \partial y}$$

$$Q_{2,2}(t,x,y) := \mathbb{P}\{\xi(t) = (2,2),\ \omega_1 < x,\ \omega_2 < y\}, \qquad q_{2,2}(t,x,y) := \frac{\partial^2 Q_{2,2}(t,x,y)}{\partial x \partial y}$$

$$Q_{1_1,1}(t,x,z) := \mathbb{P}\{\xi(t) = (1_1,1),\ \omega_1 < x,\ \theta_2 < y\}, \qquad q_{1_1,1}(t,x,z) := \frac{\partial^2 Q_{1_1,1}(t,x,z)}{\partial x \partial z}$$

$$Q_{1_2,1}(t,y,z) := \mathbb{P}\{\xi(t) = (1_2,1),\ \omega_2 < y,\ \theta_1 < z\}, \qquad q_{1_2,1}(t,y,z) := \frac{\partial^2 Q_{1_2,1}(t,y,z)}{\partial y \partial z}$$

$$Q_{1_1,0}(t,z) := \mathbb{P}\{\xi(t) = (1_1,0),\ \theta_2 < z\}, \qquad q_{1_1,0}(t,z) := \frac{\partial Q_{1_1,0}(t,z)}{\partial z}$$

$$Q_{1_2,0}(t,z) := \mathbb{P}\{\xi(t) = (1_2,0),\ \theta_1 < z\}, \qquad q_{1_2,0}(t,z) := \frac{\partial Q_{1_2,0}(t,z)}{\partial z}$$

$$Q_{0,0}(t,x,y) := \mathbb{P}\{\xi(t) = (0,0),\ \theta_1 < x,\ \theta_2 < y\}, \qquad q_{0,0}(t,x,y) := \frac{\partial^2 Q_{0,0}(t,x,y)}{\partial x \partial y}$$

Поскольку система имеет конечное число сообщающихся состояний, существует стационарный режим, т.е. при $t \to +\infty$ существуют пределы всех определённых выше функций. Введём обозначения:

$$\lim_{t\to+\infty} p_{i,j}(t) = p_{i,j} \qquad \lim_{t\to+\infty} q_{i,j}(t,...) = g_{i,j}(...)$$

Заметим, что имеют место соотношения:

$$p_{2,1_{eq}} = \int\limits_0^\infty g_{2,1_{eq}}(x)\,dx,\ p_{2,1_{neq}} = \int\limits_0^\infty dx \int\limits_0^\infty g_{2,1_{neq}}(x,y)\,dy, \ldots, p_{0,0} = \int\limits_0^\infty dx \int\limits_0^\infty g_{0,0}(x,y)\,dy$$

Вероятностные рассуждения и предельные переходы приводят к следующей системе интегро-дифференциальных уравнений и граничных условий:

$$\lambda p_{2,0} = \mu(p_{2,1_{eq}} + p_{2,1_{neq}}) + \int\limits_0^\infty dx \int\limits_0^\infty g_{2,1_{neq}}(x,y)(\mu_1(x) + \mu_2(y))\,dy +$$

$$+\int\limits_0^\infty g_{2,1_{eq}}(x)(\mu_1(x)+\mu_2(x))\,dx+\int\limits_0^\infty g_{1_1,0}(z)\beta_2(z)\,dz+\int\limits_0^\infty g_{1_2,0}(z)\beta_1(z)\,dz,$$

$$g'_{2,1_{eq}}(x)+(\lambda+2\alpha+\mu+\mu_1(x)+\mu_2(x))g_{2,1_{eq}}(x)=0, \qquad g_{2,1_{eq}}(0)=\lambda p_{2,0}$$

$$\frac{\partial g_{2,1_{neq}}(x,y)}{\partial x}+\frac{\partial g_{2,1_{neq}}(x,y)}{\partial y}+(\lambda+2\alpha+\mu+\mu_1(x)+\mu_2(y))g_{2,1_{neq}}(x,y)=0,$$

$$g_{2,1_{neq}}(0,y)=\int\limits_0^\infty g_{2,2}(x,y)(\mu+\mu_1(x))dx+\int\limits_0^\infty g_{1_2,1}(y,z)\beta_1(z)dz,$$

$$g_{2,1_{neq}}(x,0)=\int\limits_0^\infty g_{2,2}(x,y)(\mu+\mu_2(y))dy+\int\limits_0^\infty g_{1_1,1}(x,z)\beta_2(z)dz,$$

$$\frac{\partial g_{2,2}(x,y)}{\partial x}+\frac{\partial g_{2,2}(x,y)}{\partial y}+(4\alpha+2\mu+\mu_1(x)+\mu_2(y))g_{2,2}(x,y)=0,$$

$$g_{2,2}(0,y)=\lambda g_{2,1_{eq}}(y)+\lambda\int\limits_y^\infty g_{2,1_{neq}}(x,y)dx, \qquad g_{2,2}(x,0)=\lambda\int\limits_x^\infty g_{2,1_{neq}}(x,y)dy,$$

$$\frac{\partial g_{1_1,1}(x,z)}{\partial x}+\frac{\partial g_{1_1,1}(x,z)}{\partial z}+(\alpha+\mu+\mu_1(x)+\beta_2(z))g_{1_1,1}(x,z)=0,$$

$$g_{1_1,1}(0,z)=\lambda g_{1_1,0}(z), \qquad g_{1_1,1}(x,0)=\alpha g_{2,1_{eq}}(x)+2\alpha\int\limits_0^\infty g_{2,2}(x,y)\,dy+\alpha\int\limits_0^\infty g_{2,1_{neq}}(x,y)\,dy,$$

$$\frac{\partial g_{1_2,1}(y,z)}{\partial y}+\frac{\partial g_{1_2,1}(y,z)}{\partial z}+(\alpha+\mu+\mu_2(y)+\beta_1(z))g_{1_2,1}(y,z)=0,$$

$$g_{1_2,1}(0,z)=\lambda g_{1_2,0}(z), \qquad g_{1_2,1}(y,0)=\alpha g_{2,1_{eq}}(y)+2\alpha\int\limits_0^\infty g_{2,2}(x,y)\,dx+\alpha\int\limits_0^\infty g_{2,1_{neq}}(x,y)\,dx,$$

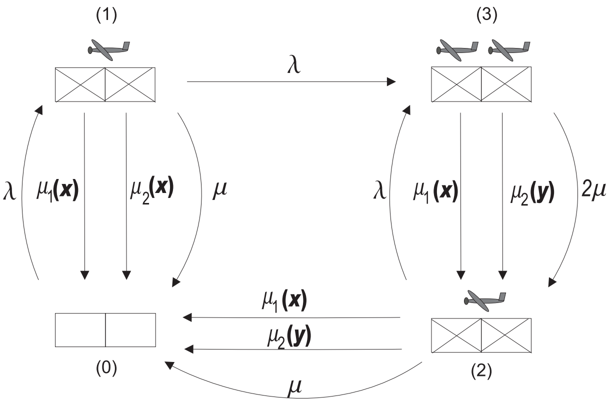
$$g'_{1_1,0}(z)+(\lambda+\beta_2(z))g_{1_1,0}(z)=\int\limits_0^\infty g_{1_1,1}(x,z)(\mu+\mu_1(x))dx+\int\limits_0^\infty g_{0,0}(x,z)\beta_1(x)dx,\ g_{1_1,0}(0)=0,$$

$$g'_{1_2,0}(z)+(\lambda+\beta_1(z))g_{1_2,0}(z)=\int\limits_0^\infty g_{1_2,1}(y,z)(\mu+\mu_2(y))dy+\int\limits_0^\infty g_{0,0}(z,y)\beta_2(y)dy,\ g_{1_2,0}(0)=0,$$

$$\frac{\partial g_{0,0}(x,y)}{\partial x}+\frac{\partial g_{0,0}(x,y)}{\partial y}+(\beta_1(x)+\beta_2(y))g_{0,0}(x,y)=0,$$

$$g_{0,0}(0,y)=\alpha\int\limits_0^\infty g_{1_1,1}(x,y)dx, \quad g_{0,0}(x,0)=\alpha\int\limits_0^\infty g_{1_2,1}(y,x)dy.$$

3. Основные результаты



Авторам не удалось решить систему уравнений и получить точные выражения для вероятностей $p_{i,j}$ состояний системы в стационарном режиме в такой общей постановке. Рассмотрим эту задачу в частном случае, когда линии не выходят из строя, т.е. $\alpha = 0$, $\beta_1(z) = 0$, $\beta_2(z) = 0$, а также при $\mu_2(x) = \mu_2 = \text{const}$. Число состояний при этом уменьшается до четырёх. Перенумеруем их: $(2,0) = (0)$, $(2,1_{eq}) = (1)$, $(2,1_{neq}) = 2$, $(2,2) = (3)$ (см. рисунок).

Система соответствующих уравнений также упрощается:

$$\lambda p_0 = \int\limits_0^\infty g_1(x)(\mu_1(x)+\mu_2)\,dx + \mu(p_1+p_2) + \int\limits_0^\infty dx \int\limits_0^\infty g_2(x,y)(\mu_1(x)+\mu_2)\,dy,$$

$$g_1'(x) + (\lambda+\mu+\mu_1(x)+\mu_2)g_1(x) = 0, \quad g_1(0) = \lambda p_0,$$

$$\frac{\partial g_2(x,y)}{\partial x} + \frac{\partial g_2(x,y)}{\partial y} + (\lambda+\mu+\mu_1(x)+\mu_2)g_2(x,y) = 0,$$

$$g_2(x,0) = \int\limits_0^\infty g_3(x,y)(\mu+\mu_2)\,dy, \qquad g_2(0,y) = \int\limits_0^\infty g_3(x,y)(\mu+\mu_1(x))\,dx$$

$$\frac{\partial g_3(x,y)}{\partial x} + \frac{\partial g_3(x,y)}{\partial y} + (2\mu+\mu_1(x)+\mu_2)g_3(x,y) = 0,$$

$$g_3(x,0) = \lambda \int\limits_x^\infty g_2(x,y)\,dy, \qquad g_3(0,y) = \lambda g_1(y) + \lambda \int\limits_y^\infty g_2(x,y)\,dx.$$

И появляется возможность выписать её решение. Точнее, в этом случае система сводится к системе линейных уравнений относительно неизвестных стационарных вероятностей $p_k, k = 0,1,2,3$ и чисел $A = \int_0^\infty g_2(0,y)\,dy$ и $B = \int_0^\infty g_3(0,y)\,dy$:

$$p_1 = \lambda p_0 \Phi_1^*(\beta_1),$$

$$p_2 = \lambda(\mu+\mu_2)A\Phi_1^{***}(\beta_1,\beta_2,\beta_1) + (\mu+\mu_2)B\Phi_1^{**}(\beta_1,\beta_2) + A\Phi_1^*(\beta_1),$$

$$p_3 = \lambda A\Phi_1^{**}(\beta_1,\beta_2) + B\Phi_1^*(\beta_2),$$

$$
\begin{aligned}
A &= \lambda A(f_1^{**}(\beta_1,\beta_2) + \mu\Phi_1^{**}(\beta_1,\beta_2)) + B(f_1^{*}(\beta_2) + \mu\Phi_1^{*}(\beta_2)),\\
B &= \lambda p_1 + \lambda^2(\mu+\mu_2)A\Phi_1^{***}(\beta_1,\beta_2,\beta_1) + \lambda(\mu+\mu_2)B\Phi_1^{**}(\beta_1,\beta_2),\\
\lambda p_0 &= \lambda p_0 f_1^{*}(\beta_1) + \lambda(\mu+\mu_2)f_1^{***}(\beta_1,\beta_2,\beta_1)A+\\
&+B(\mu+\mu_2)f_1^{**}(\beta_1,\beta_2) + Af_1^{*}(\beta_1) + (\mu+\mu_2)(p_1+p_2),\\
&p_0+p_1+p_2+p_3 = 1.
\end{aligned} \tag{1}
$$

Здесь $\Phi_1(x)$ – функция надёжности, $f_1(x)$ – плотность непрерывной случайной величины ω_1, $\beta_1 = \lambda+\mu+\mu_2$, $\beta_2 = 2\mu+\mu_2$, Φ_1^* и f_1^* – преобразования Лапласа функций $\Phi_1(x)$ и $f_1(x)$ соответственно, а преобразования F^{**} и F^{***} определяются так:

$$
F^{**}(s,t) := \int\limits_0^\infty e^{-sy}dy \int\limits_0^\infty e^{-tx}F(x+y)\,dx,
$$

$$
F^{***}(s,t,u) := \int\limits_0^\infty e^{-sz}dz \int\limits_0^\infty e^{-ty}dy \int\limits_0^\infty e^{-ux}F(x+y+z)\,dx.
$$

Полученная система линейных уравнений избыточна. Первые 6 уравнений линейно зависимы, поэтому при решении любое из них можно отбросить и использовать в дальнейшем для проверки результата.

Вероятность потери заявки в стационарном режиме равна, очевидно, p_3.

В случае, когда и $\mu_1(x) = \mu_1 = \text{const}$ стационарные вероятности p_0, p_1, p_2, p_3 легко определяются из так называемой системы уравнений равновесия (СУР) для марковского процесса $\xi(t)$, выписываемой из диаграммы «по стрелкам»:

$$
p_0 = \frac{\beta(\beta+\mu)}{\beta^2+(\lambda+\mu)\beta+\lambda\mu+\lambda^2},\quad p_1 = \frac{\lambda}{\lambda+\beta}p_0,\quad p_2 = \frac{\lambda^2}{\beta(\lambda+\beta)}p_0,\quad p_3 = \frac{\lambda^2}{\beta(\mu+\beta)}p_0.
$$

Здесь $\beta = \mu+\mu_1+\mu_2$. Это решение совпадает, как можно убедиться после некоторых выкладок, с решением системы (1).

Заключение

В данной работе рассмотрена система массового обслуживания типа M/G/2/0, т.е. система с двумя каналами обслуживания, пуассоновским потоком поступающих заявок, произвольной функцией распределения времени обслуживания, взаимопомощью между каналами, отказами (поломками) и восстановлением каналов обслуживания. Найдены интегро-дифференциальные уравнения и граничные условия, описывающие данную СМО. Авторам не удалось получить явное решение системы уравнений в общей постановке. В случае отсутствия поломок каналов поставленная задача решается. Удалось найти основные вероятностные характеристики системы в стационарном режиме, в частности, вероятность потери заявки, что в данной интерпретации задачи означает вероятность беспрепятственного пролетания самолета-бомбардировщика через зону, контролируемую ПВО. Под вопросом остается возможность нахождения явного решения для общего случая.

Список литературы

1. Овчаров Л.А. Прикладные задачи теории массового обслуживания / Л.А. Овчаров // М.: Машиностроение — 1969. — 324 с.
2. Анисимов В.В. Элементы теории массового обслуживания и асимптотического анализа систем / В.В. Анисимов, О.К. Закусило, В.С. Донченко // Киев: Выща школа — 1987. — 246 с.
3. Бочаров П.П. Теория массового обслуживания / П.П. Бочаров, А.В. Печинкин // М.: Изд-во РУДН — 1995. — 529 с.
4. Коваленко А.И. Исследование системы массового обслуживания M/G/1/1 / А.И. Коваленко, Б.Д. Марянин, В.П. Смолич // Ученые записки ТНУ — серия «Математика. Механика. Информатика и кибернетика». — 2002. — т.15(52) — №2. — С. 40-42.
5. Коваленко А.И. Исследование надёжности однолинейной системы с потерями требований / А.И. Коваленко, Б.Д. Марянин, В.П. Смолич // Таврический вестник информатики и математики. — 2003. — №2.
6. Коваленко А.И. Исследование системы M/D/1 с одной орбитой / А.И. Коваленко, Б.Д. Марянин, В.П. Смолич // Таврический вестник информатики и математики. — 2004. — №2.
7. Коваленко А.И. Исследование трёхэлементной СМО с отказами, обслуживаемой двумя наладчиками / А.И. Коваленко, Б.Д. Марянин, В.П. Смолич // Спектральные и эволюционные задачи. — Симферополь — 2007. — т.17. (КРОМШ)
8. Коваленко А.И. Стационарные характеристики системы с поочерёдным обслуживанием заявок двумя линиями / А.И. Коваленко, Б.Д. Марянин, В.П. Смолич // Динамические системы. — ТНУ, 2008. — Вып. 24 — с. 69-82.
9. Коваленко А.И. Исследование надежности однолинейной системы, обслуживающей два потока заявок, с конечной очередью / А.И. Коваленко, Б.Д. Марянин, В.П. Смолич // Таврический вестник информатики и математики. — 2009. — №2 — С. 63-70.

Статья поступила в редакцию 06.09.2011

УДК 519.7

ОПТИМИЗАЦИОННЫЕ МОДЕЛИ С ПРЕЦЕДЕНТНОЙ НАЧАЛЬНОЙ ИНФОРМАЦИЕЙ

© Анафиев А. С., Блыщик В. Ф.

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

Abstract. The main methods for solving optimization problems with incomplete (precedent) initial information are considered. Highlighted the problem of synthesis of optimization models with precedent initial information.

Введение

«*В экономике как неформализованной науке важна опора на прецеденты, на эмпирические закономерности. Поэтому в ней большое значение имеют методы обучения диагностике и выбору на основе опыта, на материале наблюдений.*» [7]

В общем случае задачу оптимизации можно сформулировать следующим образом

$$\text{extr}\, f(x) \,/\, x \in \Omega \subseteq X, \tag{1}$$

где $f : X \to Y$ — целевая функция, оптимальное (близкое к оптимальному) значение x^* которой требуется отыскать в ходе решения задачи, X — множество объектов, Y — множество значений (ответов) и Ω — множество (область) допустимых решений, из которого выбирается оптимальное решение x^*.

Согласно постановке (1) задачу оптимизации можно рассматривать как четверку $\langle X, Y, f, \Omega \rangle$.

Чаще всего множество допустимых решений задается в виде системы ограничений–неравенств. Неравенства и целевая функция могут быть линейными и/или нелинейными, в связи с чем выделяют классы задач линейного и нелинейного программирования, некоторые случаи из которых достаточно хорошо изучены, но при условии, что и *целевая функция и ограничения заданы полностью.* Однако в реальных практических задачах, в частности экономических, когда мы имеем дело с моделированием нестационарных и плохо определенных процессов, тяжело точно выписать все ограничения и целевую функцию. В следствии чего имеет место *неполнота начальных* *данных.* В этом случае трудна даже сама постановка (формализация) задачи, не говоря уже о выборе адекватной поставленной задаче модели и непосредственно поиска оптимального решения.

Если в оптимизационной задаче функция f и/или область допустимых решений Ω заданы не полностью, а имеется лишь некоторая частичная информация $I(f,\Omega)$ о них, то говорят о задаче оптимизации с *неполными (частичными) данными.*

1. Типы задач оптимизации с неполными данными

В работе В. И. Донского [1] выделяются следующие типы задач оптимизации в зависимости от начальной информации $I(f,\Omega)$ о целевой функции и ограничениях.

A1. Целевая функция f задана точно, а множество Ω задано частично перечислением точек из двух конечных множеств W_1 и W_2, для которых заведомо изветсно, что $W_1 \subset \Omega$ и $W_2 \subset X \setminus \Omega$; $I(f,\Omega) = \{f, W_1, W_2\}$.

A2. Дополнительно к A1 может быть задано множество $\Omega_\Delta \supset \Omega$ (часть ограничений, заданная явно); $I(f,\Omega) = \{f, W_1, W_2, \Omega_\Delta\}$.

B1. Задачи с заданной системой ограничений, но с частичной заданной целевой функцией. Информация о целевой функции представляет собой набор прецедентов $\left\{(x_i, f_i = f(x_i))_{i=1}^{\ell}\right\}$, который обозначим f_Δ; $I(f,\Omega) = \{f_\Delta, \Omega\}$.

B2. Задачи, отличающиеся от случая B1 тем, что информация о функции f задана частичным бинарным отношением $\rho_\Delta \subset \rho \stackrel{\Delta}{=} \left\{(x,x') : f(x) < f(x')\right\}$, иначе говоря, конечным набором ρ_Δ пар векторов $(x,x') \in \Omega \times \Omega$ таких, что имеет место $x\,\rho\,x'$; $I(f,\Omega) = \{\rho_\Delta, \Omega\}$.

B3. Задачи, отличающиеся от случая B2 тем, что существует эксперт, который для любых $x, x' \in \Omega$ дает ответ на вопрос о принадлежности пары (x,x') отношению ρ; $I(f,\Omega) = \{\rho, \Omega\}$.

Комбинируя различные варианты неполной начальной информации о целевой функции и ограничениях из задач A и B, можно сформулировать следующие типы задач оптимизации с частичными данными:

C11. $I(f,\Omega) = \{f_\Delta, W_1, W_2\}$.
C12. $I(f,\Omega) = \{\rho_\Delta, W_1, W_2\}$.
C13. $I(f,\Omega) = \{\rho, W_1, W_2\}$.
C21. $I(f,\Omega) = \{f_\Delta, W_1, W_2, \Omega_\Delta\}$.
C22. $I(f,\Omega) = \{\rho_\Delta, W_1, W_2, \Omega_\Delta\}$.
C23. $I(f,\Omega) = \{\rho, W_1, W_2, \Omega_\Delta\}$.

Задачи можно разделить на подклассы и в зависимости от некоторой дополнительной информации о множествах X и Y, о модели неизвестной целевой зависимости f и о форме области допустимых решений Ω:

D1. Вид пространства X и Y. Например, $X = \mathbb{R}^n$ и $Y = \mathbb{R}$, $X = \mathbb{B}^n \triangleq \{0,1\}^n$ и $Y = \mathbb{R}$.

D2. Предполагаемая (или явная для точно известных компонент) линейность задачи.

D3. Предполагаемая (или явная) выпуклость.

D4. Предположения о непрерывности, гладкости.

D5. Предположение о существовании вероятностной меры на X.

D6. В каком смысле понимается оптимальное решение.

2. Обзор методов решения задач оптимизации с неполными данными

Первые работы, в которых рассматриваются задачи математического программирования с неполной информацией, появились в начале 70-х годах прошлого века и принадлежат Владимиру Даниловичу Мазурову [5, 6]. В них описан подход основанный на методах теории распознавания образов в оптимальном планировании и управлении, который в дальнейшем развивался в более широком направлении синтеза итерационных методов оптимизации, порождаемых итерационными операторами с нестационарной системой параметров, и алгоритмов доопределения плохо формализуемых элементов моделей [4, 6].

Были разработаны итерационные процедуры основанные на использовании фейеровских отображений, представляющие собой реализацию методов недифференцируемой оптимизации в рамках предположения выпуклости.

Существенное место в развитии теории решения слабо формализованных задач занимает теория комитетных моделей оптимизации и классификации [4].

Определение. [4]. Комитетом системы линейных неравенств

$$\langle c_j, x\rangle = a_{j_1}x_1 + \ldots + a_{j_n}x_n \leqslant b_j, \qquad j = 1,\ldots,m,$$

где $x, c_j \in \mathbb{R}^n$, $b_j \in \mathbb{R}$, называется множество $K = \{x^1,\ldots,x^q\} \subset \mathbb{R}^n$ такое, что для любого j неравенство $\langle c_j, x^i\rangle \leqslant b_j$ выполняется более чем для половины элементов x^i множества K.

Рассмотрим пример противоречивой и неформализованной задачи $z = \langle \mathbb{R}^n, \mathbb{R}, f, I(\Omega)\rangle$, где информация $I(\Omega)$ представляет собой набор ограничений двух видов:

1) ограничения c_j, $j \in J$, модель и параметры модели которых известны;
2) ограничения γ_s, $s \in S$, модель и/или параметры модели которых неизвестны;

$$I(\Omega) = \begin{cases} a_j \leqslant c_j(x_1, \ldots, x_n) \leqslant b_j, & j \in J; \\ \alpha_s \leqslant \gamma_s(x_1, \ldots, x_n) \leqslant \beta_s, & s \in S. \end{cases} \tag{2}$$

Если система (2) совместна, γ_s — найдены, то имеем множество допустимых решений Ω и выбор оптимального решения осуществляется согласно функции f. Если же система несовместна, то ей ставится в соответствие комитетное решение, определяющее «размытое» решение — некоторое «облако» элементов, обладающее основными свойствами решения, сосредоточенного в одном элементе.

В общем случае система (2) может быть как несовместной, так и неформализованной. Тогда необходимо применить алгоритмы классификации для восстановления неизвестных ограничений γ_s, $s \in S$, а затем использовать понятие комитета системы (2) как конструкцию для неоднозначной интерпретации противоречивых данных. В этом и заключается основная идея метода комитета для решения плохо формализуемых и противоречивых задач оптимизации [8].

Отсюда можно сделать вывод, что основная направленность метода комитетов в задачах оптимизации — это преодоление проблемы противоречивости системы ограничений основанное на идеи консилиума — выработки результирующего коллективного решения на основе группы решающих правил, позволяющего ослабить «жесткие» противоречивые ограничения и осуществлять поиск решений для несобственных задач оптимизации.

Хочется отметить тот факт, что при неполноте начальных данных и восстановлении недостающих параметров модели важным является *согласованность* восстанавливаемой модели исходной задаче. Если не учитывать этого, а навязывать какую-ту «удобную» для решения модель и восстанавливать ее параметры (например, с помощью методов распознавания образов), то можно получить сколь угодно неточное решение исходной задачи.

В работе [1] детально изучаются задачи $\langle \mathbb{B}^n, \mathbb{R}, f, \Omega \rangle$ псевдобулевой оптимизации с частичной информацией; разработан алгоритм решения задач оптимизации для линейной целевой функции f, предложены методы решения для нелинейного случая и алгоритм направленного перебора на основе линеаризующей схемы ветвления для решения задач безусловной оптимизации псевдобулевой функции.

Синтетический метод [1] решения задачи оптимизации слабоопределенной линейной псевдобулевой функции с дизъюнктивным ограничением, основан на нахождении при помощи прецедентной информации системы образующих $\widetilde{C} = \{C_1, \ldots, C_q\}$ выпуклого многогранного конуса $K(\widetilde{C}) \subset \mathbb{R}^n$, которому *принадлежит неизвестный вектор* C_0 *коэффициентов линейной целевой функции.* Доказано, что при выполнении условий $C_0 \in K(\widetilde{C})$, $C_0 \neq 0$, $\forall C \in \operatorname{conv}(\widetilde{C})$ $\langle C, x^* \rangle = \max\limits_{x \in \Omega} \langle C, x \rangle$ выполняется

$$\langle C_0, x^* \rangle = \max \langle C_0, x \rangle \,/\, x \in \Omega,$$

где Ω — множество допустимых решений, определяемое дизъюнктивным ограничением. На основе представления Ω его покрытием $\Omega_1 \cup \ldots \cup \Omega_j \cup \ldots \cup \Omega_m$ таким, что $(x \in \Omega_j) \Leftrightarrow (K_j(x) = 1)$, $j = \overline{1, m}$, осуществляется решение m задач на выпуклой оболочке системы $\widetilde{C}$:

$$\max \langle \lambda_1 C_1 + \ldots + \lambda_q C_q, x \rangle \,/\, K_j(x) = 1; \quad \sum_{k=1}^{q} \lambda_k = 1; \quad \lambda_k \geqslant 0.$$

Использование дополнительной информации о знаках неизвестного вектора C_0 коэффициентов целевой функции позволяет получить простое правило вычисления булевых векторов $\alpha_j \in \Omega_j$ таких, что

$$\forall C \in \operatorname{conv}(\widetilde{C}) \quad \forall x \in \Omega \quad \langle C, \alpha_j \rangle \geqslant \langle C, x \rangle,$$

получить оценки

$$\forall C \in \operatorname{conv}(\widetilde{C}) \quad A_j \leqslant \max_{x \in \Omega_j} \langle C, x \rangle = \langle C, \alpha_j \rangle \leqslant B_j$$

и условие:

$$\exists A_{j^*} : A_{j^*} \geqslant B_j, \quad j^* \neq j, \quad j \in \overline{1, m}.$$

Показано, что при выполнении этого условия *исходная задача решается точно*:

$$\langle C_0, \alpha_{j^*} \rangle = \max \langle C_0, x \rangle \,/\, x \in \Omega;$$

в противном случае применяется теоретико-игровой подход к выбору решения, приводящий к решению матричной игры.

В случае, когда целевая функция известна точно, а ограничения заданы в виде набора прецедентов и известно, что они линейны, то при условии, что характеристическая функция принадлежности произвольного объекта $x \in X$ области допустимых решений Ω является монотонной, то применение корректного алгоритма распознавания обеспечивает существование точного решения исходной задачи.

Вывод

Несмотря на рассмотренные выше методы и подходы к решению задач оптимизации с неполными данными, на сегодняшний день проблема выбора адекватной оптимизационной модели и решения задач оптимизации с прецедентной начальной информацией является слабо изученной. Еще далеко не все методы и подходы к решению задач распознавания образов, восстановления регрессии, прогнозирования и кластеризации, в которых основной исходной информацией является набор прецедентов, применяются для решения слабоопределенных задач оптимизации с прецедентной начальной информацией о целевой функции и/или ограничениях. Задачи оптимизации подобного рода, конечно же, можно решать в два этапа: восстановление неизвестных элементов задачи и последующего решения полученной оптимизационной задачи. Но такой подход, на наш взгляд, не совсем корректен. Например, согласно [2, 3] имеются существенные различия в подходах к решению задач при неполноте информации о целевой функции по сравнению с задачами с точно заданной целевой функции и частично заданной области допустимых решений. Это лишний раз доказывает, что рассматривать задачу оптимизации с частичной информацией просто как объединение задач распознавания и оптимизации с полностью определенными данными не всегда уместно и может приводить к не адекватным решениям. Кроме того, некоторые алгоритмы машинного обучения необходимо адаптировать для решения задач оптимизации.

До сих пор не ясно, какие методы машинного обучения и как применять для решения конкретной задачи оптимизации с прецедентной начальной информацией для получения адекватных реальным практическим задачам моделей, для получения решений близким к истинным, а иногда и точно совпадающих с ними, для получения оценок надежности таких решений. Таким образом, актуальной является проблема синтеза алгоритмов машинного обучения (алгоритмов обучения по прецедентам) и методов оптимизации.

Используя особенности подходов, методов и алгоритмов машинного обучения, таких как метрические алгоритмы классификации, байесовский подход к классификации, алгоритмы кластеризации и различные методы восстановления регрессии, можно получать большое количество оптимизационных моделей с прецедентной начальной информацией, изучать их свойства и анализировать полученные решения. По мнению авторов статьи данное направление является одним из актуальных и перспективных в современной информатики и кибернетики.

Список литературы

1. Донской В. И. Дискретные модели принятия решений при неполной информации на основе синтетического подхода: дис. доктора физ.-мат. наук : 05.13.17 / Донской Владимир Иосифович. — Симферополь, 1993. — 267 с.
2. Донской В. И. Логическое управление плохо формализованными системами / Донской Владимир Иосифович // Динамические системы. — К.: Вища школа, 1985. — Вып. 4. — С. 90-96.
3. Донской В. И. Слабоопределенные задачи линейного булева программирования с частично заданным множеством допустимых решений / Донской Владимир Иосифович // Журнал вычисл. математики и матем. физики, 1988. — т. 28, № 9. — С. 1379-1385.
4. Мазуров В.Д. Метод комитетов в задачах оптимизации и классификации / Мазуров Владимир Данилович. — М.: Наука, 1990. — 248 с.
5. Мазуров В.Д. Об одном итерационном методе планирования, использующем распознавание образов для учета плохо формализуемых факторов / Мазуров Владимир Данилович // Изв. АН СССР. Техн. кибернетика. — 1973. — №3. — С. 205–207.
6. Мазуров В.Д. Применение методов теории распознавания образов в оптимальном планировании и управлении / Мазуров Владимир Данилович // Метод комитетов в распознавании образов. — Свердловск: УНЦ АН СССР, 1974. — С. 10–40.
7. Математические методы в экономике / [Еремин И. И., Мазуров Вл. Д., Скарин В. Д., Хачай М. Ю.]. — Под ред. Еремина И.И. и Мазурова Вл.Д. — Екатеринбург: Изд-во «У-Фактория», 2000. — 280 с.
8. Мазуров В. Д. Модели интерпретации противоречивых данных и метод комитетов / Мазуров В. Д. // Сборник научных трудов, Тр. ИММ УрО РАН. — 1992. — №1. — С. 193—203.

Статья поступила в редакцию 11.12.2011

УДК 519.6

НАБЛИЖЕННЯ РОЗРИВНИХ ФУНКЦІЙ ДВОХ ЗМІННИХ РОЗРИВНИМИ СПЛАЙН-ІНТЕРЛІНАНТАМИ З ВИКОРИСТАННЯМ ТРАПЕЦЕВИДНИХ ЕЛЕМЕНТІВ

© Литвин О. М., Першина Ю. І.

Українська інженерно-педагогічна академія
вул. Університетська, 16, м. Харків, 61003, Україна
e-mail: *academ@kharkov.ua*

Abstract. In work the general method of construction explosive a spline-interlination for rectangular element having the trapeze form for the purpose of their use for approach of explosive functions of two variables is offered, which too can have (and can and not have) ruptures of the first sort on the lines forming rectangular elements having the trapeze form. The constructed splines as a special case, explosive splines and continuous splines include. Are formulated and proved interlinational properties of such explosive designs.

Вступ

На даний час основна увага в теорії наближення функцій багатьох змінних сплайнами приділена наближенню неперервних і диференційовних функцій неперервними та диференційовними сплайнами [1-4]. В той же час практика показує, що серед багатовимірних об'єктів, які потрібно досліджувати, значно більша їх кількість описується розривними функціями. Наприклад, в методах комп'ютерної томографії на даний час не достатньо є дослідженим використання інформації про те, що внутрішня структура тіла людини складається з органів різної форми та різної щільності, тобто ми маємо розривну функцію. При дослідженні кори Землі за допомогою даних з кернів свердловинного буріння виникає задача відновлення внутрішньої структури між свердловинами. При цьому очевидним є той факт, що щільність ґрунту в різних точках кори є неоднорідною і найчастіше має розриви першого роду при переході від однієї складової кори до іншої (чорнозем, пісок, глина, граніт тощо. Тому актуальною є розробка методів наближення розривних функцій.

Весь розвиток обчислювальної та прикладної математики говорить про те, що використання кожної додаткової інформації про досліджуваний об'єкт може привести до більш точного і якісного відновлення цього об'єкту. Наприклад, в роботі [5] пропонується використовувати рівняння поверхні черепа людини і, таким чином, більш точно відновлювати внутрішню структур тіла.

В статті [6] авторами були побудовані розривні лінійні інтерполяційні сплайни для наближення функцій однієї змінної, що може мати розриви першого роду. В роботі [7] був запропонований метод наближення розривних функцій двох змінних з ректангульованою областю визначення розривними інтерполяційними білінійними сплайнами.

Розроблені методи в подальшому будуть використовуватися для розв'язання плоскої задачі радонівської комп'ютерної томографії. Для цього доцільніше використовувати оператори інтерлінації функцій, оскільки ці оператори відновлюють функції (можливо, наближено) за відомими їх слідами на даній системі ліній. Тобто, вони надають можливість будувати оператори, інтеграли від яких по вказаних лініях (лінійні інтеграли) будуть дорівнювати інтегралам від самої відновлюваної функції. Звідси витікає, що інтерлінація є математичним апаратом, природно пов'язаним із задачею відновлення характеристик об'єктів за відомими їх проекціями. В роботі [8] був запропонований метод наближення розривних функцій двох змінних з ректангульованою областю визначення розривними інтерлінаційними сплайнами. Були також побудовані розривні інтерлінаційні сплайни для наближення функцій двох змінних, область визначення яких розбивається на прямокутні трикутники [9].

В даній роботі будуються та досліджуються інтерлінаційні розривні сплайни для наближення розривних функцій з областю визначення, що розбивається на прямокутні трапеції та прямокутні трикутники.

Постановка задачі. Нехай задана розривна функція двох змінних $f(x,y)$ в області $D=[0,1]^2$. Будемо вважати, що область D розбивається прямими $x_0=0<x_1<x_2<...<x_m=1$, $y_0=0<y_1<y_2<...<y_n=1$ на прямокутні елементи, а кожний прямокутник розбивається похилою лінією на прямокутну трапецію та прямокутний трикутник. Трапеції та трикутники не вкладаються один в один, а їх сторони не перетинаються. Функція $f(x,y)$ має розриви першого роду на границях між цими трапеціями та трикутниками (не обов'язково між всіма).

Метою роботи є побудува та дослідження операторів розривної сплайн-інтерлінації таких, які в кожній трапеції є операторами сплайн-інтерлінації функції $f(x,y)$.

1. Метод побудови наближуючого розривного сплайн-інтерлінанта

Якщо (x_i,y_j) — вузол, в якому знаходиться прямий кут прямокутника, то може зустрітися чотири типи трапецій (рис. 1):

$$\mathrm{TP}_{ij}^{(1)}=\{x_i<x<x_{i+1},y_j<y<g_{j+1}^{(1)}(x)\},\mathrm{TP}_{ij}^{(2)}=\{x_{i-1}<x<x_i,y_j<y<g_{j+1}^{(2)}(x)\},$$
$$\mathrm{TP}_{ij}^{(3)}=\{x_i<x<x_{i+1},g_{j-1}^{(3)}(x)<y<y_j\},\mathrm{TP}_{ij}^{(4)}=\{x_i<x<x_{i+1},g_{j-1}^{(4)}(x)<y<y_j\},$$

$$\mathrm{TP}_{ij}^{(5)} = \{x_i < x < q_{i+1}^{(5)}(y), y_j < y < y_{j+1}\}, \mathrm{TP}_{ij}^{(6)} = \{q_{i-1}^{(6)}(y) < x < x_i, y_j < y < y_{j+1}\},$$
$$\mathrm{TP}_{ij}^{(7)} = \{q_{i-1}^{(7)}(y) < x < x_i, y_{j-1} < y < y_j\}, \mathrm{TP}_{ij}^{(8)} = \{x_i < x < q_{i+1}^{(8)}(y), y_{j-1} < y < y_j\},$$

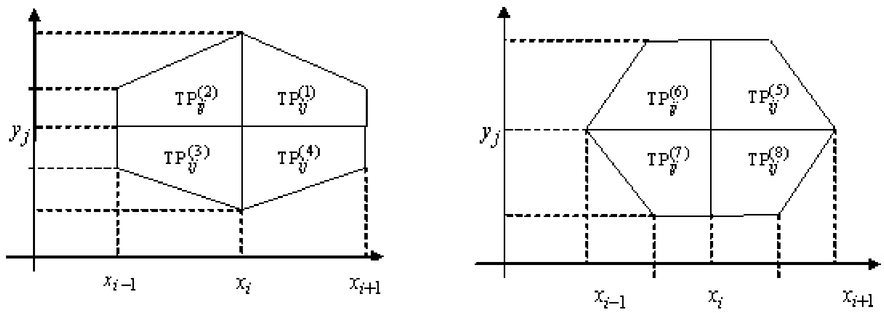


Рис. 1. Зображення можливих трапецевидних елементів з прямим кутом у вузлі (x_i, y_j)

Вважаємо, що на кожній із сторін заданих трапецій функція $f(x, y)$ може мати (а може і не мати) розриви першого роду. Розглянемо трапецію типу $\mathrm{TP}_{\mathrm{ij}}^{(1)}$.

Вважаємо заданими:

1. Сліди функції $f(x, y)$ на прямій $x = x_i$ (справа та зліва прямої відповідно):

$$\varphi p_i(y) = \lim_{x \to x_i + 0} f(x, y) = f(x_i + 0, y),$$

$$\varphi m_i(y) = \lim_{x \to x_i - 0} f(x, y) = f(x_i - 0, y).$$

Значення в кутових точках (x_i, y_j) та $(x_i, g_{j+1}^{(1)}(x_i))$ визначаються наступним чином:

$$\varphi pp_{ij} = \varphi p_i(y_j) = f(x_i + 0, y_j + 0),$$

$$\varphi pm_{i,j+1} = \varphi p_i(g_{j+1}^{(1)}(x_i)) = f(x_i + 0, g_{j+1}^{(1)}(x_i) - 0).$$

2. Сліди функції $f(x, y)$ на прямій $x = x_{i+1}$ (справа та зліва прямої відповідно):

$$\varphi p_{i+1}(y) = \lim_{x \to x_{i+1} + 0} f(x, y) = f(x_{i+1} + 0, y),$$

$$\varphi m_{i+1}(y) = \lim_{x \to x_{i+1} - 0} f(x, y) = f(x_{i+1} - 0, y).$$

Значення в кутових точках (x_{i+1}, y_j) та $(x_{i+1}, g_{j+1}^{(1)}(x_{i+1}))$ визначаються наступним чином:

$$\varphi mp_{i+1,j} = \varphi p_{i+1}(y_j) = f(x_{i+1} - 0, y_j + 0),$$

$$\varphi mm_{i+1,j+1} = \varphi m_{i+1}(g_{j+1}^{(1)}(x_{i+1})) = f(x_{i+1} - 0, g_{j+1}^{(1)}(x_{i+1}) - 0).$$

3. Сліди функції $f(x, y)$ на прямій $y = y_j$ (над та під прямою відповідно):

$$\psi p_j(x) = \lim_{y \to y_j + 0} f(x, y) = f(x, y_j + 0),$$

$$\psi m_j(x) = \lim_{y \to y_j - 0} f(x, y) = f(x, y_j - 0),$$

та значення у відповідних кутових точках:

$$\psi pp_{i,j} = \psi p_j(x_i) = f(x_i + 0, y_j + 0),$$

$$\psi mp_{i+1,j} = \psi p_j(x_{i+1}) = f(x_{i+1} - 0, y_j + 0).$$

4. Сліди функції $f(x, y)$ на прямій $y = g_{j+1}^{(1)}(x)$ (під та над прямою відповідно):

$$\psi m_{j+1}(x) = f(x, g_{j+1}^{(1)}(x) - 0),$$

$$\psi p_{j+1}(x) = f(x, g_{j+1}^{(1)} + 0)$$

та значення у відповідних кутових точках:

$$\psi pm_{i,j+1} = \psi m_{j+1}(x_i) = f(x_i + 0, g_{j+1}^{(1)}(x_i) - 0),$$

$$\psi mm_{i+1,j+1} = \psi m_{j+1}(x_{i+1}) = f(x_{i+1} - 0, g_{j+1}^{(1)}(x_{i+1}) - 0).$$

Означення. Будемо називати розривним інтерлінаційним поліноміальним сплайном в трапецевидному елементі $\mathrm{TP}_{ij}^{(1)}$ наступну функцію:

$$Lf(x, y) = (L_1 + L_2 - L_2 L_1) f(x, y), \tag{1}$$

де

$$L_1 f(x, y) = \frac{x - x_{i+1}}{x_i - x_{i+1}} \varphi p_i(y) + \frac{x - x_i}{x_{i+1} - x_i} \varphi m_{i+1}(y),$$

$$L_2 f(x,y) = \frac{y-y_j}{g^{(1)}_{j+1}(x)-y_j}\psi m_{j+1}(x) + \frac{y-g^{(1)}_{j+1}(x)}{y_j-g^{(1)}_{j+1}(x)}\psi p_j(x).$$

Теорема 1. Якщо сліди функції $f(x,y)$ задовольняють співвідношенням:

$$\varphi pp_{ij} = \psi pp_{ij}, \varphi mp_{i+1,j} = \psi mp_{i+1,j},$$

$$\varphi pm_{i,j+1} = \psi pm_{i,j+1}, \varphi mm_{i+1,j+1} = \psi mm_{i+1,j+1},$$

то оператор (1) інтерлінує $f(x,y)$ на $\partial \mathrm{TP}^{(1)}_{ij}$: $Lf(x,y)|_{\partial \mathrm{TP}^{(1)}_{ij}} = f(x,y)|_{\partial \mathrm{TP}^{(1)}_{ij}}$, тобто

$$Lf(x_i,y) = \varphi p_i(y), Lf(x_{i+1},y) = \varphi m_{i+1}(y), \tag{2}$$

$$Lf(x,y_j) = \psi p_j(x), Lf(x,g^{(1)}_{j+1}(x)) = \psi m_{j+1}(x). \tag{3}$$

Доведення. Щоб перевірити виконання цих умов, знайдемо

$$\begin{aligned} L_2 L_1 f(x,y) &= L_2\left(\frac{x-x_{i+1}}{x_i-x_{i+1}}\varphi p_i(y) + \frac{x-x_i}{x_{i+1}-x_i}\varphi m_{i+1}(y)\right) = \\ &= \frac{y-y_j}{g^{(1)}_{j+1}(x)-y_j}\left[\frac{x-x_{i+1}}{x_i-x_{i+1}}\varphi pm_{i+1,j} + \frac{x-x_i}{x_{i+1}-x_i}\varphi mm_{i+1,j+1}\right] + \\ &+ \frac{y-g^{(1)}_{j+1}(x)}{y_j-g^{(1)}_{j+1}(x)}\left[\frac{x-x_{i+1}}{x_i-x_{i+1}}\varphi pp_{i,j} + \frac{x-x_i}{x_{i+1}-x_i}\varphi mp_{i+1,j}\right]. \end{aligned} \tag{4}$$

Перевіримо виконання умов (2), (3):

$$Lf(x_i,y) = L_1 f(x_i,y) + L_2 f(x_i,y) - L_2 L_1 f(x_i,y) =$$

$$= \varphi p_i(y) + \frac{y-y_j}{g^{(1)}_{j+1}(x_i)-y_j}\psi m_{j+1}(x_i) + \frac{y-g^{(1)}_{j+1}(x_i)}{y_j-g^{(1)}_{j+1}(x_i)}\psi p_j(x_i) -$$

$$-\frac{y-y_j}{g^{(1)}_{j+1}(x_i)-y_j}\varphi pm_{i+1,j} - \frac{y-g^{(1)}_{j+1}(x_i)}{y_j-g^{(1)}_{j+1}(x_i)}\varphi pp_{i,j} = \varphi p_i(y),$$

оскільки з умови теореми маємо, що $\psi m_{j+1}(x_i) = \psi pm_{i,j+1} = \varphi pm_{i,j+1}$ та $\psi p_j(x_i) = \psi pp_{i,j} = \varphi pp_{i,j}$.

$$Lf(x_{i+1},y) = L_1 f(x_{i+1},y) + L_2 f(x_{i+1},y) - L_2 L_1 f(x_{i+1},y) =$$

$$= \varphi m_{i+1}(y) + \frac{y-y_j}{g^{(1)}_{j+1}(x_{i+1})-y_j}\psi m_{j+1}(x_{i+1}) + \frac{y-g^{(1)}_{j+1}(x_{j+1})}{y_j-g^{(1)}_{j+1}(x_{i+1})}\psi p_j(x_{i+1}) -$$

$$-\frac{y-y_j}{g_{j+1}^{(1)}(x_{i+1})-y_j}\varphi mm_{i+1,j+1}-\frac{y-g_{j+1}^{(1)}(x_{i+1})}{y_j-g_{j+1}^{(1)}(x_{i+1})}\varphi mp_{i+1,j}=\varphi m_{i+1}(y),$$

оскільки з умови теореми маємо, що $\psi m_{j+1}(x_{i+1}) = \psi mm_{i+1,j+1} = \varphi mm_{i,j+1}$ та $\psi p_j(x_{i+1}) = \psi mp_{i+1,j} = \varphi mp_{i+1,j}$.

$$Lf(x,y_j)=L_1f(x,y_j)+L_2f(x,y_j)-L_2L_1f(x,y_j)=$$
$$=\frac{x-x_{i+1}}{x_i-x_{i+1}}\varphi p_i(y_j)+\frac{x-x_i}{x_{i+1}-x_i}\varphi m_{i+1}(y_j)+$$
$$+\psi p_j(x)-\frac{x-x_{i+1}}{x_i-x_{i+1}}\varphi pp_{i,j}-\frac{x-x_i}{x_{i+1}-x_i}\varphi mp_{i+1,j}=\psi p_j(x).$$
$$Lf(x,g_j((1))(x))=L_1f(x,g_j^{(1)}(x))+L_2f(x,g_j^{(1)}(x))-L_2L_1f(x,g_j^{(1)}(x))=$$
$$=\frac{x-x_{i+1}}{x_i-x_{i+1}}\varphi p_i(g_{j+1}^{(1)}(x))+\frac{x-x_i}{x_{i+1}-x_i}\varphi m_{i+1}(g_{j+1}^{(1)}(x))+$$
$$+\psi m_{j+1}(x)-\frac{x-x_{i+1}}{x_i-x_{i+1}}\varphi pm_{i+1,j}-\frac{x-x_i}{x_{i+1}-x_i}\varphi mm_{i+1,j+1}]=\psi m_{j+1}(x).$$

Таким чином, перевірено виконання інтерлінаційних властивостей (2),(3) оператора (1).

Теорема 1 доведена. □

Зауваження. Перестановність операторів відсутня, тобто $L_1L_2 \neq L_2L_1$.

Для знаходження загального вигляду залишкового члена та його оцінки скористаємося результатами роботи [10], в якій представлений залишковий член для наближення неперервної функції трьох змінних оператором інтерфлетації на паралелепіпеді з криволінійною гранню.

Теорема 2. Нехай виконуються умови теореми 1, тоді для залишкового члена $Rf(x,y) = (I-L)f(x,y)$ виконується рівність

$$Rf(x,y)=\sum_{k=1}^{2}\sum_{m=1}^{2}P_{1,k}(x)P_{2,m}(x,y)\int\limits_{x_k}^{x}\int\limits_{y_m(x)}^{y}f^{(p,q)}(\xi,\eta)\frac{(x_k-\xi)^{p-1}(y_m-\eta)^{q-1}}{(p-1)!(q-1)!}d\xi d\eta, \quad (5)$$

$1 \le p,q \le 2$, $y_1(x) = y_j, y_2(x) = g_{j+1}^{(1)}(x)$, а поліноми $P_{1,k}(x)$, $P_{2,m}(x,y)$ мають вигляд

$$P_{1,1}(x)=\frac{x-x_{i+1}}{x_i-x_{i+1}}, P_{1,2}(x)=\frac{x-x_i}{x_{i+1}-x_i},$$

$$P_{2,1}(x,y) = \frac{y - y_j}{g^{(1)}_{j+1}(x) - y_j}, P_{2,2}(x,y) = \frac{y - g^{(1)}_{j+1}(x)}{y_j - g^{(1)}_{j+1}(x)}.$$

Доведення. Спочатку зауважимо, що оскільки одна із сторін трапеції $\mathrm{TP}^{(1)}_{\mathrm{ij}}$ задана функцією від змінної x, то $L_1L_2 \neq L_2L_1$.

Використаємо рівність

$$1 - (1 - l_1)(1 - l_2) = l_1 + l_2 - l_2l_2,$$

де l_1, l_2-деякі дійсні числа.

Підставимо замість 1 тотожній оператор I, а замість чисел l_1, l_2 оператори L_1, L_2 відповідно.

$$I - (I - L_1)(I - L_2) = L_1 + L_2 - L_2L_1 = L,$$

тобто отримали оператор $Lf(x,y)$.

Тоді для залишку $Rf(x,y)$ запишемо рівність

$$Rf(x,y) = (I - L)f(x,y) = (I - L_2)(I - L_1)f(x,y). \tag{6}$$

Для доведення формули (5) скористаємося тим фактом, що залишкові члени формули Лагранжа в інтегральній формі за кожною із змінних мають вигляд

$$(I - L_1)f(x,y) = \frac{x - x_{i+1}}{x_i - x_{i+1}} \int\limits_{x_i}^{x} \frac{\partial^p f(\xi,y)}{\partial \xi^p} \frac{(x_i - \xi)^{p-1}}{(p-1)!} d\xi +$$

$$+ \frac{x - x_i}{x_{i+1} - x_i} \int\limits_{x_{i+1}}^{x} \frac{\partial^p f(\xi,y)}{\partial \xi^p} \frac{(x_{i+1} - \xi)^{p-1}}{(p-1)!} d\xi,$$

$$(I - L_2)f(x,y) = \frac{y - g^{(1)}_{j+1}(x)}{y_j - g^{(1)}_{j+1}(x)} \int\limits_{y_j}^{y} \frac{\partial^q f(x,\eta)}{\partial \eta^q} \frac{(y_j - \eta)^{q-1}}{(q-1)!} d\eta +$$

$$+ \frac{x - y_j}{g^{(1)}_{j+1}(x) - y_j} \int\limits_{g^{(1)}_{j+1}(x)}^{y} \frac{\partial^q f(x,\eta)}{\partial \eta^q} \frac{(g^{(1)}_{j+1}(x) - \eta)^{q-1}}{(q-1)!} d\eta.$$

Підставимо ці рівності у формулу (6)

$$Rf(x,y) = (I - L_2)(I - L_1)f(x,y) =$$

$$= \frac{y - g_{j+1}^{(1)}(x)}{y_j - g_{j+1}^{(1)}(x)} (\int\limits_{y_j}^{y} \frac{\partial^q f(x,\eta)}{\partial \eta^q} \frac{(y_j - \eta)^{q-1}}{(q-1)!} d\eta +$$

$$+ \frac{x - y_j}{g_{j+1}^{(1)}(x) - y_j} \int\limits_{g_{j+1}^{(1)}(x)}^{y} \frac{\partial^q f(x,\eta)}{\partial \eta^q} \frac{(g_{j+1}^{(1)}(x) - \eta)^{q-1}}{(q-1)!} d\eta) \times$$

$$\times \frac{x - x_{i+1}}{x_i - x_{i+1}} (\int\limits_{x_i}^{x} \frac{\partial^p f(\xi,y)}{\partial \xi^p} \frac{(x_i - \xi)^{p-1}}{(p-1)!} d\xi + \frac{x - x_i}{x_{i+1} - x_i} \int\limits_{x_{i+1}}^{x} \frac{\partial^p f(\xi,y)}{\partial \xi^p} \frac{(x_{i+1} - \xi)^{p-1}}{(p-1)!} d\xi)$$

Після розкриття скобок отримаємо рівність (5).

Теорема 2 доведена. □

Оцінимо похибку наближення розривної функції $f(x,y)$ побудованим розривним інтерлінантом $Lf(x,y)$, визначеним формулою (1) в трапецевидному елементі $\mathrm{TP}_{\mathrm{ij}}^{(1)}$.

Теорема 3. Нехай $f(x,y) \in C^{p,q}(\mathrm{TP}_{\mathrm{ij}}^{(1)})$, $p = \overline{1,2}$, $q = \overline{1,2}$ та виконуються умови теореми 1, тоді для залишкового члена $Rf(x,y)$ має місце оцінка

$$\|Rf(x,y)\|_{C(\mathrm{TP}_{\mathrm{ij}}^{(1)})} \le M \int\limits_{x_i}^{x_{i+1}} \int\limits_{y_j}^{g_{j+1}^{(1)}(x)} |G_1(x,\xi) G_2(x,y,\eta)| d\xi d\eta, \qquad (7)$$

де

$$M = \max_{(x,y)\in \mathrm{TP}_{\mathrm{ij}}^{(1)}} |f^{(p,q)}(x,y)|,$$

$$G_1(x,\xi) = \begin{cases} \frac{x - x_{i+1}}{x_i - x_{i+1}} \frac{(x_i - \xi)^{p-1}}{(p-1)!}, & x_i \le \xi < x; \\ -\frac{x - x_i}{x_{i+1} - x_i} \frac{(x_{i+1} - \xi)^{p-1}}{(p-1)!}, & x \le \xi \le x_{i+1}, \end{cases}$$

$$G_2(x,y,\eta) = \begin{cases} \frac{y - g_{i+1}^{(1)}(x)}{y_j - g_{i+1}^{(1)}(x)} \frac{(y_j - \eta)^{q-1}}{(q-1)!}, & y_j \le \eta < y; \\ -\frac{y - y_j}{g_{i+1}^{(1)}(x) - y_j} \frac{(g_{i+1}^{(1)}(x) - \eta)^{q-1}}{(q-1)!}, & y \le \eta \le g_{i+1}^{(1)}(x). \end{cases}$$

Доведення. Згідно теореми 2, формулу для залишкового члена можна записати у вигляді

$$Rf(x,y) = \int\limits_{x_i}^{x_{i+1}} \int\limits_{y_j}^{g_{j+1}^{(1)}(x)} f^{(p,q)}(\xi,\eta) G_1(x,\xi) G_2(x,y,\eta) d\xi d\eta.$$

Застосовуючи для цього інтеграла нерівність Гельдера, одержуємо

$$|Rf(x,y)| \leq \ |f^{(p,q)}(x,y)\|_{L_\mu(\mathrm{TP}_{\mathrm{ij}}^{(1)})}[\int\limits_{x_i}^{x_{i+1}}\int\limits_{y_j}^{g_{j+1}^{(1)}(x)}(G_1(x,\xi)G_2(x,y,\eta))^\nu d\xi d\eta]^{\frac{1}{\nu}},$$

$$\mu \geq 1, \nu \geq 1, \frac{1}{\mu}+\frac{1}{\nu}=1.$$

Тому для похибки наближення отримаємо нерівність (7).

Теорема 3 доведена. □

Зауваження. Якщо односторонні сліди функції на відповідних лініях, що утворюють границі трапецевидних елементів, збігаються, то розривна функція перетворюється в неперервну.

2. Приклад

Нехай функція $f(x,y)$ задана в одиничному квадраті $[0,1]^2$ таким чином (рис. 2):

$$f(x,y)=\begin{cases} x+y, & 0.5<x<1, 0.5<y<0.5+\sqrt{0.09[1-\frac{(x-0.5)^2}{0.49}]};\\ x^2+y, & 0<x<0.5, 0.5<y<0.5+\sqrt{0.09[1-\frac{(x-0.5)^2}{0.49}]};\\ x+y^2, & 0<x<0.5, 0.5-\sqrt{0.09[1-\frac{(x-0.5)^2}{0.49}]}<y<0.5;\\ x^2+y^2, & 0.5<x<1, 0.5-\sqrt{0.09[1-\frac{(x-0.5)^2}{0.49}]}<y<0.5.\end{cases}$$

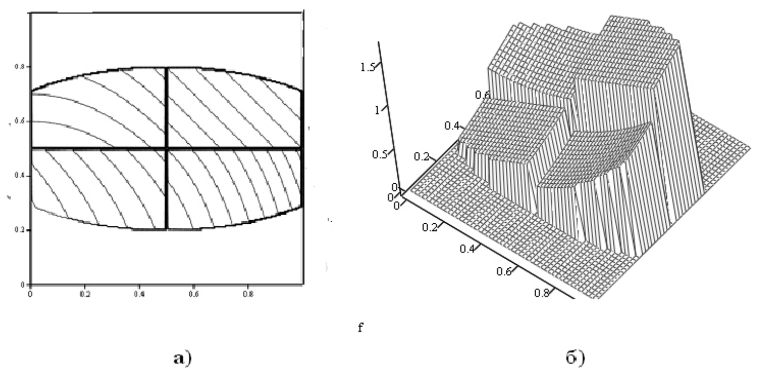


Рис. 2. Графічне зображення: а) області визначення функції $f(x,y)$; б) функції $f(x,y)$

Тобто на лінії еліпса $\frac{(x-0.5)^2}{0.49}+\frac{(y-0.5)^2}{0.09}=1$ функція $f(x,y)$ має розриви першого роду. Нехай задані лінії:

$$x_1=0, x_2=0.5, x_3=1,$$

$$y_1=0.5-\sqrt{0.09[1-\frac{(x-0.5)^2}{0.49}]},$$

$$y_2=0.5+\sqrt{0.09[1-\frac{(x-0.5)^2}{0.49}]}.$$

Вони розбивають область визначення функції $f(x,y)$ на вісім трапецевидних елементів з однією криволінійною стороною в кожному елементі.

Спочатку побудуємо розривний інтерполяційний сплайн на заданій сітці, який, наприклад, для трапеції $\mathrm{TP}_{\mathrm{ij}}^{(1)}$ задається формулою

$$S(x,y)=L_1L_2f(x,y)=\varphi pp_{ij}\frac{x-x_{i+1}}{x_i-x_{i+1}}\frac{y-g_{j+1}^{(1)}(x)}{y_j-g_{j+1}^{(1)}(x)}+\varphi mp_{i+1,j}\frac{x-x_i}{x_{i+1}-x_i}\frac{y-g_{j+1}^{(1)}(x)}{y_j-g_{j+1}^{(1)}(x)}+$$

$$+\varphi pm_{i,j+1}\frac{x-x_{i+1}}{x_i-x_{i+1}}\frac{y-y_j}{g_{j+1}^{(1)}(x)-y_j}+\varphi mm_{i+1,j+1}\frac{x-x_i}{x_{i+1}-x_i}\frac{y-y_j}{g_{j+1}^{(1)}(x)-y_j},$$

при умові, що виконуються рівності $\varphi pm_{i,j+1}=\psi pm_{i,j+1}$, $\varphi pp_{i,j}=\psi pp_{i,j}$, $\varphi mm_{i+1,j+1}=\psi mm_{i+1,j+1}$, $\varphi mp_{i+1,j}=\psi mp_{i+1,j}$.

Графічний вигляд такого інтерполяційного сплайна наведений на рис. 3.

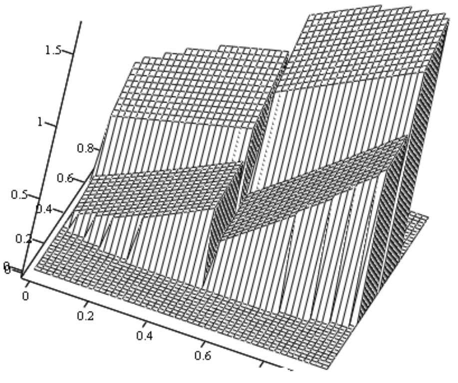


Рис. 3. Графічний вигляд розривного сплайн-інтерполянта для функції $f(x,y)$

Знайдемо оцінку похибки наближення розривної функції $f(x,y)$ побудованою розривною конструкцією $S(x,y)$

$$\max|f(x,y) - S(x,y)| \simeq 0.025$$

Тепер побудуємо на заданій сітці розривний інтерлінаційний сплайн $Lf(x,y)$ за формулою (1). Після перетворень можна побачити, що аналітично цей сплайн повністю збігається із заданою функцією $f(x,y)$, тобто $Lf(x,y) = f(x,y)$.

Можемо зробити висновок, що інтерлінаційний розривний сплайн точно відновлює задану розривну функцію на заданій сітці вузлів.

Висновки

Таким чином, в даній статті запропоновано загальний метод побудови розривних сплайн-інтерлінантів для трапецевидних елементів. Ці сплайни, як частинний випадок, включають в себе розривні сплайни та неперервні сплайни. Сформульовано і доведено теореми про інтерлінаційні властивості таких розривних конструкцій. Зокрема, з цих властивостей витікає наступна точка зору авторів: розривні в деяких точках або на деяких лініях функції від двох змінних краще наближувати розривними сплайн-інтерлінантами. При цьому можна отримати однаково високі оцінки похибки наближення в кожному елементі розбиття, притаманні неперервно-диференційовним сплайн-інтерлінантам.

Наступним кроком автори планують застосувати розроблену теорію наближення розривних функцій розривними сплайн-інтерлінантами до розв'язання двовимірної задачі комп'ютерної томографії.

Список літератури

1. Стечкин С.Б. Сплайны в вычислительной математике / С.Б. Стечкин, Ю.Н. Субботин. — М.: Наука, 1976.
2. Завьялов Ю.С. Методы сплайн-функций / Ю.С. Завьялов, Б.И. Квасов, В.Л. Мирошниченко — М.: Наука. 1976.
3. Василенко В.А. Сплайн-функции: теория, алгоритмы, программы / В.А. Василенко. — Новосибирск; Наука, 1983
4. Литвин О.М. Інтерлінація функцій та деякі її застосування / О.М. Литвин — Х.: Основа, 2002. — 504 с.
5. Литвин О.М. Про один метод розв'язання 3D задачі комп'ютерної томографії / О.М. Литвин, О.О. Литвин // Тезисы докладов Международной конференции АППММ'06. — Харків: ІПМАШ ім. А.М. Підгорного. — 2006. — С 18.

6. Литвин О.М. Наближення розривної функції за допомогою розривних сплайнів / О.М. Литвин, Ю.І. Першина // Математичне та комп'ютерне моделювання. Серія: Фізико-математичні науки: зб. наук. праць. — Кам'янець-Подільський: Кам'янець-Подільський національний університет ім. Івана Огієнка, 2010. — Вип. 3. — С. 122–131.
7. Литвин О.М. Побудова кусково-білінійних сплайнів для наближення функцій з розривами першого роду у вузлах ректангуляції двовимірної області / О.М. Литвин, Ю.І. Першина // Таврійский вісник інформатики та математики. — Симферополь. — 2011. — №1. — С. 63–72.
8. Литвин О.Н. Приближение разрывной функции двух переменных с помощью разрывных сплайнов двух переменных (прямоугольные элементы) / О.Н. Литвин, Ю.И. Першина // Компьютерная математика. — Киев, 2011. — №1. — С. 96–105.
9. Литвин О.М. Наближення розривних функцій двох змінних з розривами на лініях тріангуляції двовимірної області за допомогою операторів сплайн-інтерлінації / О.М. Литвин, Ю.І. Першина // Інформатика та системні науки (ІСН-2011): матеріали ІІ Всеукраїнської науково-практичної конференції 17-19 березня 2011 р. — Полтава: РВВ ПУЕТ, 2011. — С. 178–181.
10. Литвин О.М. Інтерфлетація функцій при розв'язуванні тривимірної задачі теплопровідності / О.М. Литвин, Л.І. Гулік. — К.: Наукова думка, 2011. — 210 с.

Статья поступила в редакцию 24.10.2011

УДК 519.713.1:51.681.3

О КОМПОНЕНТНОМ АНАЛИЗЕ ПАРАЛЛЕЛЬНЫХ РАСПРЕДЕЛЕННЫХ СИСТЕМ

© Лукьянова Е. А.

Киевский национальный университет
имени Тараса Шевченко, г. Киев.
e-mail: *lukyanovaea@mail.ru*

Abstract. The method of verification of parallel and distributed system by Component Petry Nets is offered. The using of Component Petry Nets for the design of parallel and distributed systems considerably diminishes model sizes and abbreviates time of verification. The presented methodology is studied on the example of verification of task of five philosophers.

Введение

Сети Петри являются удобным средством моделирования и анализа различных параллельных и распределенных систем. Однако при детальном моделировании реальных систем и объектов возникают проблемы, связанные с большими размерами получающихся сетей. Такие сети Петри могут содержать сотни, а иногда и тысячи элементов, что делает анализ детальных моделей практически неосуществимым. Для сокращения размеров модели и ее большей наглядности в работе [1] была введена в рассмотрение и предложена для моделирования параллельных и распределенных систем компонентная сеть Петри (CN-сеть).

Компонентная сеть Петри характеризуется следующими особенностями:

- конечные множества мест и переходов включают подмножества, состоящие из составных компонент (компонент-мест и компонент-переходов);
- функционирование составных компонент в CN-сети понимается как мгновенное выполнение.

Таким образом, места и переходы в CN-сети могут быть различных типов, что дает возможность двух аспектного подхода к функционированию CN-сети. С одной стороны, игнорирование внутренней работы составной компоненты, с другой стороны, рассмотрение отдельных представителей из групп одинаковых составных компонент в виде отдельной сети, находящейся некоторое время в активном состоянии. Такой подход к функционированию сети позволяет устанавливать структурные свойства модели следующим образом:

1) если исследуемое структурное свойство не выполняется на CN-сети, то это структурное свойство не выполняется и для детальной (базовой) модели исходной системы;

2) если исследуемое структурное свойство выполняется на CN-сети, то это структурное свойство выполняется для детальной модели системы, если оно выполняется на одном представителе из групп одинаковых составных компонент CN-сети.

Анализ CN-сетей предлагается проводить с помощью формальных методов, основанных на применении фундаментального уравнения и инвариантов. Использование CN-сетей в качестве средства моделирования систем позволяет эффективно применять эти методы и значительно уменьшить время верификации.

Целью данной работы является:

- подробное описание структуры составных компонент CN-сети, возможных условий, накладываемых на правила их построения (или выделения (нахождения) их в базовой модели) и описание работы составных компонент в CN-сети;
- демонстрация методологии использования CN-сети для верификации параллельных процессов на примере моделирования и анализа широко известной задачи о пяти философах.

Представленные в работе модели Петри задачи о пяти философах полностью учитывают все возможные особенности (опасные ситуации: ловушки, дедлоки) работы данной параллельной системы, и позволяют продемонстрировать особенности применения предложенной методологии в рамках данной статьи.

1. Компонентная сеть Петри

Построение компонентной сети Петри (CN-сети) для исследуемой параллельной системы, начинается с выявления возможных составных компонент (компонент-мест и компонент-переходов) в проектируемой детальной модели. Данное выявление начинается на начальном этапе моделирования еще при анализе исходной сложной системы. Результат такого анализа — выделение групп одинаковых или однотипных процессов. Это позволит на этапе построения модели заранее определить и неоднократно уточнить группы одинаковых или однотипных процессов и оформить их в виде блоков составных компонент модели. В результате получим модель, являющуюся детальной (подробной) моделью исходной системы, но в которой однотипные процессы заключены в соответствующие блоки — составные компоненты. При этом, рассматривая составные компоненты как места и переходы, получим компактную модель (CN-сеть) исследуемой системы.

Компонента-место C_p представляет собой участок сети, моделирующий некоторый однотипный процесс, начинающийся и заканчивающийся местом (местами), компонента-переход C_t — участок сети, моделирующий некоторый однотипный процесс, начинающийся и заканчивающийся переходом (переходами).

На адекватность модели предложенные преобразования не повлияют.

2. Формальное определение компонентной сети Петри

Компонентная сеть Петри (CN-сеть) — это ориентированный граф, описываемый множеством:

$$CN = (P, T, F, W, M_0),$$

где P — конечное множество мест, состоящее из подмножеств P_1 и P_2 (P_1 — конечное множество компонент-мест, P_2 — конечное множество мест, понимаемое в обычном смысле мест сетей Петри, оставшихся после выделения компонент-мест); T — конечное множество переходов, состоящее из подмножеств T_1 и T_2 (соответственно множество компонент-переходов, и множество переходов понимаемое в обычном смысле переходов сетей Петри, оставшихся после выделения компонент-переходов), $F \subseteq P \times T \cup T \times P$ — отношение инцидентности между местами и переходами, $W : F \to N \backslash \{0\}$ — функция кратности дуг, M_0 — начальная разметка сети.

Множества P и T удовлетворяют следующим условиям: $P \neq \emptyset$, $T \neq \emptyset$, $P \cap T = \emptyset$ (граф CN-сети должен содержать хотя бы один переход и одно место, причем вершина графа не может быть одновременно элементом множеств P и T).

Отношение инцидентности F и функция кратности дуг W определяют функцию инцидентности I , задающую правило: $I : ((P \times T) \cup (T \times P) \to N)$, и определяющую то, что элементы одного множества дугами соединены быть не могут, а также описывающую наборы входных и выходных элементов.

Компонентная сеть функционирует, переходя от разметки к разметке, как и регулярная сеть. Составные компоненты в CN-сети свои функции выполняют мгновенно: компонента-место C_p — мгновенное использование условия реализации события, компонента-переход C_t — мгновенная реализация события, приводящая к изменению разметки мест всех типов.

Исследование процесса функционирования составных компонент согласно условиям (1), (2) является неотъемлемой частью исследования свойств исследуемой системы на модели CN-сети.

Условия функционирования составной компоненты-места. Функционирование составной компоненты-места начинается после срабатывания ее входного перехода, в этот момент компонента C_p получает фишку. Это означает, что фишка помещается (фишки помещаются) в начальное место (во все начальные места) компоненты C_p. Выходной переход компоненты C_p сработает только тогда, когда компонента отработает и фишка переместиться (фишки переместятся) в последнее место (во все последние места) компоненты C_p. То есть, пока фишка находится в C_p, компонента-место работает от начального до финального своего состояния, при этом начальная и финальная разметки компоненты C_p не совпадут.

Условия функционирования составной компоненты-перехода. Функционирование составной компоненты-перехода начинается с запуска начального перехода (всех начальных переходов) компоненты C_t. Завершение работы компоненты-перехода — срабатывание последнего перехода (всех последних переходов) компоненты C_t. Начальная и финальная разметки компоненты C_t совпадут.

Отдельно от CN-сети составные компоненты не работают, реализация или выполнение происходят лишь после выполнения входного условия для C_t или после срабатывания входного перехода для C_p.

3. Моделирование задачи о пяти философах компонентной сетью Петри

Рассмотрим задачу о пяти философах. Пять размышляющих философов гуляют в саду. Если философ чувствует голод, он заходит в столовую, где стоит круглый стол с пятью стульями и блюдом спагетти посреди стола. На столе — пять вилок, по одной слева и справа от каждого стула. Философ берет вилки (обязательно нужно две вилки: в левую и в правую руки) и ест спагетти. Утолив голод, философ кладет вилки на стол и выходит в сад размышлять, пока вновь не проголодается. Таким образом, каждый из философов Φ_i $(i = 1, 2, 3, 4, 5)$ выполняет семь действий: 1 — Φ_i входит в столовую, 2 — Φ_i берет левую вилку, 3 — Φ_i берет правую вилку, 4 — Φ_i ест спагетти, 5 — Φ_i кладет левую вилку, 6 — Φ_i кладет правую вилку, 7 — Φ_i выходит из столовой.

Модель данной задачи в виде CN-сети (Рис. 1) учитывает все возможные варианты поведения философов и синхронизирует их независимые действия.

Регламентация использования вилок двумя соседними философами (обеспечение взаимного исключения) достигается в модели наличием мест p_{16}, p_{17} , p_{18}, p_{19} , p_{20}, которые разрешают срабатывание только одного из переходов t_{15} или t_6 , t_7 или t_8, t_9 или t_{10} , t_{11} или t_{12} , t_{13} или t_{14} соответственно.

Не допущение состояния вечного ожидания (ловушки), когда один из философов так и не сумеет получить доступ к ресурсу (вилке) и заговора соседей, когда обедают одни и те же, обеспечивается тем, что в модели прописаны действия, которые должен выполнить пообедавший философ прежде, чем он снова будет претендовать на вилку.

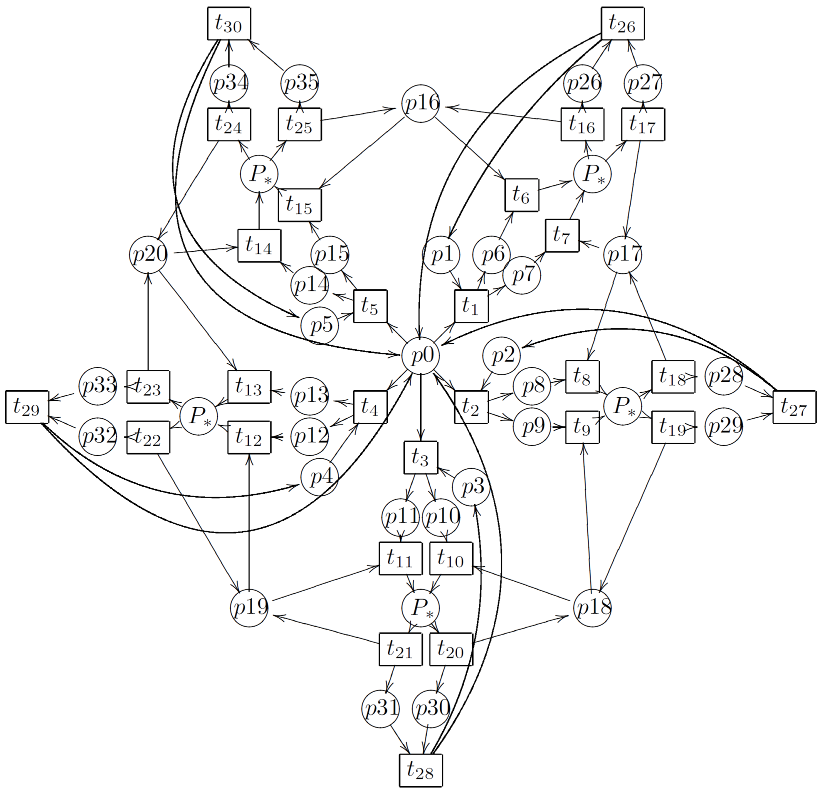


Рис. 1. CN-сеть, моделирующая задачу о пяти философах , где P_* — компонента-место, моделирующая один и тот же процесс

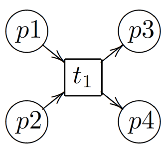


Рис. 2. Компонента-место P_* в CN-сети, моделирующей задачу о пяти философах

Например, процесс срабатывания переходов t_5 , t_{30} для Φ_5 дает возможность(время) кушать либо Φ_1 , либо Φ_4.

Проблема дедлока, когда все философы сидят за столом, каждый из них взял по одной вилке, и никто не может начать есть спагетти решается в модели за счет запрета на одновременное присутствие в столовой более, чем четырех философов, что обеспечивается местом p_0, количество фишек в котором равно четырем. Это место связано входными дугами с переходами t_i $(i = 1, 2, 3, 4, 5)$, соответствующими входу Φ_i $(i = 1, 2, 3, 4, 5)$ в столовую. Начало работы CN-сети обуславливается наличием фишек во входных местах p_i $(i = 1, 2, 3, 4, 5)$ (условие, что Φ_i проголодался) переходов t_i $(i = 1, 2, 3, 4, 5)$. Срабатывание этих переходов возможно до тех пор, пока в дополнительном месте имеется фишка — обед разрешен, в противном случае (если в столовой уже обедают четыре человека) философ будет ждать, пока один из обедающих не освободит ему место.

Предлагаемая CN-сеть имеет только один тип составных компонент — одну компоненту-место P_* (Рис. 2). Компонента-место P_* , моделирует следующее поведение отдельного i-го философа: Φ_i держит в руках левую и правую вилки, ест, заканчивает есть и готов положить левую, правую вилки.

4. Моделирование задачи о пяти философах компонентной сетью Петри с ингибиторными дугами

Синхронизацию независимых действий философов, при использование вилок двумя соседними философами, можно осуществить за счет использования дополнительных мест с ингибиторными дугами. Ингибиторная дуга была предложена в [2] для преодоления одного из ограничений классических сетей Петри — невозможности проверки места на отсутствие фишки. Ингибиторные дуги запрещают срабатывания перехода, если во входной позиции, связанной с переходом ингибиторной дугой, находится фишка. В предлагаемой модели ингибиторная дуга является индикатором свободы правого столового прибора: если прибор занят, то соседний философ ждет, пока он освободится. Этим в модели обеспечивается невозможность одновременного использования одной и той же вилки двумя обедающими философами.

Использование ингибиторной дуги, позволило при моделировании задачи выделить уже два типа составных компонент и значительно уменьшить модель. Модель задачи о пяти философах в виде CN-сети с ингибиторными дугами (Рис. 3) содержит следующие составные компоненты: компоненту-место P_* (Рис. 4) и компоненту-переход T_{**} (Рис. 5). Компонента-переход T_{**} моделирует процесс взятия Φ_i правой вилки, принятия еды, возвращение левой, правой вилки, выхода из столовой. Компонента- место P_* моделирует условие пребывания философа в столовой и взятия им левой вилки.

При формальном определении CN-сети с ингибиторными дугами необходимо отразить наличие множества обычных дуг и множества ингибиторных дуг. Поэтому для определения CN-сети с ингибиторными дугами в определении компонентной сети Петри во множестве F (отношении инцидентности между местами и переходами) выделим подмножество F_1 дуг (p_i, t_j), которые являются ингибиторными.

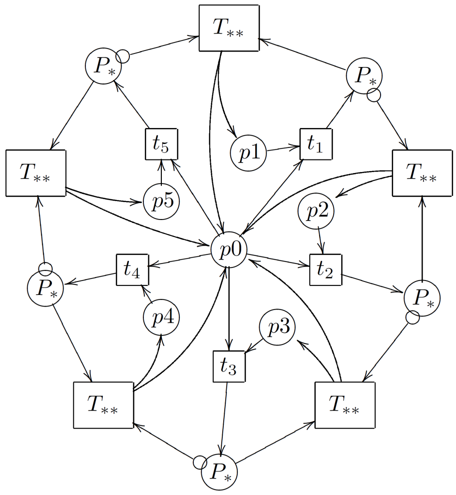


Рис. 3. CN-сеть с ингибиторными дугами, моделирующая задачу о пяти философах, где P_* — компонента-место, T_{**} — компонента-переход

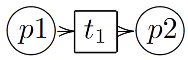


Рис. 4. Компонента-место P_* в CN-сети с ингибиторными дугами (Рис. 3)

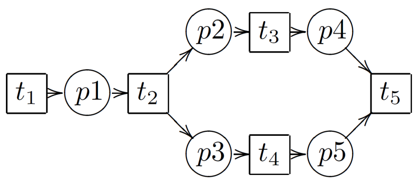


Рис. 5. Компонента-переход T_{**} в CN-сети с ингибиторными дугами (Рис. 3)

5. Анализ модели

В качестве методов исследования CN-сетей, используемых при верификации, выберем методы линейной алгебры [3, 4]. Система уравнений состояний (фундаментальное уравнение) CN-сети имеет вид

$$A \cdot x = d \qquad (1),$$

где $d = M_k - M_0$, (M_0, M_k — соответственно начальная и конечная разметки CN-сети), x — вектор счета срабатывания переходов, A — матрица инцидентности.

Если в качестве модели системы используется CN-сеть с ингибиторными дугами, то в матрице инцидентности такой сети факт применения ингибиторной дуги отразится следующим образом: соответствующая координата $(p_i, t_j,)$ в матрице инцидентности при ингибиторной дуге будет равна нулю, а не -1, так как соответствующая связь в модели существует, но движение фишки не осуществляет — переход срабатывает, не отнимая фишку у места.

Разрешимость фундаментального уравнения в неотрицательных целых числах является необходимым условием достижимости заданной разметки [3]. Решения системы (1) используются для построения искомых последовательностей срабатывания переходов. Для CN-сети строятся множества T- и S- инвариантов. Решение x системы линейных однородных диофантовых уравнений (СЛОДУ) $A \cdot x = 0$, когда в (1) $= M_0 = M_k$, есть T-инвариант CN-сети, а решение y СЛОДУ $A^T \cdot y = 0$ — S-инвариант CN-сети. S-инвариант — это линейное отношение на разметке подмножества мест, выражающееся в том, что взвешенная сумма различных фишек в местах является константой и равна значению, определяемому начальной разметкой. По S- и T-инвариантам устанавливаются такие структурные свойства CN-сети, как ограниченность, безопасность, повторяемость, непротиворечивость.

Теорема 1. *Если компонентная сеть Петри имеет только компоненты-переходы и они живы, то структурное свойство для детальной модели исследуемой системы выполняется, если это структурное свойство выполняется на CN-сети.*

Доказательство теоремы основывается на следующих фактах: живость компоненты-перехода обуславливается наличием у этой компоненты T-инвариантов, которые включают все переходы компоненты, что обеспечивает возможность срабатывания любого перехода и достижения любой разметки при функционировании компоненты-перехода. Это означает, что при запуске компоненты-перехода ее работа обеспечивает свободное движение фишки, а следовательно не мешает выполнимости свойств на CN-сети.

Теорема 2. *Если компонентная сеть Петри имеет только компоненты-места и соответствующие этим компонентам-местам системы линейных неоднородных диофантовых уравнений (СЛНДУ) совместны, то структурное свойство для детальной модели исследуемой системы выполняется, если это структурное свойство выполняется на CN-сети.*

Доказательство теоремы основывается на том, что из совместности соответствующей СЛНДУ следует существование в компоненте-месте возможности перехода из одного заданного состояния в другое, что обеспечивает сохранность свойств, которыми обладает CN-сеть.

Имеет место следующая теорема.

Теорема 3. *Если у компонентной сети Петри компоненты-переходы являются живыми, а соответствующие компонентам-местам СЛНДУ совместны, то структурное свойство для детальной модели исследуемой системы выполняется, если это структурное свойство выполняется на CN-сети.*

6. Анализ модели задачи о пяти философах

Выделение одной составной компоненты P_* в модели задачи о пяти философах позволило построить матрицу инцидентности, отвечающую данной модели, размерности не 56×35, а 36×30. Матрица инцидентности модели задачи о пяти философах CN-сетью с ингибиторными дугами имеет размерность 11×10 и приведена ниже.

$$\begin{matrix}
-1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 \\
-1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1
\end{matrix}$$

Поскольку разметка на k-м шаге M_k равна начальной разметке M_0 ($M_k = M_0 = (4,1,1,1,1,1,0,0,0,0,0)$) строим СЛОДУ $A \cdot x = 0$. Ее положительные целочисленные решения являются T-инвариантам CN-сети с ингибиторными дугами:

$$T_1 = (1,0,0,0,0,1,0,0,0,0), \;\; T_2 = (0,1,0,0,0,0,1,0,0,0),$$
$$T_3 = (0,0,1,0,0,0,0,1,0,0), \;\; T_4 = (0,0,0,1,0,0,0,0,1,0),$$
$$T_5 = (0,0,0,0,1,0,0,0,0,1).$$

Для поиска S-инвариантов строим СЛОДУ $A^T \cdot y = 0$. Ее решения — S-инварианты CN-сети с ингибиторными дугами:

$$S_1 = (0,1,0,0,0,0,1,0,0,0,0), \;\; S_2 = (0,0,1,0,0,0,0,1,0,0,0),$$

$$S_3 = (0,0,0,1,0,0,0,0,1,0,0), \quad S_4 = (0,0,0,0,1,0,0,0,0,1,0),$$

$$S_5 = (0,0,0,0,0,1,0,0,0,0,1), \quad S_6 = (1,0,0,0,0,0,1,1,1,1,1).$$

Матрицы инцидентности, соответствующие составным компонентам T_{**} и P_* в CN-сети с ингибиторными дугами, представлены ниже.

$$\begin{matrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{matrix}$$

$(-1,1)^T$

Для составной компоненты T_{**} финальная и начальная разметки равны ($M_k = M_0$=(0,0,0,0,0)), соответствующая СЛОДУ $A \cdot x = 0$ совместна и имеет одно решение (1,1,1,1,1) — T-инвариант составной компоненты T_{**} . Для составной компоненты P_* $M_k - M_0 = (1,-1)$, и соответствующая СЛНДУ имеет очень простой вид и единственное решение x=1.

Зная S-инварианты CN-сети с ингибиторными дугами и используя неравенство $M(p) \leq (M_0^T y)/y(p)$ ($y(p)$ означает p-ю координату вектора y) , дающее верхнюю оценку числа фишек, которые помещаются в место p_5 [5], можно определить максимальное количество фишек, помещающихся в места исследуемой CN-сети: $M(p_i) = 4$ $(i = 0,6,7,9,10)$, $M(p_j) = 1$ $(j = 1,2,3,4,5,)$.

Проведенный анализ CN-сети с ингибиторными дугами, позволяет установить свойства данной модели.

Так, из того, что все места CN-сети с ингибиторными дугами покрываются позитивными S-инвариантами следует, что CN-сеть с ингибиторными дугами является ограниченной. Более того, верхняя оценка числа фишек при начальной разметке $M_0 = (4,1,1,1,1,1,0,0,0,0,0)$ равна 4. Это свойство означает, что в столовой не может находиться более четырех человек.

Все переходы исследуемой CN-сети покрываются ненулевыми координатами из множества T-инвариантов следовательно CN-сеть с ингибиторными дугами обладает свойством повторяемости. Физическая интерпретация этого свойства означает, что любой из философов, проголодавшись, может утолить голод практически бесконечное число раз.

Исследуемая CN-сеть обладает свойством непротиворечивости, так как любая разметка М является достижимой из самой себя, что для задачи означает возможность всегда приходить в исходное состояние.

С учетом установленных фактов: компонента-переход жива, соотетствующая СЛНДУ для компоненты-места совместа, делаем вывод, что свойства, которыми обладает CN-сеть с ингибиторными дугами, сохраняются и для детальной (подробной) модели задачи о пяти философах.

Заключение

Предложенная в работе методология верификации параллельных и распределенных систем предполагает компонентное моделирование исследуемых систем CN-сетями и компонентный анализ свойств детальной модели, основанный на методах линейной алгебры. Компонентное моделирование дает возможность построить обозримую модель, удобную для анализа. Компонентный анализ позволяет значительно ускорить процесс проверки свойств на модели.

Список литературы

1. Лук'янова О. О. Про компонентне моделювання систем з паралелізмом / О. О. Лук'янова // Наукові записки НАУКМА, серія комп'ютерні науки. — прийнято до друку.
2. Котов В. Е. Сети Петри / В. Е. Котов — М.: Наука, 1984. — 157 с.
3. Murata T. Petri Nets: Properties, Analysis and Applications / Murata T. // in "Proceedings of the IEEE". — 1989. — vol. 77, № 4. — P. 541–580.
4. Крывый С. Л. О некоторых методах решения и критериях совместности линейных диофантовых уравнений в области натуральных чисел / С. Л. Крывый // Кибернетика и системный анализ. — 1999. — №4. — С. 12–36.
5. Крывый С. Л. О вычислении минимального множества инвариантов сетей Петри / С. Л. Крывый // Штучний інтелект. — 2001. — №3. — С. 199–206.

Статья поступила в редакцию 10.12.2011

Ибрагимов Н. С. Об одной задаче идентификации для одномерного нелинейного стационарного уравнения квазиоптики / Н. С. Ибрагимов // Таврический вестник информатики и математики. — 2011. — № 2. — С. 17-29.

УДК 517.97

В даній роботі розглядається задача ідентифікації для одномірного нелінійного стаціонарного рівняння квазіоптики про визначення коефіцієнтів рівняння в класі квадратично сумовних функцій, коли нелінійна частина рівняння містить суто уявний коефіцієнт. При цьому доведені теореми існування та єдиності розв'язку задачі ідентифікації. Крім того, доведена теорема існування та єдиності розв'язку відповідної прямої задачі.

В данной работе рассматривается задача идентификации для одномерного нелинейного стационарного уравнения квазиоптики об определении коэффициентов уравнения в классе квадратично суммируемых функций, когда нелинейная часть уравнения содержит чисто мнимый коэффициент. При этом доказаны теоремы существования и единственности решеня задачи идентификации. Кроме того, доказана теорема существования и единственности решения соответствующей прямой задачи.

Донской В. И. Эмпирическое обобщение и распознавание: классы задач, классы математических моделей и применимость теорий. Часть II / В. И. Донской // Таврический вестник информатики и математики. — 2011. — № 2. — С. 31-42.

УДК 519.95

Запропоновано класифікацію задач розпізнавання по їх основних властивостям. Обгрунтовується доцільність вибору таких методів рішення, які є узгодженими з особливостями класів задач.

Предложена классификация задач распознавания по их основным свойствам. Обосновывается целесообразность выбора методов решения, согласованных с особенностями классов задач.

Коваленко А. И. Исследование функционирования системы ПВО из двух ЗРК / А. И. Коваленко, Б. Д. Марянин, В. П. Смолич // Таврический вестник информатики и математики. — 2011. — № 2. — С. 43-49.

УДК 519.872

Аналізується СМО типа M/G/2/0 з обмеженим часом перебування заявки у системи. Одна заявка у системи обслуговується відразу у двох лініях. Під час з'явлення другой заявки лінії обслуговують заявки окремо. У роботі отримано ймовірнісні характеристики системи у стаціонарному режиму.

Рассматривается СМО типа M/G/2/0 с ограниченным временем пребывания заявки в системе. Одна заявка в системе обслуживается сразу двумя линиями. В случае появления ещё одной заявки линии обслуживают заявки раздельно. Найдены вероятностные характеристики системы в стационарном режиме.

Анафиев А. С. Оптимизационные модели с прецедентной начальной информацией / А. С. Анафиев, В. Ф. Блыщик // Таврический вестник информатики и математики. — 2011. — № 2. — С. 51-57.

УДК 519.7

У роботі розглянуті основні методи оптимізації з неповною (прецедентною) початковою інформацією. Виділена проблема синтезу моделей оптимізації з прецедентною початковою інформацією.

В работе рассмотрены основные методы оптимизации с неполной (прецедентной) начальной информацией. Выделена проблема синтеза моделей оптимизации с прецедентной начальной информацией.

Литвин О. М. Наближення розривних функцій двох змінних розривними сплайн-інтерлінантами з використанням трапецевидних елементів / О. М. Литвин, Ю. І. Першина // Таврический вестник информатики и математики. — 2011. — № 2. — С. 59-70.

УДК 519.6

В роботі запропоновано загальний метод побудови розривних сплайн-інтерлінантів для прямокутних трапецевидних елементів з метою використання їх для наближення розривних функцій від двох змінних, які теж можуть мати(а можуть і не мати) розриви першого роду на лініях, що утворюють прямокутні трапецевидні елементи. Побудовані сплайни, як частинний випадок, включають в себе розривні сплайни та неперервні сплайни. Сформульовано і доведено теореми про інтерлінаційні властивості таких розривних конструкцій.

В работе предложен общий метод построения разрывных сплайн-интерлинантов для прямоугольных трапецевидных элементов с целью использования их для приближения разрывных функций двух переменных, которые тоже могут иметь (а могут и не иметь) разрыввы первого рода на линиях, образующих прямоугольные трапецевидные элементы. Построенные сплайны, как частный случай, включают в себя разрывные сплайны и непрерывные сплайны. Сформулированы и доказаны интерлинационные свойства такиз разрывных конструкций.

Лукьянова Е. А. О компонентном анализе параллельных распределенных систем / Е. А. Лукьянова // Таврический вестник информатики и математики. — 2011. — № 2. — С. 71-81.

УДК 519.713.1:51.681.3

Представлена методологія верифікації паралельних і розподілених систем з допомогою компонентних мереж Петрі. Використання компонентних мереж Петрі для моделювання паралельних і розподілених систем значно зменшує розміри моделі і скорочує час верифікації. Представлена методологія вивчена на прикладі задачи про п'ять філософів.

Представлена методология верификации параллельных и распределенных систем с помощью компонентных сетей Петри. Использование компонентных сетей Петри для моделирования параллельных и распределенных систем значительно уменьшает размеры модели и сокращает время верификации. Представленная методология изучена на примере верификации задачи о пяти философах.

СПИСОК АВТОРОВ НОМЕРА

Анафиев Айдер Сератович	к. ф.-м. н., доцент кафедры информатики факультета математики и информатики Таврического национального университета им. В. И. Вернадского, ученый секретарь редакции журнала ТВИМ, Украина *e-mail: anafiyev@gmail.com*
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*
Донской Владимир Иосифович	д. ф.-м. н., профессор, заведующий кафедры информатики факультета математики и информатики Таврического национального университета им. В. И. Вернадского, главный редактор журнала ТВИМ *e-mail: donskoy@tnu.crimea.ua*
Ибрагимов Натик Сохраб оглы	к. ф.-м. н., доцент кафедры экономической информатики Бакинского государственного университета, Азербайджанская Республика *e-mail: ns.ibragimov@gmail.com*
Коваленко Александр Ильич	к. ф.-м. н., доцент кафедры математического анализа Таврического национального университета им. В. И. Вернадского, Украина
Литвин Олег Николаевич	д. ф.-м. н., профессор, заведующий кафедры высшей и прикладной математики Украинской инженерно-педагогической академии, Украина *e-mail: academ@kharkov.ua*
Лукьянова Елена Александровна	к. ф.-м. н., докторант Киевского национального университета им. Тараса Шевченко, Украина *e-mail: lukyanovaea@mail.ru*
Марянин Борис Давыдович	к. ф.-м. н., доцент кафедры математического анализа Таврического национального университета им. В. И. Вернадского, Украина

Першина Юлия Игоревна	к. ф.-м. н. докторант кафедры высшей и прикладной математики Украинской инженерно-педагогической академии, Украина *e-mail: yulia_pershina@mail.ru*
Смолич Владимир Павлович	к. ф.-м. н., доцент кафедры математического анализа Таврического национального университета им. В. И. Вернадского, Украина *e-mail: svp54@mail.ru*

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. Падати заявку на сайті журналу **www.tvim.info**.

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 орігіналу, відповідно до форми Ф23 бюлетеню ВАК України, 2008, № 3.
6. Стаття має бути підготовлена за допомогою видавничої системи LATEX з використанням стильового пакету twim.sty, який можно отримати за адресою **www.tvim.info**. Файли статті у форматі TeX і PDF (плюс графічні файли, якщо потрібні) необхідно прикріпити до заявки на публікацію статті на сайті журналу.

Робота редакції з авторами

1. Матеріали необхідно надіслати за допомогою сайта **www.tvim.info**, 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**.
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 22.12.2011. Формат 38x30/2. Бумага тип ОП. Объем 9,72 п.л. Тираж 500 экз. Заказ 335.
Издано в редакционном отделе КНЦ НАНУ
просп. Вернадского, 2, г. Симферополь, АРК, 95007, Украина