

**Институт проблем управления
им. В.А. Трапезникова РАН**

УПРАВЛЕНИЕ БОЛЬШИМИ СИСТЕМАМИ

Выпуск 39

Сентябрь 2012

**СБОРНИК
ТРУДОВ**

ISSN 1819-2467

Регистрационный номер Эл №ФС77-44158 от 09 марта 2011 г.

Москва – 2012

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Институт проблем управления
им. В.А. Трапезникова

**УПРАВЛЕНИЕ
БОЛЬШИМИ
СИСТЕМАМИ**

СБОРНИК ТРУДОВ

Выпуск 39

Москва – 2012

КООРДИНАЦИОННЫЙ СОВЕТ

Академики РАН: Васильев С.Н., Емельянов С.В., Коровин С.К., Куржанский А.Б., Федосов Е.А., Черноусько Ф.Л.; члены-корреспонденты РАН: Желтов С.Ю., Каляев И.А., Пархоменко П.П., Попков Ю.С.; д-ра техн. наук: Бутковский А.Г., Дорофеев А.А., Кузнецов О.П., Кульба В.В., Кротов В.Ф., Лотоцкий В.А., Павлов Б.В., Поляк Б.Т., Рутковский В.Ю.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Главный редактор: член-корр. РАН Новиков Д.А. **Отв. секретарь:** к.т.н. Губко М.В.
Д-ра техн. наук: проф. Алескерев Ф.Т. (ГУ ВШЭ), проф. Артамонов Е.И. (ИПУ РАН), д-р экон. наук, проф. Архипова М.Ю. (ИПИ РАН), д-ра техн. наук: проф. Афанасьев В.Н. (МИЭМ), проф. Бахтадзе Н.Н. (ИПУ РАН), проф. Бурков В.Н. (ИПУ РАН), проф. Вишневский В.М. (ИППИ РАН), д-р экон. наук, проф. Голиченко О.Г. (ЦЭМИ РАН), д-р физ.-мат. наук, проф. Добровидов А.В. (ИПУ РАН), д-ра техн. наук: проф. Заложнев А.Ю. (ИПУ РАН), проф. Ириков В.А. (МФТИ), проф. Калянов Г.Н. (ИПУ РАН), проф. Касаткин С.И. (ИПУ РАН), проф. Каравай М.Ф. (ИПУ РАН), канд. техн. наук Квинто Я.И. (ИПУ РАН), д-р экон. наук, проф. Клочков В.В. (ИПУ РАН), д-ра техн. наук: проф. Кононенко А.Ф. (ВЦ РАН), канд. техн. наук Куливец С.Г. (ИПУ РАН), проф. Курдюков А.П. (ИПУ РАН), проф. Лебедев В.Г. (ИПУ РАН), к-т техн. наук, доцент Лебедев В.Н. (ИПУ РАН), д-р экон. наук, проф. Ловчиновский Э.В. (ИПУ РАН), д-р техн. наук, проф. Мандель А.С. (ИПУ РАН), д-р экон. наук, проф. Нижегородцев Р.М. (ИПУ РАН), д-ра техн. наук: проф. Новосельцев В.Н. (ИПУ РАН), проф. Орлов А.И. (МВТУ), д-р физ.-мат. наук, проф. Рапопорт Л.Б. (ИПУ РАН), д-р техн. наук, проф. Рыков А.С. (МИСИС), д-р экон. наук, проф. Секерин В.Д. (ИПУ РАН), д-ра техн. наук: проф. Сидельников Ю.В. (МАИ), проф. Совлуков А.С. (ИПУ РАН), д-р экон. наук, проф. Сухарев О.С. (Ин-т экономики РАН), д-ра техн. наук: проф. Уткин В.А. (ИПУ РАН), проф. Хоботов Е.Н. (МВТУ), д-ра физ.-мат. наук: доцент Чеботарев П.Ю. (ИПУ РАН), проф. Чхартишвили А.Г. (ИПУ РАН), проф. Щербаков П.С. (ИПУ РАН).

РЕГИОНАЛЬНЫЕ РЕДАКЦИОННЫЕ СОВЕТЫ

Волгоград – д-ра физ.-мат. наук: проф. Воронин А.А., проф. Лосев А.Г. (ВолГУ); **Воронеж** – д-р техн. наук, проф. Баркалов С.А., д-р физ.-мат. наук, проф. Головинский П.А. (ВГАСУ), д-р техн. наук, проф. Подвальный С.Л. (ВГТУ); **Ижевск** – д-р физ.-мат. наук, проф. Непейвода Н.Н., к-т физ.-мат. наук, проф. Родионов В.И. (УдмГУ); **Иркутск** – д-ра физ.-мат. наук: проф. Бычков И.В., проф. Лакеев А.В. (ИДСТУ СО РАН); **Казань** – д-р физ.-мат. наук, проф. Маликов А.И., д-р техн. наук, проф. Сиразетдинов Р.Т. (КГТУ-КАИ); **Липецк** – д-ра техн. наук: проф. Кузнецов Л.А., проф. Погодаев А.К. (ЛГТУ); **Самара** – д-ра экон. наук: проф. Богатырев В.Д., проф. Гераскин М.И., д-р техн. наук, проф. Засканов В.Г. (СГАУ); **Санкт-Петербург** – д-ра физ.-мат. наук: проф. Петросян Л.А. (СПбГУ), проф. Фрадков А.Л. (ИПМ РАН); **Старый Оскол** – д-р техн. наук, проф. Еременко Ю.И. (СТИ); **Тверь** – д-ра техн. наук: проф. Кузнецов В.Н., проф. Палюх Б.В. (ТГТУ).

Адрес редакции: 117997, г. Москва, ул. Профсоюзная, д. 65.

Адрес в Интернет: ubs.mtas.ru.

Номер гос. регистрации электронного научного издания (ЭНИ): 0421200023.

© ИПУ РАН, 2012

СОДЕРЖАНИЕ

Системный анализ

Крянев А. В., Семенов С. С.

Особенности развития современной техники и метод оценки технического уровня сложных технических систем, основанных на использовании зарождающихся технологий 5

Математическая теория управления

Исмаилов И. Г.

Проекционно-итерационные процедуры приближенного построения вынужденных колебаний в нелинейных системах 37

Ломов А. А.

Вариационные методы идентификации линейных динамических систем и проблема локальных экстремумов 53

Чайковский М. М.

Синтез анизотропных субоптимальных регуляторов заданного порядка на основе полуопределенного программирования и алгоритма поиска взаимобратных матриц 95

Анализ и синтез систем управления

Гордеев А. А., Юркевич В. Д., Зиновьев Г. С.

Исследование системы управления двигателем постоянного тока с многоуровневым преобразователем напряжения 138

Усков А. А.

Устойчивость систем с блоками нечеткого логического вывода в объекте управления 155

Управление в социально-экономических системах

Воронин А. А., Харитонов М. А.

Модель численной оптимизации структуры операционного ядра организации 165

Клочков В. В., Гривский С. А., Игнатъева А. И.
Налоговые механизмы стимулирования повышения экологичности оборудования длительного пользования 184

Управление в медико-биологических и экологических системах

Веденяпин Д. А., Лосев А. Г.
Об одной нейросетевой модели диагностики венозных заболеваний 219

Управление техническими системами и технологическими процессами

Андрюенко А. Я., Тропова Е. И., Чадаев А. И.
Анализ возможностей повышения безопасности эксплуатации перспективных ракетных средств выведения на орбиты 230

Сочнев А. Н.
Распределение ресурсов производственной системы с использованием сетей Петри и генетического алгоритма 238

Цуканов М. А., Боева Л. М.
Мультиагентная система поддержки принятия решений по оперативному планированию и технологической координации сложноструктурированных производственных систем 254

Управление подвижными объектами и навигация

Парсегов С. Э.
Сцепление координат и иерархические алгоритмы в задаче равноудаленного расположения агентов на отрезке 264

Технические и программные средства управления

Тюгашев А. А., Ильин И. А., Ермаков И. Е.
Пути повышения надежности и качества программного обеспечения в космической отрасли 288

УДК 658.562.6; 629;7.01; 007

ББК 30.1

**ОСОБЕННОСТИ РАЗВИТИЯ СОВРЕМЕННОЙ
ТЕХНИКИ И МЕТОД ОЦЕНКИ ТЕХНИЧЕСКОГО
УРОВНЯ СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ,
ОСНОВАННЫХ НА ИСПОЛЬЗОВАНИИ
ЗАРОЖДАЮЩИХСЯ ТЕХНОЛОГИЙ**

Крянев А. В.¹,

*(Национальный исследовательский ядерный университет
«МИФИ», г. Москва)*

Семенов С. С.²

*(ОАО «Государственное научно-производственное пред-
приятие «Регион», г. Москва)*

Данная статья посвящена проблеме оценки технического уровня (ТУ) сложных технических систем (СТС), основанных на новых зарождающихся технологиях, при этом в статье подробно отражены особенности развития современной техники, обоснована актуальность и описаны существующие методы определения момента перехода на новые технологии, описан принцип формирования прогнозных значений основных характеристик новой технологии и представлен метод оценки ТУ СТС с зарождающимися технологиями.

Ключевые слова: технический уровень, коэффициент технического уровня, сложная техническая система, жизненный цикл, зарождающаяся новая технология, прогнозные значения, темп роста новой технологии, момент перехода на новую технологию.

¹ Александр Витальевич Крянев, доктор физико-математических наук, профессор кафедры «Прикладная математика», (avkryanev@terphi.ru).

² Сергей Сергеевич Семенов, кандидат технических наук, руководитель группы анализа и перспективных исследований (gnppregion@sovintel.ru).

1. Актуальность определения момента перехода на новые технологии

Для общества начала XXI в. характерно дальнейшее наращивание технологических возможностей во всех сферах человеческой деятельности. Если судить с позиций теории русского экономиста Н.Д. Кондратьева (1892–1931 гг.), то мир вступил в пятый цикл технологического уклада, для которого определяющими факторами является не только усложнение машин, различных аппаратов, но и интенсивное развитие новых технологий, опирающихся на последние достижения в области микроэлектроники, информатики, биотехнологии, геномной инженерии, новых видов энергии, освоение космического пространства, спутниковой связи и т.д. (таблица 1) [3, 11].

По данным научной школы в сфере экономики наукоемких технологий МИФИ, полученным за последнее десятилетие [8], были уточнены временные рамки четвертого и пятого укладов и введен шестой цикл, относящийся к периоду 2010–2050 гг., в котором преобладают нанотехнологии.

Современная продукция строится на новых высоких технологиях, в особенности информационных. В связи с тем, что значительная доля общественного труда будет связана с информационными технологиями, по оценке корпорации РЭНД (США) в ближайшие 10–20 лет произойдет углубление неравенства наций и, как следствие, – небывалый рост напряженности во всем мире [4]. Поэтому считается, что нация, которая освоит новые информационные технологии, бесспорно получит преимущества в своем развитии.

Следует также отметить, что по мере прогресса в области науки и техники заметно сократился период между научными исследованиями и созданием технических средств [24]. Этот период исчисляется для фотографии в 102 года, для телефона – 56 лет (1820–1876), для радио – 28 лет (1867–1895), для телевидения 14 лет (1922–1936), для беспилотного летательного аппарата³ – 6 лет (1903–1909), для радара – 14 лет (1926–1940), для

³ Файвуш Я., Аррисон В. Самолет без летчика и управления им по радио. – М.: Авиакхим, 1925. – 45 с.

транзистора – 5 лет (1948–1953), для лазера – 5 лет (1956–1961), рис. 1.

В таблице 2 представлен период создания некоторых новых технологий и машин [1]. Необходимо заметить, что приведенные данные по времени создания ЭВМ на современной технологической базе в 21 год основываются на приведенном источнике [1], всего же процесс создания ЭВМ занял несколько столетий и начался с первых счетных машин Леонардо да Винчи (1517) и Шиккарта (1622), а первым программистом считается графиня Лавлейс (1815–1852) – дочь лорда Байрона.

Как считал Э. Янч, консультант Организации экономического сотрудничества и развития в 1970-е годы, общий промежуток времени от открытия до нововведения при эффективном проведении всего цикла исследований и разработок в общем составляет 15 лет.

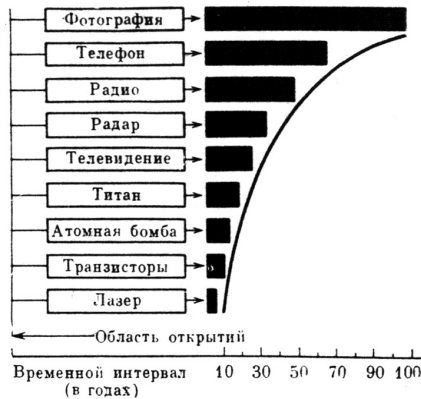


Рис.1. Временной интервал между датой открытия и практическим его использованием

Цикл «жизни» машин (рис. 2) включает ряд фаз, каждая из которых отличается других степенью научного и технического совершенства изделия, организационными особенностями, экономическими характеристиками, уровнем спроса [22].

Таблица 1. Технологические циклы научно-технического преобразования

Циклы	Страны – лидеры технического прогресса	Общая характеристика	Ядро технологического цикла	Ключевые факторы	Ядро нового цикла
Третий цикл (1880–1940 гг.)*	Германия, США, Великобритания, Швейцария, Нидерланды, Италия, Дания, Австро-Венгрия, Канада, Япония, Испания, Россия, Швеция	Активное использование в промышленном производстве электроэнергии, развитие тяжелого машиностроения и электротехнической промышленности на базе стального проката, новых открытий в области химии. Развитие химической промышленности. Нефтяной бум в США, формирование военно-промышленного комплекса в Европе, изобретение и внедрение телеграфа и радио. Производство автомо-	Электротехническая промышленность, тяжелое машиностроение, неорганическая химия	Электродвигатель; металлы (в том числе сталь)	Автомобилестроение, органическая химия, добыча и переработка нефти, цветная металлургия, автомобильное строительство

Циклы	Страны – лидеры технического прогресса	Общая характеристика	Ядро технологического цикла	Ключевые факторы	Ядро нового цикла
		билей и самолетов, добыча цветных металлов (в том числе алюминия), разработка пластмасс и полимеров, производство товаров длительного пользования, средств радио- и телекоммуникаций. Открытие фирм, картелей, трестов, корпораций, монополий, олигополий. Государство либо осуществляет контроль за ними, либо цели ком ими владеет. Концентрация финансового и банковского капитала			

Циклы	Страны – лидеры технического прогресса	Общая характеристика	Ядро технологического цикла	Ключевые факторы	Ядро нового цикла
Четвертый цикл (1930–1990 гг.)	США, страны Западной Европы, Канада, Австралия, Япония, Швеция, Швейцария, СССР, Бразилия, Мексика, Китай, Тайвань, Индия	Развитие энергетики на основе нефти и нефтепродуктов, развитие средств связи. Развитие автомобилестроения. Разработка новых видов вооружения и военной техники. Выпуск товаров длительного пользования. Строительство скоростных автомагистралей и аэропортов. Создание компьютера и разработка программных продуктов. Разработка и активное использование продукции электронной промышленности	Автомобилестроение и тракторостроение, цветная металлургия, производство товаров длительного пользования. Синтетические материалы, органическая химия, добыча и переработка нефти	Двигатель внутреннего сгорания	Микроэлектронные компоненты

Циклы	Страны – лидеры технического прогресса	Общая характеристика	Ядро технологического цикла	Ключевые факторы	Ядро нового цикла
		ности. Развитие атомной промышленности. Массовое производство на основе конвейерной технологии. Господство олигополической конкуренции, формирование транснациональных корпораций, осуществляющих прямые инвестиции в экономику различных стран. Перенос массового производства в страны «третьего мира»			
Пятый цикл (1980–2040 гг.)	Япония, США, Германия, Швеция, страны Западной Европы, Тайвань, Южная	Бурное развитие микроэлектроники, информатики, биотехнологий, генной инженерии. Получение	Электронная промышленность, оптоволоконная техника, программное обеспечение,	Микроэлектронные компоненты	Биотехнологии, космическая техника, тонкая химия

Циклы	Страны – лидеры технического прогресса	Общая характеристика	Ядро технологического цикла	Ключевые факторы	Ядро нового цикла
	Корея, Австралия, Бразилия, Мексика, Аргентина, Венесуэла, Китай, Индия, Турция, Индонезия, страны Восточной Европы, Россия (?)	энергии из нетрадиционных источников, освоение космического пространства. Распространение спутниковой связи. Формирование единой сети крупных и средних фирм, осуществляющих тесное взаимодействие в области технологий, контроля качества продукции, планирования инвестиций и т.д., логистика	телекоммуникации, роботостроение, добыча и переработка газа, информатика		
* Первый (1785–1835 гг.) и второй (1830–1890 гг.) циклы не показаны.					

Таблица 2. Темпы создания научно-технических новшеств

Новшество	Год начала промышленных разработок	Продолжительность (в годах)		
		Инкубационного периода	Промышленных разработок	Всего процесса создания
Автомобильный транспорт	1891	23	4	27
Авиационный транспорт	1903	6	8	14
Радиовещание	1913	17	9	26
Электронные лампы	1914	7	6	13
Телевидение	1933	22	12	34
ЭВМ	1944	15	6	21
Атомная электростанция	1954	11	3	14
Интегральные схемы	1958	2	3	5

Научный прогноз (фаза $0 - t_1$) дает толчок к активизации НИР и ОКР (фаза $t_1 - t_2$) в заданном направлении, результатом является постепенное улучшение характеристик будущей машины, которая после прохождения определенного комплекса испытаний может быть передана в производство (фаза $t_2 - t_3$). Отработка изделия продолжается и на этой фазе, но уже более медленными темпами. Наконец, конструкция и технологические процессы достигают степени совершенства при данном уровне развития науки и техники (кривая 1 на рис. 3). Одновременно увеличивается спрос на новую продукцию с лучшими эксплуатационными показателями (кривая 2), следовательно, растут прибыли изготовителя и потребителя (кривая 3) по фазам $(t_3 - t_4)$ и $(t_4 - t_5)$. Фаза $(t_5 - t_6)$ характеризует насыщение рынка сбыта и возникновение симптомов морального старения конструкции, вызванные бурным научно-техническим прогрессом. Изготовитель вынужден

изыскивать способы продления жизни изделия, производство которого хорошо отлажено. Так возникает потребность в модернизации конструкции ($t_6 - t_7$), которая на некоторое время вновь повышает заинтересованность потребителей в приобретении данной продукции. Но наука неуклонно движется вперед, появляются изделия, основанные на принципиально новом конструктивно-технологическом подходе, а выпуск старого изделия постепенно свертывают, и наступает фаза ($t_7 - t_8$) его моральной гибели. **На основе изучения информационных источников и обработки накопленного статистического материала существующими методами прогнозирования возможно определять временные интервалы развития конструкций и процессов разного назначения и устанавливать момент перехода к новым принципам создания изделий.**

Поскольку наука развивается опережающими темпами по отношению к производству, то возможности производства, как правило, оказываются ограниченными и не позволяют использовать наиболее перспективные технические решения сразу после их появления (таблица 3) [1].

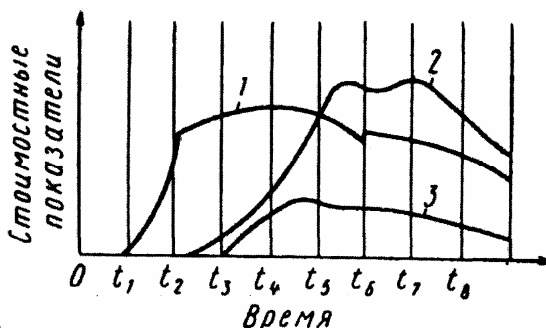


Рис. 2. Фазы цикла жизни изделия: кривая 1 – затраты на разработку, испытания и серийное производство; кривая 2 – цена реализации продукции; кривая 3 – прибыли изготовителя и потребителя; фазы (0 – t_1) – научный прогноз; ($t_1 - t_2$) – НИР и ОКР; ($t_2 - t_3$) – внедрение в производство; ($t_3 - t_4$) – рост; ($t_4 - t_5$) – зрелость; ($t_5 - t_6$) – насыщение; ($t_6 - t_7$) – частичная модернизация; ($t_7 - t_8$) – прекращение использования

Таблица 3. Оценка продолжительности отдельных этапов жизненного цикла некоторых видов техники (в годах)

Наименование объекта	Выработка проектных и условных основных спецификаций	Конструирование, изготовление, испытание опытного образца	Рабочие чертежи на массовое производство и наладку оборудования	Время производства до выпуска первых промышленных образцов	Общая продолжительность
Танко-десантный корабль	0,5	0,7	0,9	0,8	2,9
Минный тральщик	0,8	0,8	0,8	0,8	3,2
Безоткатное орудие	0,3	1,7	1,0	1,3	4,3
Средний танк	0,13	2,3	0,5	1,1	4,2
Эсминец	0,8	2,1	1,3	1,0	5,2
Транспортный самолет	0,5	2,4	0,5	1,8	5,2
Бомбардировщик	0,5	3,0	0,8	1,8	6,1
Реактивный истребитель	0,8	3,3	0,7	2,4	7,2
Автоматическая телефонная станция	–	–	–	–	6,0
Спутник связи	–	–	–	–	5,0
Реактивный двигатель	–	–	–	–	14,0
Радар	–	–	–	–	10–13
Магнитофон	–	5	–	35	40
Гироскоп	13	–	–	40	53

В настоящее время определен жизненный цикл современных сложных технических систем (СТС)*: НИР, аван-проект, эскизный проект, рабочий проект, испытания, серийное изготовление, эксплуатация. При этом, как установил чл.-корр. АН СССР Д.М. Гвишиани, для новых изделий и их разработки в США в среднем менее 2% выдвигаемых идей воплощаются в новую, пользующуюся спросом на рынке продукцию [22]. Тенденция отказа от идей или проектов по мере продвижения по стадиям процесса создания новой конструкции представлена на рис. 3.

Поиску новых перспективных технологий особое значение придается в США. С этой целью Управление перспективных исследовательских программ национальной разведки США (IARPA) ищет новые способы выявления и приоритеты перспективных технологий по всему миру [8]. Разработана и осуществляется исследовательская программа *FUSE (Foresight and Under Standing from Scientific Exposition* – «Предвидение и понимание на научной основе») с участием ученых университетов, институтов, компаний, по которой предполагается получить надежную и автоматизированную систему отбора и оценки технической информации из открытых источников с тем, чтобы выбрать лучшие для глубокого изучения. Для российского оборонно-промышленного комплекса необходим, как отмечено в статье [16], технологический рывок, который немислим без научно обоснованного методического аппарата по выявлению новых передовых и зарождающихся технологий.

* Под системой понимается внутренне организованная на основе того или иного принципа целостность, в которой все элементы настолько тесно связаны друг с другом, что выступают по отношению к окружающим условиям и другим системам как нечто единое (Спиркин А.Г. Основы философии. – М.: Политиздат, 1988. – 592 с. – С. 179). Создаваемые и проектируемые в настоящее время системы характеризуются исключительной сложностью. Сложная система определяется большим числом элементов и выполняемых ими операций, высокой степенью связности элементов, сложностью алгоритмов выбора тех или иных управляющих воздействий и большими объемами перерабатываемой информации (Цвиркун А.Д. Структура сложных систем. – М.: Сов. радио, 1976. – 200 с. – С.7).

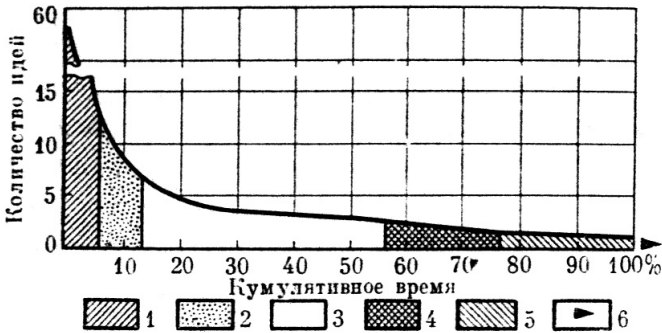


Рис. 3. Жизнеспособность проектов новых изделий. (по результатам обследования 51 фирмы). Эволюция по стадиям: 1 – отбор идей; 2 – анализ конъюнктуры; 3 – разработка проекта; 4 – испытания изделия; 5 – сбыт; 6 – одно удачное новое изделие

В связи с важностью выявления перспективных технологий, а также учитывая то, что интервал времени от идеи до внедрения для современных объектов техники сокращен (например, смена поколений компьютеров происходит за 1–3 года и менее), первостепенное значение приобретает фактор выявления новых прогрессивных технологий, которые обеспечат в будущем фирме-разработчику успех и опережающее развитие.

2. Характер развития современной техники

Анализ сложных СТС с зарождающимися технологиями приводит к тому, что в начале своего развития они имеют низкие характеристики, но по мере их совершенствования и использования происходит существенное улучшение технического уровня (ТУ) СТС.

На рис. 4. представлена динамика развития техники, основанной на старой технологии и новой [24, 25]. S-образная кривая отражает зависимость между затратами, связанными с улучшением продукта или технологического процесса, и результатами, полученными от вложения средств. Успех возникает вначале медленно, затем набирает темп, ускоряется, но потом неизбежно

сходит на нет, технология приходит в упадок. Фирмам важно знать о технологическом пределе и вовремя перейти к новой технологии. «Технологический сдвиг» – это периоды перехода от одной технологии к качественно другой (или от одного продукта к качественно другому, удовлетворяющему ту же потребность). С наступлением современного этапа научно-технической революции фирмы высокоразвитых стран вступили, по словам П. Дракера, в «век разрыва непрерывности»: растет частота технологических разрывов. Возникают новые условия конкуренции. Последние десятилетия дают много примеров, когда технологические разрывы означали исчезновение не только отдельных видов продукции, но и целых отраслей и приводили к падению и даже банкротству одних корпораций и к взлету других. Широко известна ситуация в электронной промышленности западных стран: здесь в середине 50-х годов, в начале эры современной электроники, объем продаж электронных ламп составлял примерно 700 млн долл., продажи транзисторов были ничтожны – 7 млн. долл. Но за четверть века произошел переход к транзисторам и интегральным схемам [21]. В отрасли произошла почти полная смена лидеров. Технологические разрывы – это сегодня одна из самых серьезных угроз, игнорировать которую не могут даже самые благополучные, процветающие фирмы западных стран.

Ярким примером неучета новых развивающихся технологий является деятельность фирмы «Кодак», которая была основана Джорджем Истманом в 1888 г. и являлась лидером пленочной технологии фотосъемки. Первый цифровой фотоаппарат также был создан этой фирмой. *Eastman Kodak* – ведущим брендом мировой фотоиндустрии. Однако впоследствии она утратила передовые позиции, вовремя не перешла на цифровую технику фотографирования и в настоящее время терпит банкротство, ее долг составляет 7 млрд. долл. [5].

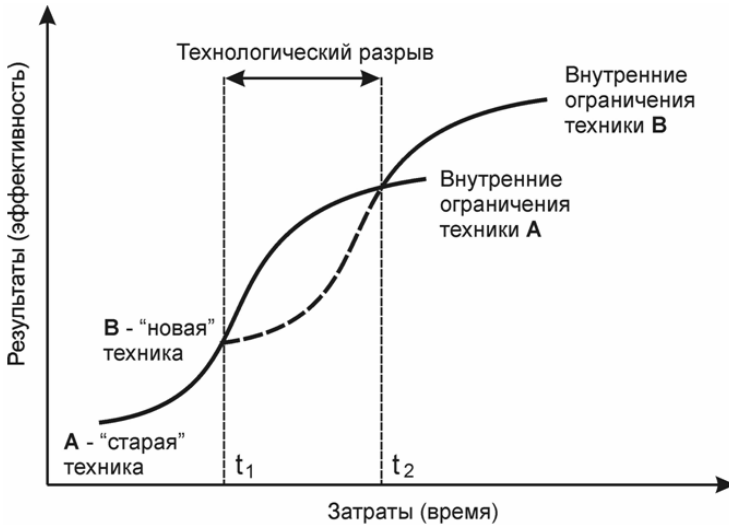


Рис. 4. S-образные кривые замещения одной технологии другой

В монографии [24] отмечено, что всегда бывает важным выбрать оптимальный момент начала работ (не слишком рано, не слишком поздно) по созданию продукции на основе новых технологий. Ведь лидер в какой-либо области всегда совершает большие ошибки и должен приложить больше усилий, чем те, кто следует за ним. За право быть первым в какой-либо области, иногда приходится платить слишком большую цену, особенно если работы были начаты слишком рано. Вот почему все нововведения следует патентовать, так как наличие патента дает право монопольного владения изобретением на определенный срок для компенсации финансовых затрат, а возможность использования этого нововведения другими фирмами может быть осуществлена путем приобретения лицензии у фирмы-патентообладателя. Опоздание с использованием новых технологий может привести к утрате уже завоеванных (достигнутых) позиций на мировом рынке. Но чтобы добиться успеха и освоить вовремя высококонкурентные технологии, необходимо научиться определять моменты перехода на новые технологии.

Можно привести пример из области новых высоких технологий, в том числе и информационных. Так, в 1990-е годы, когда

стала создаваться глобальная навигационная система позиционирования объектов *GPS NAVSTAR*, точность определения местоположения объектов была невелика (среднеквадратическое отклонение σ составляло около 60 м). Поэтому возможно еще к этой системе не был проявлен интерес со стороны потенциальных потребителей, хотя она являлась всепогодной и круглосуточной. И только после ее ввода в 1995–1996 гг., когда была обеспечена точность, позволяющая определять координаты положения объекта в любой зоне Земли с точностью 8–13 м, спутниковая навигация стала широко использоваться в гражданской сфере и военных технологиях, в особенности для обеспечения навигации и высокоточного наведения оружия, например управляемых авиационных бомб, на цель (рис. 5) [15, 18]. Повышение точности наведения с $\sigma = 14$ м до $\sigma = 3$ м увеличивает вероятность попадания в 4,5 раза.

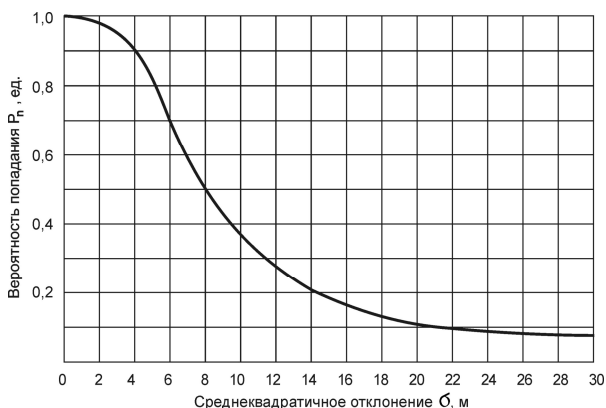


Рис. 5. Вероятность попадания УАБ $P_n = F(a)$ в цель при кругловом рассеивании и нормальном законе распределения

Поэтому решение проблемы выбора оптимального времени уже начатых работ по новым технологиям занимает важное место в менеджменте организаций и предприятий.

Одним из элементов системы управления качеством продукции и повышения ее конкурентоспособности является так называемая базовая оценка, под которой подразумевается

«оценка процессов, продукции и всех признаков компонент на основе сопоставлений с лучшими мировыми достижениями, в том числе и в других отраслях» [13]. Чтобы сделать метод базовых оценок более доступным для производителей, были созданы базы данных о наивысших достижениях. Например, фирма «Эрнст энд Янг» и Американский фонд качества провели исследования международных достижений в области качества. Было обследовано более 500 компаний из 4-х стран в автомобильной промышленности, банковской сфере, производства компьютеров, а также в отраслях, обслуживающих здравоохранение.

При этом были получены следующие результаты:

– сравнению с лучшими мировыми достижениями придается первостепенное значение (США и Япония около 30%, Канада – 25%);

– на предварительную оценку проектов технико-экономического обоснования планирования фирмами выделяется в среднем в США не менее 25% и Японии – до 40% от общих затрат.

О характере развития СТС в области военных технологий можно судить на примере развития УАБ и беспилотных летательных аппаратов (БЛА) во второй половине XX в. Анализируя динамику развития зарубежных УАБ с лазерными и телевизионными системами самонаведения по $K_{ТУ}$ за почти 30-летний период можно отметить, что со временем путем вложения значительных ассигнований постепенно совершенствуется технология создания УАБ и наблюдается плавный, почти линейный характер возрастания технического совершенствования УАБ с 1,0 до 1,35 (рис. 6) [18]. Это дает нам основание считать, что в соответствии с концепцией долгосрочного стратегического планирования развитие такого технического объекта, как УАБ, соответствует верхнему участку S -образной кривой B «результаты–затраты» или «эффективность–время» (рис. 3) [18].

Кривая A соответствует развитию обычных авиабомб (АБ) (рис. 3). Их развитие шло эволюционным путем (как за счет совершенствования свойств АБ, так и прицельно-навигационного комплекса самолета-носителя) также с плавным увеличением значения $TУ$. Появление УАБ в период 40-60-х годов, как показывают расчеты, отмечен скачкообразным возрастанием коэф-

фициента ТУ до 2–3 [17], что можно объяснить за счет перехода со старой технологии на новую (кривая *B*) (рис. 4).

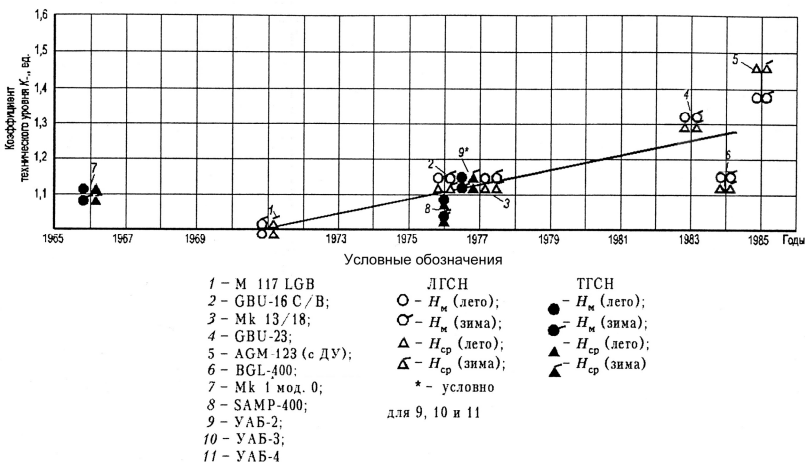


Рис. 6. Зависимость коэффициента технического уровня $K_{ТУ}$ УАБ калибра 500 кг в функции года окончания их разработки ($K_{ТУ}$ базы сравнения принята за единицу)

Полученные выводы подтверждаются данными о эффективности действия боевой авиации в различные исторические периоды [23]. Во вторую мировую войну для поражения одной цели размером 60×199 м (типа промышленного здания) было сброшено порядка 9000 бомб и было сделано около 1500 вылетов самолета типа В-17 (статистика по результатам бомбардировок Германии союзной авиацией). Во Вьетнаме истребитель-бомбардировщик В-4 «Фантом» для поражения одной цели делал 30 боевых вылетов, сбрасывая 176 неуправляемых бомб калибра 500 фунтов. В операции «Буря в пустыне» использовались уже управляемые авиабомбы с лазерными и телевизионными системами наведения. Самолет F-117 нес две 2000-фунтовые бомбы с лазерной ГСН и в одном вылете уничтожал две цели. В последней иракской войне за один вылет бомбардировщик В-2 нес 16 шт. 2000-фунтовых авиабомб и поража

новременно 16 целей. Эта наглядная демонстрация влияния технического прогресса на развитие оружия.

Дальнейшее эволюционное развитие УАБ безусловно можно связать с улучшением ТТХ УАБ (повышение точности наведения, увеличение дальности применения, расширение условий боевого применения по режимам сброса, временным и погодным условиям и т.д.), однако значительный прирост $K_{ТУ}$ следует ожидать от использования революционизирующих технологий, например, интернет-технологий, т.е. необходимо устранить «технологический разрыв»¹. Это использование высокоточных всепогодных систем управления и наведения (*GPS/INS*-технология), тепловизионной системы наведения *DAMASK* и т.п., способных надежно функционировать в любых географических и климатических условиях по широкому спектру целей (поверхностных/подповерхностных, в том числе подвижных) с осуществлением принципа «сбросил – забыл» и автоматической селекцией заданного класса целей, а также эффективных специализированных или универсальных боевых частей, аэродинамических схем, обеспечивающих маневренность и максимальную дальность полета УАБ, новых конструктивных технологий, снижающих радиолокационную заметность и т.п. Рост эффективности боевого применения УАБ в составе ударного авиационного комплекса по мере их совершенствования представлен на рис. 7 [19].

Как следует из рис. 6, вероятность поражения возросла с момента появления микроэлектроники в период с 1950 г. по 1967 г. в 3,5 раза и в период с 1990 г. по 1995 г. на 2–5%.

В диссертации [10] по комплексному коэффициенту качества K_K были определены характер и тенденции развития крылатых ракет или БЛА (рис. 8).

Приведенные примеры позволяют сделать вывод о том, что используя данные ретроспективного анализа о ТУ объекта ис-

¹ «Технологический разрыв» – это период перехода от одной технологии к качественно другой (или от одного продукта к качественно другому, удовлетворяющему ту же потребность). Современный этап научно-технической революции характеризуется ростом частоты технологических разрывов.

следования, а также его прогнозные значения, можно выявлять перспективные технологии развития создаваемой техники.

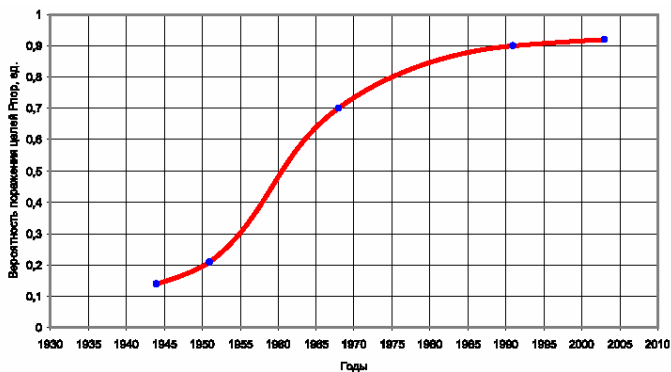


Рис. 7. Динамика потенциально достижимой вероятности поражения целей УАБ ($P_{пор}$)

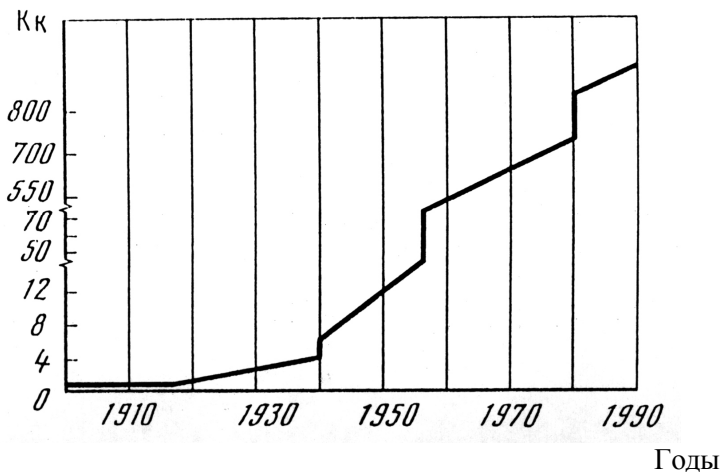


Рис. 8. Изменение по времени комплексного коэффициента качества БЛА

3. Существующие методы определения момента перехода на новые технологии

В монографии [18] был предложен метод оценки ТУ УАБ на основе весовых коэффициентов, который может быть также использован для определения тенденций развития различных СТС.

Для оценки методом сравнения всех показателей предложен критерий ТУ ($K_{\text{ТУ}}$), который характеризует разрабатываемый образец по отношению к существующим аналогам (отечественным или зарубежным):

$$(1) \quad K_{\text{ТУ}} = \frac{\sum_{i=1}^n K_{n_i} \varphi(i)}{\sum_{i=1}^n \varphi(i)}$$

где $K_{\text{ТУ}}$ – критерий технического уровня; $\varphi(i)$ – функция, нормирующая вес i -го относительного показателя, входящего в ранжированную последовательность; n – общее число показателей; K_{n_i} – относительное значение i -го показателя, значимость которого определяется местом, занимаемым в ранжированной последовательности.

Эти показатели применяются в двух модификациях:

если повышению ТУ соответствует увеличение рассматриваемого i -го показателя; например, увеличение дальности сброса КАБ (УАБ) (D_{max}), то относительное значение i -го показателя определяется по формуле

$$(2) \quad K_{n_i} = \frac{K_{o_i}}{K_{z_i}},$$

где K_{o_i} – величина i -го показателя разрабатываемого (сравниваемого) образца; K_{z_i} – величина i -го показателя аналога, принятого для сравнения;

если повышению ТУ соответствует уменьшение рассматриваемого показателя; например, при уменьшении КВО ($E_{\text{КВО}}$), то относительное значение i -го показателя определяется по формуле

$$(3) \quad K_{n_i} = \frac{K_{z_i}}{K_{o_i}}.$$

На основе данного метода был проведен ретроспективный анализ развития зарубежных УАБ и разработана аттестационная шкала (таблица 4) для оценки перспективности УАБ [18].

Таблица 4. Аттестационная шкала перспективности УАБ

Пределы изменения $K_{ТУ}$	Прогнозная оценка изделия в целом
1,05–1,065	Не перспективная УАБ
1,07–1,13	Малоперспективная УАБ
1,135–1,265	Перспективная УАБ
1,27 и выше	Весьма перспективная УАБ

Оценка технического уровня альтернативных образцов УАБ с использованием метода весовых коэффициентов осуществлялась при следующих оценочных показателях [18]:

- $E_{КВО}$, м – круговое вероятное отклонение;
- $m_{БЧ}$, кг – масса БЧ;
- $m_{ВВ}$, кг – масса ВВ БЧ;
- $D_{пр. макс}$, км – максимальная дальность сброса УАБ;
- $t = (1 - t_{упр. и нав.}/t_6)$ – коэффициент автономности полета УАБ после ее сброса с учетом степени участия летчика-оператора ($t_{упр. и нав.}$ – суммарное время управления и наведения УАБ на цель; t_6 – общее время полета УАБ после сброса);
- $T = T_{прим.}/24$ – коэффициент круглосуточности применения УАБ ($T_{прим.}$ – возможное время применения УАБ в течение суток, ч).;
- $B = B_{прим.}/100$ – коэффициент всепогодности применения УАБ ($B_{прим.}$ – суммарное время возможности применения (выполнения задачи наведения) по условиям продолжительности при метеоусловиях (дождь, снег, туман, дымка) в % к дням года).;
- $Об = Об_{прим.}/100$ – коэффициент применения УАБ по условиям отсутствия облачности ($Об_{прим.}$ – суммарное время возможности применения УАБ по условиям выполнения задачи

наведения при облачности в % к дням года).

Для СТС с зарождающимися технологиями в начальный период развития t_1 характерен низкий ТУ $K_{\text{ТУ}-1}$. Однако исследователь может ожидать в будущем развитие СТС, основанной на новой технологии. Тогда при проведении оценок по определению ТУ следует использовать для описания СТС с новыми технологиями прогнозные значения ее основных характеристик и произвести сравнение t_2 . (рис. 3). В этом случае при рассмотрении СТС на основании данных по новой технологии среди остальных альтернативных решений возможно получение положительных решений. При таком подходе исключается отсеивание новых технологий ввиду их неготовности на начальных стадиях разработки, как это было бы в случае сравнения СТС с имеющимися характеристиками в момент времени (t_1).

Известна также работа [11], в которой представлена модель оптимизации выпуска продукции на рынок, причем момент перехода определяется путем сравнения достигнутого уровня продукции с мировым. На рис. 9 представлена динамика мирового уровня качества и уровня качества продукции конкретного предприятия. Мировой уровень качества (наивысшие достижения основных характеристик продукции, обеспечиваемый передовыми фирмами) непрерывно растет, в то время как уровень качества продукции конкретного предприятия меняется скачкообразно. Он заметно увеличивается при выпуске на рынок новой марки продукции, а затем остается постоянным вплоть до выпуска следующей марки. В течение жизненного цикла очередной марки продукции ее уровень качества сначала заметно выше мирового, затем преимущество уменьшается, наконец, уровень качества оказывается ниже мирового, и через некоторое время марка снимается с производства.

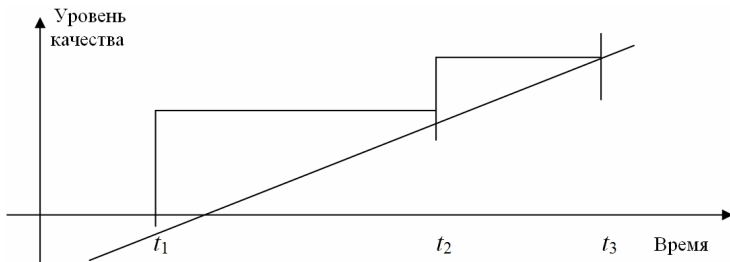


Рис. 9. Сравнение динамики мирового уровня качества и уровня качества продукции конкретного предприятия

4. Основные принципы и метод оценки технического уровня сложных технических систем с зарождающимися технологиями

Предполагаем, что процесс развития исследуемых СТС является эволюционным. Так как в предлагаемом методе оценки ТУ СТС с применением теории прогнозирования используется S-образный вид развития технологий, то рассмотрим эту зависимость более подробно.

Линейный и экспоненциальный рост функции, наблюдаемый на определенных этапах развития технологии, можно рассматривать как аппроксимированные отрезки логистической кривой [2, 14]. На рис. 10 экспоненциальный характер имеет часть логистической кривой в интервале $(t_1 - t_2)$, характеризующем период зарождения и экспериментального опробования технологии, которая повышает производительность процесса изготовления продукции. Этот период, когда СТС имеет малый удельный вес среди других аналогичных СТС, называется латентным периодом. Линейный характер зависимости появляется на этапе $(t_2 - t_3)$ наращивания выпуска продукции, изготавливаемой по новой технологии, который можно характеризовать как бурное развитие нового поколения СТС. Период насыщения соответствует интервалу $(t_3 - t_4)$, когда становится необходимым переход на принципиально новую технологию. Время жизни поколения СТС составляет интервал времени $(t_4 - t_1)$.

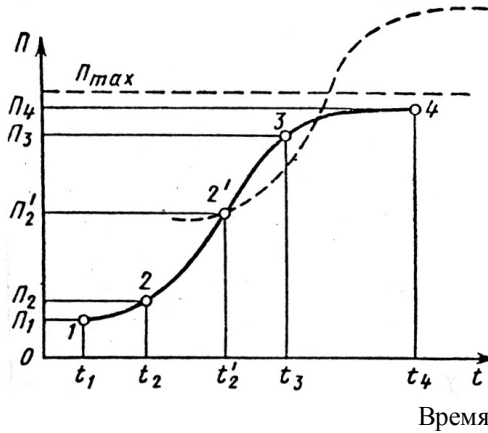


Рис. 10. Логистическая кривая роста производительности технологического процесса (Π): — старая технология; — — новая технология, зарождающаяся при существующей

Новую технологию разрабатывают не только после окончательного устаревания и отмирания прежних технологических методов. Иногда новая технология рождается еще при существовании экономически жизнеспособной старой технологии (точка t_2'). Определение момента замены технологических процессов, а, следовательно, оборудования, оснастки и т.д., является сложной задачей, решение которой возможно на базе всестороннего анализа современного состояния и прогнозирования перспектив технологии и производства, исходя из характера развития рассматриваемого процесса.

Функция, соответствующая логистической кривой (применительно к динамике производительности технологического процесса), имеет вид [26]

$$(4) \quad \Pi(t) = \frac{\Pi_{\max}}{1 + A \cdot \exp(-u \cdot t)}$$

где Π_{\max} — предельное значение производительности рассматриваемой технологии; u определяет темп роста производительности технологии, а поскольку

$$(5) \quad \frac{\Pi_{\max}}{\Pi(0)} = 1 + A,$$

то $A \times 100\%$ показывает – насколько процентов вырастет производительность рассматриваемой технологии по отношению к её начальному значению.

Анализ логистической зависимости (4) показывает, что при $A > 1$ наибольшее значение темпа роста производительности достигается для момента времени $t^* = \frac{\ln(A)}{u}$, откуда имеем соотношение

$$(6) \quad u = \frac{\ln(A)}{t^*}$$

Таким образом, для прогнозирования динамики развития новой технологии достаточно спрогнозировать числовые значения двух характеристик:

максимальный уровень производительности новой технологии;

время достижения максимального темпа роста новой технологии.

Используя затем соотношения (5)–(6), получаем прогнозные значения параметров A , u прогнозной модели, а из равенства (4) и прогноз самой динамики развития производительности новой технологии.

Для определения погрешности в прогнозных значениях $\Pi(t)$, зная порядок погрешностей прогнозных значений вышеуказанных двух характеристик, подсчитываем дисперсии прогнозных значений параметров A , u и дисперсии значений производительности для любой даты в будущем, используя формулы

$$(7) \quad \sigma^2(u) = \frac{\sigma^2(A)}{A^2 \cdot t^{*2}} + \left(\frac{\ln A}{t^{*2}} \right)^2 \cdot \sigma^2(t^*), \quad \sigma^2(A) = \frac{\sigma^2(\Pi_{\max})}{\Pi^2(0)},$$

$$\sigma^2(\Pi(t)) = \sigma^2(\Pi_{\max}) \cdot \frac{t}{t^*},$$

где значения дисперсий $\sigma^2(t^*)$, $\sigma^2(\Pi_{\max})$ определяются на основе экспертных оценок.

В частности, используя формулы (7), можно вычислять доверительные интервалы для параметров u , A и значений $\Pi(t)$ для любого заданного уровня доверия [6].

В случае, когда отсутствует возможность получить экспертную оценку максимальной производительности новой технологии [12], можно оценить два параметра логистической кривой по уже реализованной части роста производительности новой технологии, используя метод наименьших квадратов [6].

Следует отметить, что логистическая кривая широко используется для описания и прогнозирования различных процессов в природе [26].

Пример 1. Пусть определены следующие значения:

$$(8) \quad \frac{\Pi_{\max}}{\Pi(0)} \approx N(2,3; 0,1), \quad t^* \approx N(5; 0,2).$$

Имеем:

$$(9) \quad \sigma^2(u) \cong 0,0024, \quad \sigma^2\left(\frac{\Pi(t)}{\Pi(0)}\right) = 0,02 \cdot t.$$

Пример 2. Рассмотрим пример развития новой технологии, связанной с ЛА. В 1990 г. точность наведения ЛА на объекты составляла $T = 60$ м (1990 год возьмем за начало отсчета). Известно, что в 1995 г. точность составляла $T = 11 \pm 2$ м, а в 2003 г. уже $T = 4 \pm 1$ м. Определим производительность новой технологии согласно равенству $\Pi = 100 - T$. Тогда $\Pi(0) = 100 - 60 = 40$, $\Pi(5) = 89 \pm 2$, $\Pi(13) = 96 \pm 1$. Таким образом, в 2003 году с помощью новой технологии по существу было достигнуто максимальное значение её производительности. Следовательно, для рассматриваемого примера имеем следующие значения:

$$\frac{\Pi_{\max}}{\Pi(0)} = \frac{96}{40} = 2,4; \quad A = \frac{\Pi_{\max}}{\Pi(0)} - 1 = 1,4.$$

Значение u находим из равенства

$$\frac{\Pi(5)}{\Pi(0)} = 2,225 = \frac{2,4}{1 + 1,4 \cdot \exp(-5 \cdot u)},$$

откуда получаем $u = 0,576$.

Следовательно, функция, соответствующая динамике производительности технологического процесса, имеет вид

$$(10) \quad \frac{\Pi(t)}{\Pi(0)} = \frac{2,4}{1 + 1,4 \cdot \exp(-0,576 \cdot t)}$$

На рис. 11 представлен график функции (10), на оси абсцисс отложены годы (нулю соответствует 1990 г.), на оси ординат отложено отношение $\frac{\Pi(t)}{\Pi(0)}$.

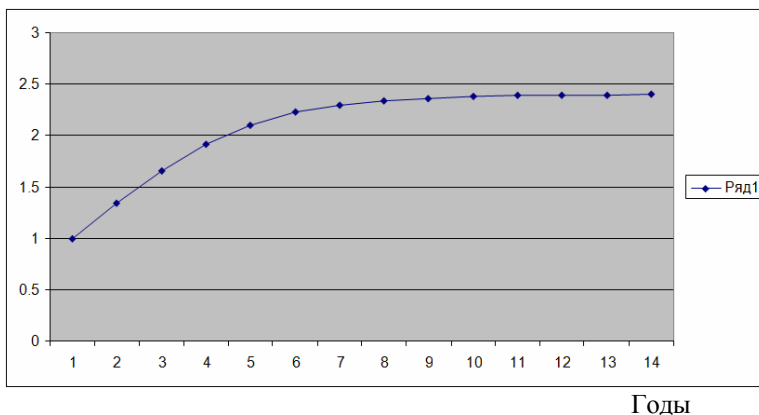


Рис. 11. Динамика производительности технологического процесса ЛА

С учетом вышесказанного, общая схема метода оценки ТУ СТС, известная по статье [20], при оценке СТС с зарождающимися технологиями будет изменена таким образом, как представлено на рис. 12. Модифицированный метод оценки ТУ СТС, в числе которых имеются СТС, основанные на зарождающихся технологиях, содержит этап 6, на котором производится определение прогнозных значений основных характеристик СТС с зарождающимися технологиями.

Данный метод реализован в информационно-аналитической системе «Оценка и выбор» [17]. Следует заметить, что при оценке ТУ СТС принимается во внимание совокупность оценочных показателей, при этом может получиться, что даже прогрессивный характер одного из анализируемых параметров новой технологии может не дать значительного прироста ТУ.

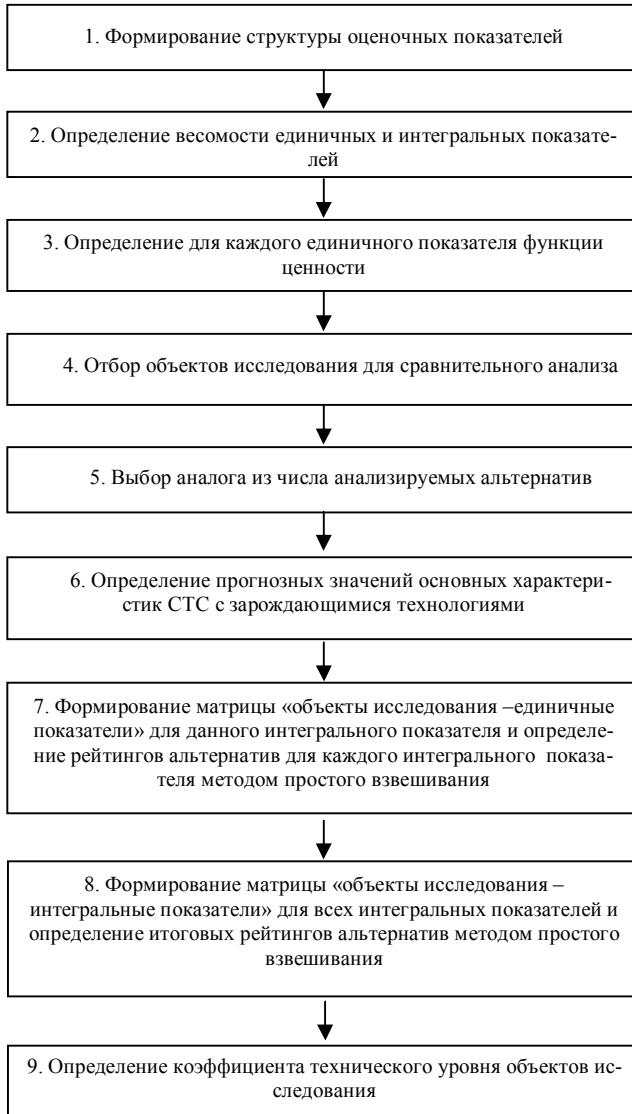


Рис. 12. Этапы решения задачи сравнительного анализа СТС с помощью ИАС «Оценка и выбор»

5. Выводы

1. Анализ создания новых технологий показал, что развитие современных СТС соответствует логистической кривой роста

та эффективности, при этом только около 2% выдвигаемых идей воплощается в новую продукцию, а период времени от зарождения до внедрения в среднем составляет около 15 лет в XX в., с тенденцией сокращения в XXI в.

2. Предложенный метод определения момента перехода при создании СТС со старой технологии на новую базируется на прогнозной экспертной оценке развития показателей СТС.

3. Разработанный алгоритм оценки смены технологий при создании СТС обоснован и апробирован. Внедрение данного метода будет способствовать созданию высокотехнологичной продукции.

Литература

1. БИЛИНКИС В.Д. *Методы оценки технического уровня и конкурентоспособности продукции.* – Воронеж: Изд-во ВГТУ, 2000. – 118 с.
2. ГАЙКОВИЧ А.И. *Основы теории проектирования сложных технических систем.* – СПб.: НИЦ «МОРИНТЕХ», 2001. – 432 с.
3. ГЛАЗЬБЕВ С.Ю. *Технологические сдвиги в экономике России // Экономика и математические методы.* – 1992. – Том 33, вып. 2. – С. 5–25.
4. ГРИНЯЕВ С. *Цифровое неравенство наций // Независимое военное обозрение.* – 2004. – №3(363). – С. 4.
5. КОРОЛЕВА А. *Конец эпохи Kodak. Компания Eastman Kodak объявила о банкротстве // Expert Online.* – 20 января 2012 года. (expert.ru/2012/01/20/konets-epohi-kodak/-80k)
6. КРЯНЕВ А.В., ЛУКИН Г.В. *Метрический анализ и обработка данных.* – М.: Физматлит, 2010. – 280 с.
7. *Нанотехнологии как ключевой фактор нового технологического уклада в экономике /* Под ред. акад. РАН С.Ю. Глазьева и проф. В.В. Харитонова. – М.: Изд-во «Триват», 2009. – 304 с.
8. *Национальная разведка США ищет перспективные технологии // Зарубежное военное обозрение.* – 2010. – №10. – С. 89.
9. НЕЩАДИН А.А., КАШИН В.К., ЛИПСИЦ И.В., ВИГДОРЧИК Е.А. *Экономическое развитие страны и конверсия // Конверсия в машиностроении.* – 2000. – №1. – С. 6–20.

10. НОВИЧКОВ Н.Н. *Развитие крылатых ракет самолетных схем*. Дис. канд. техн. наук. – М., 1982. – 341 с.
11. ОРЛОВ А.В. *Менеджмент: Организационно-экономическое моделирование*. – Ростов н/Д.: Феникс, 2009. – 475 с.
12. ОРЛОВ А.В. *Организационно-экономическое моделирование*. Ч.2. Экспертные оценки. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2011. – 486 с.
13. ПОРТ О., КЭРИ Д. *Стремление к совершенству. Специальный репортаж* // Бизнес Уик. – 1992. – №11. – С. 20–27.
14. ПРОСКУРЯКОВ А.В. *Организация создания и освоения новой техники*. – М.: Машиностроение, 1975. – 224 с.
15. *Проблемы функционирования глобальной спутниковой навигационной системы GPS после 2000 г.: Экспресс-информация* // ГосНИИАС. – 1998. – №27–28. – С. 8–9..
16. РОГОЗИН Д. России нужна «умная оборонка» // Газета «Красная звезда». – 2012. – 7 февраля.
17. СЕМЕНОВ С.С., ХАРЧЕВ В.Н. *Проблемы создания корректируемых и управляемых авиационных бомб*. – М.: Инженер, 2003. – 528 с.
18. СЕМЕНОВ С.С., ХАРЧЕВ В.Н., ИОФФИН А.И.. *Оценка технического уровня образцов вооружения и военной техники*. – М.: Радио и связь, 2004. – 552 с.
19. СЕМЕНОВ С.С., ХАРЧЕВ В.Н. *Корректируемые авиабомбы российских ВВС*. – М.: Изд. группа «Бедретдинов и Ко», 2005. – 88 с.
20. СЕМЕНОВ С.С., ЩЕРБИНИН В.В. *Метод оценки технического уровня систем наведения управляемых авиационных бомб* // Вопросы оборонной техники. Сер. 9. Специальные системы управления, следящие приводы и их элементы. – М.: ФГУП «НТЦ «Информтехника». – 2010. – Вып. 1(242)–2(243). – С. 29–32.
21. СМИРНОВ А.Д., МАКСИМОВ В.Ф., АКУЛЕНКО Д.Н. и др. *Рыночная экономика*. Т. 2. Часть 1.– М.: Соминтэк, 1992. – 160 с.
22. УВАРОВА Л.И. *Научный прогресс и разработки технических средств*. – М.: Наука, 1973. – 272 с.
23. ФЕДОСОВ Е.А. *Высокоточное оружие и технологии сетевого целеуказания и управления – основы глобального удара*

боевой авиации в XXI веке // Труды юбилейной научно-технической конференции «Авиационные системы в XXI веке». Том 1. – М.: ГИЦ РФ «ФГУП «ГосНИИАС», РАРАН, 2006. – С. 9–29.

24. Эйрес Р. *Научно-техническое прогнозирование и долгосрочное планирование* / Под ред. Г.М. Доброва. – М.: Мир, 1971. – 296 с.
25. Янч Э. *Прогнозирование научно-технического прогресса*. – М.: Прогресс, 1974. – 591 с.
26. KINGSLAND S.E. *Modeling nature*. – Chicago: University of Chicago, 1985.

DEVELOPMENT OF MODERN TECHNOLOGY AND METHOD TO ESTIMATE TECHNOLOGICAL LEVEL OF COMPLEX TECHNICAL SYSTEMS BASED ON ARISING TECHNOLOGIES

Kryanev Alexander, National research nuclear university “МЭФТИ”, Moscow, Dr. of sciences, professor of chair “Applied mathematics” (*avkryanev@mephi.ru*).

Semenov Serguey, Open Society “State research-and-production enterprise “Region”, Moscow, PhD, the head of analysis and perspective researches group (*gnppregion@sovintel.ru*).

We consider the problem of technological level estimation of complex technical systems, which are based on new arising technologies. We characterize the processes of contemporary technology development, survey known methods of choosing the moment of new technology adoption, and explain motivation for new methods elaboration. We suggest the principle of forecasting basic characteristics of a new technology and introduce the novel method for technological level estimation of complex technical systems with arising technologies.

Keywords: technological level, technological level factor, complex technical system, life cycle, arising new technology, forecasteed values, rate of new technology growth, the moment of transition to new technology.

Статья представлена к публикации членом редакционной коллегии Р. М. Нижегородцевым

УДК 513.88:519.177:519.173:519.816

ББК 22.18

ПРОЕКЦИОННО-ИТЕРАЦИОННЫЕ ПРОЦЕДУРЫ ПРИБЛИЖЕННОГО ПОСТРОЕНИЯ ВЫНУЖДЕННЫХ КОЛЕБАНИЙ В НЕЛИНЕЙНЫХ СИСТЕМАХ

Исмаилов И. Г.¹

(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Приводятся результаты, относящиеся к проблемам приближенного построения режимов функционирования сложных нелинейных систем и специальным численным процедурам. Исследуются некоторые аспекты теории итерационных и проекционных процедур в задачах приближенного построения колебательных режимов в нелинейных системах, в частности в системах автоматического регулирования. Предлагается процедура выбора начального приближения итерационного метода, основанная на использовании априорной информации об отыскиваемом колебательном режиме.

Ключевые слова: оптимальное управление, автоматическое регулирование, вынужденные колебания, приближенные методы.

1. Введение

Итерационные и проекционные методы составляют два важнейших класса численных процедур, применяемых для построения решений различных задач, возникающих в теории управления и оптимизации. Построение общей теории таких методов базируется на идеях и методах современного функ-

¹ Илхам Гусейнкулу оглы Исмаилов, кандидат физико-математических наук, старший научный сотрудник (Москва, ул. Профсоюзная, д.65, тел. (495) 334-79-00, iig07@mail.ru).

ционального анализа, которые не только позволили существенно упростить взгляд на природу многих численных методов, но и привели к разработке принципиально новых вычислительных схем в проблемах современной теории управления. Одна из таких схем описана далее в задачах приближенного построения колебательных режимов в нелинейных системах.

Многочисленная литература посвящена приближенному построению колебательных режимов в нелинейных системах (см., например, работы [1–18]). Особое место в числе прочих занимает метод гармонического баланса; он сводит приближенное построение к решению конечной системы скалярных уравнений (в общем случае нелинейных). Решая эту систему, находят коэффициенты тригонометрических многочленов, являющихся приближениями к отыскиваемому колебательному режиму.

Для решения системы уравнений гармонического баланса часто применяют итерационные процедуры. Скорость сходимости таких процедур существенно зависит от выбора начального приближения. Мы показываем зависимость скорости сходимости проекционно-итерационной процедуры от выбора вспомогательного аппроксимирующего подпространства малой размерности. Если отыскиваемый колебательный режим достаточно хорошо аппроксимируется тригонометрическим многочленом низкого порядка, то в качестве начального приближения следует брать решение, найденное методом гармонического баланса этого порядка. Информация о структуре колебания используется и для оценки погрешности метода гармонического баланса с большим числом гармоник.

2. Постановка задачи

Рассмотрим систему

$$(1) \quad \frac{dx}{dt} = Ax + f(t, x).$$

Здесь $x = \{x_1, \dots, x_N\}$ – вектор евклидова пространства \mathbf{R}^N , $A = \{a_{ij}\}$, $i, j = 1, \dots, N$, – постоянная матрица; $f: \mathbf{R} \times \mathbf{R}^N \rightarrow \mathbf{R}^N$ –

непрерывная по совокупности переменных, T -периодическая по t : $f(t + T, x) = f(t, x)$ вектор-функция $f(t, x)$, удовлетворяющая условию Липшица

$$(2) \quad |f(t, x_1) - f(t, x_2)| \leq q |x_1 - x_2|$$

(здесь и ниже через $|\cdot|$ обозначается евклидова норма в \mathbf{R}^N).

Предположим, что система (1) имеет T -периодический режим $x_*(t) = \{x_*^{(1)}(t), \dots, x_*^{(N)}(t)\}$.

В методе гармонического баланса приближения $x_n(t) = \{x_{n,1}(t), \dots, x_{n,N}(t)\}$ ищутся в виде тригонометрических полиномов

$$(3) \quad x_{n,s}(t) = a_{0,s} + \sum_{k=1}^n \left(a_{k,s} \cos \frac{2\pi kt}{T} + b_{k,s} \sin \frac{2\pi kt}{T} \right).$$

Для отыскания коэффициентов $a_{k,s}$, $b_{k,s}$ составляется равенство

$$(4) \quad \frac{dx_n(t)}{dt} = Ax_n(t) + P_n f(t, x_n(t)),$$

где $P_n g(t)$ – отрезок ряда Фурье T -периодической функции $g(t)$, в котором удержаны гармоники до порядка n . Приравнивая коэффициенты при одинаковых гармониках в левой и правой частях равенства (4), получаем систему $(2n + 1)N$ уравнений

$$(5) \quad \mathbb{R}_k(a_{0,1}, \dots, a_{n,1}, b_{1,1}, \dots, b_{n,1}, \dots, a_{0,N}, \dots, a_{n,N}, b_{1,N}, \dots, b_{n,N}) = 0$$

с $N(2n + 1)$ неизвестными. Уравнения (5) образуют систему уравнений гармонического баланса. Находя решения этой системы и подставляя их в (3), получаем компоненты приближения $x_n(t)$ к отыскиваемому периодическому режиму $x_*(t)$. Системы (4) и (5) эквивалентны.

Систему (5) решают приближенно и в (3) подставляют найденные приближения.

Метод последовательных приближений применяют непосредственно к системе (4). Предположим, что линейное уравнение $dx/dt = Ax$ не имеет ненулевых T -периодических решений, а начальное приближение $x_n^0(t)$ – это тригонометрический полином с компонентами вида (3). Если приближение $x_n^k(t)$ уже построено, то приближение $x_n^{k+1}(t)$ определяется из уравнения

$$\frac{dx_n^{k+1}(t)}{dt} = Ax_n^{k+1}(t) + P_n f(t, x_n^k(t)), \quad k = 0, 1, \dots$$

3. Теоремы сходимости

Задача отыскания T -периодических режимов системы (1) эквивалентна отысканию решений интегрального уравнения

$$(6) \quad x(t) = \int_0^T H(t-s; T) f(s, x(s)) ds.$$

Функция $H(t; T)$ определена на всей оси, T -периодична и $H(t; T) = (I - e^{TA})^{-1} e^{tA}$, $0 \leq t < T$.

Уравнение (6) будем изучать в пространстве L_2 вектор-функций $x(t)$, определенных на $[0, T]$, измеримых и суммируемых с квадратом. Норма и скалярное произведение в L_2 определяются равенствами

$$\|x\| = \left(\int_0^T \|x(t)\|^2 dt \right)^{1/2}; \quad (x, y) = \int_0^T (x(t)y(t)) dt.$$

Положим

$$\mathcal{H}x = \int_0^T H(t-s; T) x(s) ds,$$

$$f(x) = f(t, x(t)).$$

Линейный оператор \mathcal{H} действует в L_2 и ограничен, а оператор суперпозиции f действует в L_2 и удовлетворяет условию Липшица

$$(7) \quad \|f(x_1) - f(x_2)\|_{L_2} \leq q \|x_1 - x_2\|_{L_2},$$

где q – коэффициент из (2). Уравнение (6) можно записать в операторной форме:

$$(8) \quad x = \mathcal{H}f(x), \quad x \in L_2.$$

В силу (7) оператор $\mathcal{H}f : L_2 \rightarrow L_2$ удовлетворяет условию Липшица с константой $q \|\mathcal{H}\|$, где $\|\mathcal{H}\|$ – норма оператора $\mathcal{H} : L_2 \rightarrow L_2$. Если выполнено неравенство

$$(9) \quad q \|\mathcal{H}\| < 1,$$

то оператор $\mathcal{H}f : L_2 \rightarrow L_2$ удовлетворяет условиям принципа сжимающих отображений (см., например, работу [2]) и, следовательно, уравнение (8) (а значит, и уравнение (6)) имеет единственное решение $x^* = x^*(t)$, являющееся T -периодическим решением системы (1). При этом последовательные приближения (10) $x_{k+1} = \mathcal{H}f(x_k)$, $k = 0, 1, \dots$ сходятся к x^* при любом начальном приближении x_0 и

$$\|x_* - x_k\|_{L_2} \leq \frac{(q \|\mathcal{H}\|)^k}{1 - q \|\mathcal{H}\|} \|x_0 - \mathcal{H}f(x_0)\|_{L_2}.$$

При реализации метода (10) предварительно переходят от уравнения (8) к какой-либо его конечномерной аппроксимации, для отыскания решений которой и применяют метод последовательных приближений. Построение конечномерного уравнения, аппроксимирующего уравнение (8), осуществляют либо при помощи метода конечных элементов, либо заменяют интеграл механическими квадратурами, используют метод коллокации и др. Ниже используется метод гармонического баланса.

Пусть H_n – это $(2n + 1)N$ -мерное подпространство пространства L_2 , базис в котором образуют элементы

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \cos \frac{2\pi t}{T}, \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \sin \frac{2\pi t}{T}, \dots, \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \cos \frac{2\pi n t}{T}, \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \sin \frac{2\pi n t}{T},$$

.....

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ N \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ N \end{bmatrix} \cos \frac{2\pi t}{T}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ N \end{bmatrix} \sin \frac{2\pi t}{T}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ N \end{bmatrix} \cos \frac{2\pi n t}{T}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ N \end{bmatrix} \sin \frac{2\pi n t}{T}.$$

Пусть $P_n : L_2 \rightarrow H_n$ – ортопроектор пространства L_2 на H_n . Тогда приближения $X_n(t) = \{x_{n,1}(t), \dots, x_{n,N}(t)\}$ к отыскиваемому периодическому режиму $x^*(t)$ системы (1), полученные методом гармонического баланса, совпадают с решениями рассматриваемого на H_n операторного уравнения

$$(11) \quad x_n = P_n \mathcal{H} f(x_n).$$

Поскольку $\|P_n\| = 1$, $n = 1, 2, \dots$, то при выполнении условия (9) оператор $P_n \mathcal{H} f : H_n \rightarrow H_n$ является сжимающим. Таким образом, справедливо следующее утверждение:

Теорема 1. Пусть выполнено неравенство (9). Тогда последовательные приближения

$$(12) \quad x_n^{k+1} = P_n \mathcal{H} f(x_n^k), \quad k = 0, 1, \dots$$

сходятся к единственному решению $x_n^* = x_n^*(t)$ уравнения (11) и

$$(13) \quad \|x_n^* - x_n^k\|_{L_2} \leq \frac{(q \|\mathcal{H}\|)^k}{1 - q \|\mathcal{H}\|} \|x_n^0 - P_n \mathcal{H} f(x_n^0)\|_{L_2}.$$

Аналоги теоремы 1 хорошо известны. Здесь она приведена для удобства читателя.

Оценка (13) существенно зависит от начального приближения x_n^0 : если норма невязки $x_n^0 - P_n \mathcal{H} f(x_n^0)$ велика, а число $q \|\mathcal{H}\|$ близко к единице, то метод (12) может оказаться нереализуемым. Если имеется дополнительная информация о том, что отыскиваемый периодический режим достаточно хорошо аппроксимируется тригонометрическим многочленом малой степени m , то в качестве начального приближения целесообразно взять приближенное решение системы (1), построенное методом гармонического баланса с m гармониками.

Теорема 2. Пусть $m < n$, а $x^0 = x^0(t)$ – решение рассматриваемой в H_m системы

$$(14) \quad \frac{dx(t)}{dt} = Ax(t) + P_m f(t, x(t)).$$

Пусть для некоторого $\varepsilon > 0$ выполнена оценка

$$(15) \quad \|x_* - P_m x_*\|_{L_2} \leq \varepsilon.$$

Тогда

$$(16) \quad \|x_* - x^0\|_{L_2} \leq \frac{\varepsilon}{1 - q \|\mathcal{H}\|}$$

и

$$(17) \quad \|P_n \mathcal{H} f(x^0) - x^0\|_{L_2} \leq \frac{1 + q \|\mathcal{H}\|}{1 - q \|\mathcal{H}\|} \varepsilon.$$

Доказательство. Установим вначале оценку (16). Поскольку $x^0 = x^0(t)$ – это решение системы (14), то $x^0 = P_m \mathcal{H} f(x^0)$.

Но тогда в силу неравенств (7) и (15)

$$\begin{aligned} \|x_* - x^0\|_{L_2} &= \|x_* - P_m \mathcal{H} f(x^0)\|_{L_2} \leq \\ &\leq \|x_* - P_m x_*\|_{L_2} + \|P_m x_* - P_m \mathcal{H} f(x^0)\|_{L_2} \leq \\ &\leq \varepsilon + \|P_m(\mathcal{H} f(x_*) - \mathcal{H} f(x^0))\|_{L_2} \leq \varepsilon + \|P_m\| \|\mathcal{H}\| q \|x_* - x^0\|_{L_2} = \\ &= \varepsilon + \|\mathcal{H}\| q \|x_* - x^0\|_{L_2}. \end{aligned}$$

Сравнивая начало и конец этой цепочки соотношений, приходим к оценке (16).

Перейдем к доказательству оценки (17). В силу (16)

$$\begin{aligned} \|x^0 - \mathcal{H} f(x^0)\|_{L_2} &\leq \|x^0 - x_*\|_{L_2} + \|\mathcal{H} f(x_*) - \mathcal{H} f(x^0)\|_{L_2} \leq \\ &\leq \|x^0 - x_*\|_{L_2} + q \|\mathcal{H}\| \|x_* - x^0\|_{L_2} = (1 + q \|\mathcal{H}\|) \|x_* - x^0\|_{L_2} \leq \\ &\leq \frac{1 + q \|\mathcal{H}\|}{1 - q \|\mathcal{H}\|} \varepsilon. \end{aligned}$$

Поскольку $m < n$, а $x^0 \in H_m$ то $P_n x^0 = x^0$. Поэтому

$$\begin{aligned} \|P_n \mathcal{H} f(x^0) - x^0\|_{L_2} &= \|P_n \mathcal{H} f(x^0) - P_n x^0\|_{L_2} \leq \\ &\leq \|P_n\| \|\mathcal{H} f(x^0) - x^0\|_{L_2} = \|\mathcal{H} f(x^0) - x^0\|_{L_2} \leq \\ &\leq \frac{1 + q \|\mathcal{H}\|}{1 - q \|\mathcal{H}\|} \varepsilon. \end{aligned}$$

Теорема доказана.

Из теорем 1 и 2 вытекает

Теорема 3. Пусть выполнены неравенства (9) и (15).

Пусть начальное приближение метода (12) – это решение $x^0 = x^0(t)$ системы (14). Тогда

$$\|x_n^* - x_n^k\|_{L_2} \leq \frac{(q \|\mathcal{H}\|)^k (1 + q \|\mathcal{H}\|)}{(1 - q \|\mathcal{H}\|)^2} \varepsilon.$$

Теоремы 1 и 3, обосновывающие сходимость метода последовательных приближений построения решений уравнений гармонического баланса, не содержат оценок близости конструируемых приближений x_n^k к отыскиваемому T -периодическому

режиму $x_* = x_*(t)$. Однако, если известна какая-либо оценка нормы этого режима, то удастся получить и оценки норм $\|x_* - x_n^k\|$.

Пусть W_2^1 – пространство определенных на $[0, T]$ абсолютно непрерывных функций $x(t)$ со значениями в \mathbf{R}^N , производные которых суммируемы с квадратом. Пространство W_2^1 гильбертово; скалярное произведение и норма в нем определяются равенствами:

$$(x, y)_{W_2^1} = \int_0^T (x(t), y(t)) dt + \int_0^T (x'(t), y'(t)) dt;$$

$$\|x\|_{W_2^1} = \left(\int_0^T \|x(t)\|^2 dt + \int_0^T \|x'(t)\|^2 dt \right)^{1/2}.$$

Теорема 4. Пусть верна оценка (9) и

$$(18) \quad \|x_*\|_{W_2^1} \leq \rho.$$

Тогда

$$(19) \quad \|x_* - x_n^k\|_{L_2} \leq \frac{T\rho}{2\pi(1-q\|\mathcal{H}\|)(n+1)} + \frac{(q\|\mathcal{H}\|)^k}{1-q\|\mathcal{H}\|} \|x_n^0 - P_n \mathcal{H}f(x_n^0)\|_{L_2}.$$

Доказательство. Докажем вначале два вспомогательных утверждения.

Обозначим через \widetilde{W}_2^1 пространство определенных на $[0, T]$ абсолютно непрерывных вектор-функций $x(t)$ со значениями в \mathbf{R}^N , имеющих нулевые средние и производные, суммируемые с квадратом. Это пространство гильбертово; скалярное произведение и норма в нем определяются равенствами:

$$(x, y)_{\widetilde{W}_2^1} = \int_0^T (x'(t), y'(t)) dt, \quad \|x\|_{\widetilde{W}_2^1} = \left(\int_0^T \|x'(t)\|^2 dt \right)^{1/2}.$$

Пусть P_n – ортопроектор L_2 на подпространство H_n . Поскольку пространство \widetilde{W}_2^1 вложено в L_2 , то P_n можно рассматривать и как оператор, действующий из \widetilde{W}_2^1 в H_n . Непосредствен-

ная проверка показывает, что оператор P_n является ортопроектором и в \widetilde{W}_2^1 .

Рассмотрим на шаре $\|x\|_{\widetilde{W}_2^1} \leq \mathbb{R}$ функционал

$$F(x) = \frac{1}{2} \|x - P_n x\|_{L_2}^2.$$

Лемма 1. Пусть $B(\rho) = \{x \in \widetilde{W}_2^1 : \|x\| \leq \rho\}$.

Тогда справедливо равенство

$$(20) \quad \sup_{x \in B(\rho)} F(x) = \frac{1}{8} \left(\frac{T\rho}{\pi(n+1)} \right)^2.$$

Доказательство. Естественное вложение пространства \widetilde{W}_2^1 в L_2 компактно, т.е. шар $B(\rho)$ компактен в L_2 . Поэтому в силу теоремы Вейерштрасса функционал F достигает на шаре $B(\rho)$ своего максимума в некоторой точке x_* . Покажем, что $P_n x_n = 0$.

Действительно, если $P_n x_n \neq 0$, то

$$\begin{aligned} \|x_*\|_{\widetilde{W}_2^1}^2 &= \|P_n x_* + (I - P_n)x_*\|_{\widetilde{W}_2^1}^2 = \\ &= \|P_n x_*\|_{\widetilde{W}_2^1}^2 + \|(I - P_n)x_*\|_{\widetilde{W}_2^1}^2 > \|(I - P_n)x_*\|_{\widetilde{W}_2^1}^2. \end{aligned}$$

Поэтому при некотором $\lambda > 1$ точка $y_* = \lambda(I - P_n)x_*$ лежит в шаре $B(\rho)$ и, следовательно, $F(y_*) \leq F(x_*)$; с другой стороны, $F(y_*) = \lambda^2 F((I - P)x_*) = \lambda^2 F(x_*) > F(x_*)$.

Полученное противоречие показывает, что

$$(21) \quad P_n x_* = 0.$$

В силу теоремы Люстерника [1] в точке максимума x_* при некотором μ выполнено равенство

$$(22) \quad \nabla F(x_*) = \mu^2 x_*;$$

здесь через ∇F обозначен градиент функционала F на пространстве \widetilde{W}_2^1 , причем

$$(23) \quad \begin{aligned} \nabla F(x) &= \int_0^t (t-s)(P_n x(s) - x(s)) ds - \\ &- \frac{1}{T} \int_0^T \int_0^t (t-s)(P_n x(s) - x(s)) ds dt. \end{aligned}$$

Из (21), (22) и (23) вытекает равенство

$$\int_0^t (t-s)x_*(s)ds - \frac{1}{T} \int_0^T \int_0^t (t-s)x_* ds dt = -\mu^2 x_*(t).$$

Следовательно, функция $x_*(s)$ является нетривиальным решением двухточечной краевой задачи $x'' = -1/\mu^2 x$, $\dot{x}(0) = \dot{x}(T) = 0$, удовлетворяющим условию (21). Такие решения имеют вид $x_k(t) = a \cos 2k\pi t/T$, где $a \in \mathbf{R}^N$, $a \neq 0$, $k \in N$, $k > n$.

Непосредственный подсчет показывает, что максимум функционала F достигается при $k = n + 1$, $|a| = \sqrt{T/2} \cdot \rho / \pi(n + 1)$ и равен $1/8(T\rho/\pi(n + 1))^2$. Лемма доказана.

Лемма 2. Пусть выполнены неравенство (9) и оценка (18). Тогда

$$(24) \quad \|x_* - x_n^*\|_{L_2} \leq \frac{T\rho}{2\pi(1-q\|\mathcal{H}\|)(n+1)}.$$

Доказательство. Поскольку $x_* = \mathcal{H}f(x_*)$ и $x_n^* = P_n \mathcal{H}f(x_n^*)$, то

$$\begin{aligned} \|x_* - x_n^*\|_{L_2} &\leq \|x_* - P_n x_*\|_{L_2} + \|P_n x_* - x_n^*\|_{L_2} = \\ &= \|x_* - P_n x_*\|_{L_2} + \|P_n \mathcal{H}f(x_*) - P_n \mathcal{H}f(x_n^*)\|_{L_2} \leq \\ &\leq \|x_* - P_n x_*\|_{L_2} + q\|\mathcal{H}\|\|x_* - x_n^*\|_{L_2}. \end{aligned}$$

Сравнивая начало и конец цепочки соотношений (24), получаем

$$(25) \quad \|x_* - x_n^*\|_{L_2} \leq \frac{\|x_* - P_n x_*\|_{L_2}}{1 - q\|\mathcal{H}\|}.$$

Но в силу оценки (18) и равенства (20)

$$(26) \quad \|x_* - P_n x_*\|_{L_2} \leq \sup_{x \in B(\rho)} \|x - P_n x\|_{L_2} \leq \frac{T\rho}{2\pi(n+1)}.$$

Из (25) и (26) вытекает неравенство (24). Лемма доказана.

Перейдем к доказательству теоремы 4. В силу неравенства треугольника

$$\|x_* - x_n^k\|_{L_2} \leq \|x_* - x_n^*\|_{L_2} + \|x_n^* - x_n^k\|_{L_2}.$$

Для первого слагаемого в правой части верна оценка (24), а для второго – оценка (13). Но тогда справедлива оценка (19). Теорема доказана.

Аналогично теореме 4 доказывается

Теорема 5. Пусть выполнены неравенства (9), (15) и (18). Пусть начальное приближение метода (12) является решением системы (14). Тогда

$$\|x_* - x_n^k\|_{L_2} \leq \frac{T\rho}{2\pi(1-q\|\mathcal{H}\|)(n+1)} + \frac{(q\|\mathcal{H}\|)^k(1+q\|\mathcal{H}\|)}{(1-q\|\mathcal{H}\|)^2} \varepsilon.$$

4. Колебания в системах автоматического регулирования

Пусть динамика системы W автоматического регулирования описывается дифференциальным уравнением $L(p)y = M(p)g(t, y)$.

Здесь, как обычно,

$$p = \frac{d}{dt},$$

$$L(p) = p^l + a_1 p^{l-1} + \dots + a_l,$$

$$M(p) = b_0 p^m + b_1 p^{m-1} + \dots + b_m$$

и $l > m$. Предполагается, что нелинейность $g(t, y)$ непрерывна по совокупности переменных, T -периодична по t и удовлетворяет условию Липшица по y :

$$(27) \quad \|g(t, y_1) - g(t, y_2)\| \leq q \|y_1 - y_2\|.$$

Предположим, что многочлен $L(p)$ не имеет корней вида

$$\frac{2k\pi i}{T}, \quad k = 0, \pm 1, \dots$$

Тогда задача отыскания T -периодических режимов системы W эквивалентна отысканию решений интегрального уравнения

$$(28) \quad y(t) = \int_0^T h(t-s; T) g(s, y(s)) ds,$$

где $h(t; T)$ – импульсно-частотная характеристика линейного звена с передаточной функцией $W(p) = M(p)/L(p)$. Импульсно-частотная характеристика $h(t; T)$ определена на всей оси, T -периодична и $h(t, T) = ((I - e^{TA})^{-1} e^{tA} \gamma, c)$, $0 \leq t < T$, где

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ -a_1 & -a_2 & -a_3 & \dots & -a_l \end{bmatrix}, \quad c = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{l-1} \\ \gamma_l \end{bmatrix}$$

и

$$\begin{aligned} \gamma_1 &= \gamma_2 = \dots = \gamma_{l-m-1} = 0, \\ \gamma_{l-m} &= b_0, \quad \gamma_{l-m+1} + \gamma_{l-m} a_l = b_1; \\ &\dots\dots\dots \\ \gamma_l + \gamma_{l-1} a_1 + \dots + \gamma_{l-m} a_m &= b_m. \end{aligned}$$

Уравнение (28) будем изучать в пространстве L_2 .

Пусть H_n – подпространство пространства L_2 , базис в котором образуют функции

$$1, \cos \frac{2\pi t}{T}, \sin \frac{2\pi t}{T}, \dots, \cos \frac{2\pi n t}{T}, \sin \frac{2\pi n t}{T},$$

а P_n – ортопроектор из L_2 на H_n . Далее, пусть

$$\mathfrak{h} y = \int_0^T h(t-s; T) y(s) ds.$$

Линейный оператор $\mathfrak{h}: L_2 \rightarrow L_2$ вполне непрерывен, причем

$$\gamma = \|\mathfrak{h}\| = \max_{k=0,1,2,\dots} \left| W \left(\frac{2k\pi i}{T} \right) \right|,$$

$$\gamma_n = \|P_n \mathfrak{h}\| = \max_{k=0,1,2,\dots,n} \left| W \left(\frac{2n\pi i}{T} \right) \right|.$$

Предположим, что система W имеет T -периодический режим $y^* = y^*(t)$.

Построение приближений y_n^* к режиму y^* методом гармонического баланса эквивалентно отысканию решений операторного уравнения

$$(29) \quad y_n = P_n \mathfrak{h} g(y_n), \quad y_n \in H_n,$$

где g – оператор суперпозиции: $g(y) = p(t, y(t))$.

Аналогом теоремы 1 является

Теорема 6. Пусть выполнено неравенство

$$(30) \quad \gamma_n q < 1,$$

где q – коэффициент из (27). Тогда уравнение (29) имеет единственное решение

$$y_n^* = y_n^*(t),$$

а последовательные приближения

$$(31) \quad y_n^{k+1} = P_n \mathfrak{h}g(y_n^k), \quad k = 0, 1, \dots$$

сходятся к y_n^* при любом начальном приближении y_n^0 , причем

$$\|y_n^* - y_n^k\|_{L_2} \leq \frac{(\gamma_n q)^k}{1 - \gamma_n q} \|P_n \mathfrak{h}g(y_n^0) - y_n^0\|_{L_2}.$$

Если известна оценка нормы разности $y^* - P_m y^*$, то последнюю оценку можно уточнить.

Теорема 7. Пусть $m < n$ и выполнено неравенство (30). Пусть начальное приближение y_n^0 в методе (31) является решением в H_m уравнения

$$(32) \quad L(p)y(t) = M(p)P_m g(t, y(t)).$$

Пусть

$$(33) \quad \|y^* - P_m y^*\|_{L_2} \leq \varepsilon.$$

Тогда

$$\|y_n^* - y_n^k\| \leq \frac{(\gamma_n q)^k (1 + \gamma_n q)}{(1 - \gamma_n q)(1 - \gamma_m q)} \varepsilon.$$

Следующие два утверждения – аналоги теорем 4 и 5 – дают оценки уклонения последовательных приближений (31) от колебания $y^* = y^*(t)$.

Теорема 8. Пусть верны оценки (30) и

$$(34) \quad \|y^*\|_{W_2^1} \leq \rho.$$

Тогда

$$\|y^* - y_n^k\| \leq \frac{T\rho}{2\pi(1 - \gamma_n q)(n + 1)} + \frac{(\gamma_n q)^k}{1 - \gamma_n q} \|P_n \mathfrak{h}g(y_n^0) - y_n^0\|_{L_2}.$$

Теорема 9. Пусть $m < n$ и выполнены неравенства (30), (33), (34). Пусть начальное приближение y_n^0 в методе (31) принадлежит H_m и является решением уравнения (32). Тогда

$$\|y_* - y_n^k\| \leq \frac{T\rho}{2\pi(1-\gamma_n q)(n+1)} + \frac{(\gamma_n q)^k(1+\gamma_n q)}{(1-\gamma_n q)(1+\gamma_n q)} \varepsilon.$$

Смысл основных теорем 5, 9 состоит в том, что в качестве начального приближения итерационного метода (12), (31) берется решение, полученное методом гармонического баланса. Тогда скорость сходимости итерационного метода будет удовлетворять приведенным оценкам.

5. Заключение

Полученные результаты не только показывают качественную зависимость сходимости приближенного метода от выбора начального приближения, но и дают количественные оценки скорости этой сходимости. Кроме того, нами предложены направления исследования метода гармонического баланса, являющегося наиболее эффективным и практически применимым средством решения задач о вынужденных колебаниях.

Литература

1. БАБИЦКИЙ В.И., КРУПЕНИН В.Л. *Колебания в сильно нелинейных системах*. – М.: Наука, 1985. – 320 с.
2. БОБЫЛЕВ Н.А., КРАСНОСЕЛЬСКИЙ А.М. *О приближенном построении автоколебаний в системах автоматического регулирования* // Докл. АН СССР. – 1983. – Т. 272, №2.
3. БОБЫЛЕВ Н.А., КРАСНОСЕЛЬСКИЙ А.М. *О методе гармонического баланса в задаче об автоколебаниях* // Автоматика и телемеханика. – 1984. – №9.
4. БОБЫЛЕВ Н.А., КРАСНОСЕЛЬСКИЙ А.М., КРАСНОСЕЛЬСКИЙ М.А. *Устойчивость периодических колебаний и возможность их построения методом гармонического баланса* // Автоматика и телемеханика. – 1988. – №7.
5. БОБЫЛЕВ Н.А., БУРМАН Ю.М., КОРОВИН С.К. *Оценка погрешности метода гармонического баланса* // Автоматика и телемеханика. – 1992. – №5.

6. БОГОЛЮБОВ Н.Н., МИТРОПОЛЬСКИЙ Ю.А. *Асимптотические методы в теории нелинейных колебаний*. – М.: Физматгиз, 1958. – 407 с.
7. БРАВЕРМАН Э.М., МЕЕРКОВ С.М., ПЯТНИЦКИЙ Е.С. *Малый параметр в проблеме обоснования метода гармонического баланса (в случае гипотезы фильтра)* // Автоматика и телемеханика. – 1975. – №1. – С. 5–12.
8. ВАЙНИККО Г.М. *Анализ дискретизационных методов*. – Тарту: Изд-во Тартуского ун-та, 1976. – 161 с.
9. КРАСНОСЕЛЬСКИЙ М.А., ВАЙНИККО Г.М., ЗАБРЕЙКО П.П. и др. *Приближенное решение операторных уравнений*. – М.: Наука, 1969. – 457 с.
10. КРАСНОСЕЛЬСКИЙ А.М. *Направленные поправки в методе гармонического баланса* // Докл. АН СССР. – 1987. – Т. 294, №6. – С. 1314–1318.
11. НАУМОВ Б.Н., ЦЫПКИН Я.З. *Частотные критерии абсолютной устойчивости процессов в нелинейных системах автоматического регулирования* // Автоматика и телемеханика. – 1964. – №6.
12. ПОПОВ Е.П., ПАЛЬТОВ И.П. *Приближенные методы исследования нелинейных автоматических систем*. – М.: Физматгиз, 1960. – 729 с.
13. РОЗЕНВАССЕР Е.Н. *Колебания нелинейных систем*. – М.: Наука, 1969. – 576 с.
14. РОЗЕНВАССЕР Е.Н. *Апостериорные оценки применимости метода гармонического баланса* // Автоматика и телемеханика. – 1986. – №3. – С. 44–52.
15. САМОЙЛЕНКО А.М., РОНТО Н.И. *Численно-аналитические методы исследования периодических решений*. – Киев: Вища шк. 1976. – 180 с.
16. СТАРЖИНСКИЙ В.М. *Прикладные методы теории нелинейных колебаний*. – М.: Наука, 1977. – 260 с.
17. BABUSKA I., PRAGER M., VITASEK V. *Numerical processes in differential equations*. – New York: Interscience, 1966. – 351 p.

18. STOKER J.J. *Nonlinear vibrations in mechanical and electrical systems*. – New York: Interscience, 1950. – 273 p.

PROJECTION-ITERATIVE PROCEDURES FOR APPROXIMATION OF FORCED OSCILLATIONS IN NONLINEAR SYSTEMS.

Iham Ismailov, Institute of Control Sciences of RAS, Moscow, Cand.Sc., Senior Research Associate (Moscow, Profsoyuznaya st., 65, (495) 334-79-00, iig07@mail.ru).

Abstracts: Results on problems of approximation of complicated nonlinear systems regimes and special numeric procedures are described. We study some applications of iteration and projective theory to the problems of approximate construction of oscillation regimes for nonlinear systems, particularly, automatic control systems. We introduce a routine for initial approximation selection in the iteration method. This routine uses a priori information about the oscillation regime we search for.

Keywords: optimal control, automatic control, forced oscillations, approximation methods.

Статья представлена к публикации членом редакционной коллегии Б. Т. Поляком

УДК 681.5.015

ББК 32.965.4

ВАРИАЦИОННЫЕ МЕТОДЫ ИДЕНТИФИКАЦИИ ЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ И ПРОБЛЕМА ЛОКАЛЬНЫХ ЭКСТРЕМУМОВ¹

Ломов А. А.²

(ФГБУН Институт математики им. С. Л. Соболева Сибирского
отделения РАН, Новосибирский государственный университет,
Новосибирск)

Рассмотрена проблема большого числа локальных экстремумов, возникающих при использовании «прямых» методов идентификации параметров линейных динамических систем по конечным выборкам наблюдений. Определен новый класс вариационных («непрямых») оценок параметров через свойство проективности слагаемых ядра целевой функции. Построены вариационные целевые функции с числом локальных экстремумов не больше числа элементов в матрицах системы. Получены условия состоятельности вариационных оценок в пределе большого числа наблюдений независимых траекторий конечной длины.

Ключевые слова: идентификация параметров, разностные уравнения, динамические системы.

1. Введение

Вычислительная сложность метода идентификации в значительной степени определяется экстремальными свойствами целевой функции. Известно, что при идентификации по конечным выборкам «прямыми» методами число экстремумов может быть

¹ Работа выполнена при финансовой поддержке РФФИ (грант № 10-01-00035) и при поддержке Сибирского отделения Российской академии наук (междисциплинарный проект № 80).

² Андрей Александрович Ломов, кандидат физико-математических наук, доцент (lomov@math.nsc.ru).

значительным и даже расти с ростом объема выборки наблюдений [26]. Поэтому заслуживают внимания методы идентификации, целевые функции у которых обладают «хорошими» экстремальными свойствами, возможно, за счет увеличения объема вычислений. В статье определен новый класс вариационных («непрямых») методов, интересный тем, что для ряда типичных представителей этого класса удается доказать конечность числа критических точек целевой функции независимо от объема выборки наблюдений.

Опишем структуру статьи. Во введении описываются исследуемые системы, даются основные понятия и обозначения. Затем формулируется проблема большого числа локальных экстремумов при идентификации «прямыми» методами по выборкам конечного объема. Решение проблемы может быть найдено при переходе к вариационной постановке задачи идентификации. В разделе 2 определяется новый класс «непрямых» (вариационных) оценок параметров через свойство проективности слагаемых ядра целевой функции. Показано, что в этот класс входят как известные в литературе оценки (ОР [42, 48], ОРМ [31], ВИ [9, 10], TLS [39], GTLS [49], STLS [45]), так и ряд новых предложенных в статье оценок — в частности, многомерные ОРМ и ОРС. Далее изучаются свойства вариационных оценок (раздел 3). Установлено, что многие задачи идентификации состоятельно решаются вариационными методами с целевыми функциями, число экстремумов которых не превосходит числа элементов в матричных коэффициентах уравнения системы независимо от объема L , N выборки измерений. В этом смысле вариационные постановки задач идентификации обладают «регуляризующим», «сглаживающим» эффектом, и поэтому могут рассматриваться как средство для решения проблемы большого числа локальных экстремумов.

1.1. ИССЛЕДУЕМЫЕ СИСТЕМЫ, ОЦЕНКИ ПАРАМЕТРОВ, ОБОЗНАЧЕНИЯ

Будем рассматривать линейные динамические системы

$$(1) \quad \alpha_p y[k+p] + \dots + \alpha_0 y[k] = \beta_p u[k+p] + \dots + \beta_0 u[k], \\ k \in \overline{1, N-p}, \quad N \geq p+1.$$

Здесь $\alpha_i = \alpha_{i,\theta} \in \mathbb{R}^{r \times r}$, $\beta_i = \beta_{i,\theta} \in \mathbb{R}^{r \times m}$ — матричные коэффициенты, которые зависят от параметра $\theta \in \Theta$, множество $\Theta \subseteq \mathbb{R}^v$ открытое. Параметр θ фиксирован и подлежит идентификации по множеству наблюдений

$$\{\check{y}[1], \check{u}[1], \dots, \check{y}[N], \check{u}[N]\},$$

$$\check{y}[k] = y[k] + \varepsilon_y[k], \quad \check{u}[k] = u[k] + \varepsilon_u[k],$$

где $y[k] \in \mathbb{R}^r$, $u[k] \in \mathbb{R}^m$ — переменные выхода, входа, $\varepsilon_y[k] \in \mathbb{R}^r$, $\varepsilon_u[k] \in \mathbb{R}^m$ — измерительные возмущения. Существенно, что старший матричный коэффициент α_p может быть вырожденным; тогда уравнение (1) описывает широкий класс многомерных (MIMO) систем с обратными связями и запаздываниями.

Определим вектор *процесса*

$$z \doteq (z[1]; \dots; z[N]) \in \mathbb{R}^{N(r+m)}, \quad z[k] \doteq \begin{pmatrix} y[k] \\ u[k] \end{pmatrix} \in \mathbb{R}^{r+m},$$

и вектор возмущений

$$\varepsilon \doteq (\varepsilon[1]; \dots; \varepsilon[N]) \in \mathbb{R}^{N(r+m)}, \quad \varepsilon[k] \doteq \begin{pmatrix} \varepsilon_y[k] \\ \varepsilon_u[k] \end{pmatrix} \in \mathbb{R}^{r+m}.$$

Здесь и далее, следуя [34], запятыми $(*, \dots, *)$ обозначаем вектор-строку, а точками с запятой $(*; \dots; *)$ — вектор-столбец. Аналогичные обозначения используются для клеточных матриц: $(A, B) \doteq (A \ B)$, $(A; B) \doteq \begin{pmatrix} A \\ B \end{pmatrix}$.

Пусть $z_{(i)} \in \mathbb{R}^{N(r+m)}$, $i \in \overline{1, L}$, — некоторый «полный» с точки зрения идентификации набор процессов в системе, $\varepsilon_{(i)}$

— случайные возмущения, ограниченные «естественными» предположениями о распределении (точные определения даны ниже). Нас будут интересовать следующие свойства оценки как функции наблюдений:

1. Единственность (по смыслу это свойство лучше назвать точность, или состоятельность в детерминированном смысле):

$$\hat{\theta}(z_{(1)}, \dots, z_{(L)}) = \theta.$$

2. Устойчивость:

$$\hat{\theta}(z_{(1)} + \varepsilon_{(1)}, \dots, z_{(L)} + \varepsilon_{(L)}) \xrightarrow[\substack{\|\varepsilon_{(i)}\| \rightarrow 0, \\ i \in \overline{1, L}}]{\rightarrow} \theta.$$

3. Состоятельность (в статистическом смысле):

$$\hat{\theta}(z_{(1)} + \varepsilon_{(1)}, \dots, z_{(L)} + \varepsilon_{(L)}) \xrightarrow{L \rightarrow \infty} \theta.$$

Если число процессов L ограничено, то состоятельность рассматривается в предельном случае больших длин процессов N :

$$\hat{\theta}(z + \varepsilon) \xrightarrow{N \rightarrow \infty} \theta, \quad z + \varepsilon = \begin{pmatrix} z^{[1]} + \varepsilon^{[1]} \\ \vdots \\ z^{[N]} + \varepsilon^{[N]} \end{pmatrix}.$$

Последнее предполагает, что система не изменяет своих параметров в течение длительного отрезка времени Nh , где h — время дискретизации.

В общем случае оценки $\hat{\theta}(\check{z}_{(1)}, \dots, \check{z}_{(L)})$ вычисляются не по явным формулам, а минимизацией той или иной целевой функции, зависящей от наблюдений $\check{z}_{(i)}$. Под методом идентификации будем понимать целевую функцию вместе с вычислительным алгоритмом минимизации.

1.2. ПРЯМЫЕ МЕТОДЫ ИДЕНТИФИКАЦИИ И ПРОБЛЕМА ЛОКАЛЬНЫХ ЭКСТРЕМУМОВ

С вычислительной точки зрения важны экстремальные свойства целевой функции — в частности, число критических точек

в зависимости от множества наблюдений. Этим в значительной степени определяется устойчивость алгоритмов идентификации.

В статье рассмотрим экстремальные свойства двух больших групп методов идентификации во временной области:

- «прямые» [28, раздел 5.3] (или «наивные» [38], «разомкнутые» [24]);
- «непрямые» (или «замкнутые» [24], «вариационные» [10]).

По-видимому, первым, кто попытался выделить эти две группы, был К.Б. Норкин [24], затем А.О. Егоршин [9] и Л. Глезер [38]. Данные ими определения были скорее интуитивными, мотивированными желанием найти «наиболее правильные»³ постановки задач идентификации. Методы первой группы («прямые») в вычислительном отношении более просты, но с ними связан ряд проблем, описанных ниже. Этим вызвано внимание к более сложным «непрямым» методам.

В статье далее делается попытка уточнить смысл понятия «правильной постановки» задачи идентификации с точки зрения экстремальных свойств целевых функций. В рамках этой программы предлагается (по-видимому, впервые) точное определение для второй группы («непрямых») методов, и на основании исследования экстремальных свойств устанавливаются частные случаи, в которых «непрямые» методы имеют преимущества по сравнению с методами первой группы.

Кратко опишем особенности «прямых» методов. Для этой группы, насколько известно, точного определения пока не дано. Имеется следующее описание, данное А. А. Красовским [28, раздел 5.3]: «Прямыми методами параметрической идентификации» называются те, в которых «неизвестные параметры ... модели определяются на основе того или иного способа (обычно приближенного) решения системы уравнений, получающихся путем [прямой] подстановки (выделено мной — А. Л.) в оператор модели последовательности значений входных и выходных величин

³ Приблизительная формулировка.

реального объекта». Согласно этому нестрогому определению, к «прямым» относятся методы, основанные на минимизации нормы «ошибки прогноза» (PEM, «prediction error methods») и их многочисленные модификации, включающие в себя рекуррентные методы наименьших квадратов по невязке уравнения и их варианты типа алгоритмов стохастической аппроксимации и алгоритмов Качмажа [11, 23, 25, 29, 28, 35, 52], а также методы типа инструментальных переменных [51, глава 8]. Прямые методы идентификации повсеместно употребляются при решении задач адаптивного управления [1, 32]. Для линейных моделей со структурой общего вида

$$(2) \quad A(s)\check{y}[k] = \frac{B(s)}{F(s)}\check{u}[k] + \frac{C(s)}{D(s)}\varepsilon[k]$$

(где s — оператор сдвига, и \check{y} , \check{u} — наблюдаемые переменные) целевые функции прямых методов строятся исходя из нормы оценки невязки уравнения $\frac{C(s)}{D(s)}\varepsilon[k]$ или ошибки прогноза $\frac{C(s)}{A(s)D(s)}\varepsilon[k]$ [23].

Заметим, что идея прямой подстановки измерений в уравнение системы восходит к ставшим уже классическими работам А.Н. Колмогорова [12] и Н. Винера [54]. В этой связи уместно назвать и работу К. Гаусса по методу наименьших квадратов [5].

В применении к идентификации динамических систем прямые методы при сравнительной простоте вычислительных алгоритмов имеют ограничения: 1) система должна быть управляемой; 2) длина N интервала наблюдения должна быть большой в сравнении с длительностью переходных процессов в системе [35].

Кроме того, И. И. Перельманом было показано [26], что для прямых методов при конечных N существует проблема большого числа локальных экстремумов целевой функции. Наличие локальных экстремумов отмечалось также в [33, 50]. Приведем простейший пример [26]:

$$y[k] = ay[k - 1] + u[k], \quad \check{y}[k] = y[k] + \varepsilon[k], \quad \check{u}[k] = u[k].$$

Уравнение для ошибки прогноза имеет вид

$$\begin{aligned}\varepsilon[k] &= \check{y}[k] - \check{u}[k] - a\check{y}[k-1] + a\varepsilon[k-1] \doteq \\ &\doteq \check{\omega}[k] + a\varepsilon[k-1],\end{aligned}$$

где $\check{\omega}[k]$ — измерения. Принимая в качестве целевой функции квадрат нормы ошибки прогноза на конце траектории, получаем

$$\varepsilon^2[N] = (a^{N-1}\varepsilon[1] + a^{N-2}\check{\omega}[1] + \dots + \check{\omega}[N])^2 \doteq J(a).$$

Локальные экстремумы определяются из уравнения $\frac{\partial J}{\partial a} = 0$, которое может иметь до $(2N - 3)$ действительных корней. Очевидно, число корней растет с ростом N , и все они должны рассматриваться как претенденты на точки глобального минимума. Этот факт, как пишет И. И. Перельман, «приводит к обоснованным сомнениям относительно практической разрешимости» поставленной простейшей задачи идентификации. Трудности только усугубляются, если рассмотреть системы с контурами обратной связи, т. е. системы с несколькими уравнениями.

В пределе $N \rightarrow \infty$ вероятность большого числа экстремальных точек стремится к нулю [3]. Однако этот факт не снимает отмеченных трудностей идентификации «прямыми» методами по конечному множеству наблюдений.

2. Вариационные (непрямые) оценки

Решение проблемы локальных экстремумов может быть найдено при переходе к вариационной постановке задачи идентификации. В одном частном случае это было показано М. Левиным [42] и М. Аоки, П. Ю [31], которые рассматривали системы (1) из одного уравнения ($r = 1$) с простой параметризацией. Для идентификации параметров они применяли метод ортогональной регрессии [48], целевая функция которого имеет вид отношения квадратичных форм (22), и как следствие, число экстремумов ограничивается сверху размерностью вектора параметров (см. раздел 2.1).

Термины «вариационный метод», «вариационная постановка задачи идентификации» были предложены А.О. Егоршиным [10], который сконструировал один из наиболее эффективных с точки зрения приложений «замкнутый» метод идентификации [9]. Поясним примером. Пусть дано наблюдение $\check{z} \in \mathbb{R}^{N(r+m)}$. Оценка вектора параметров системы (1) может быть получена минимизацией по θ функции

$$(3) \quad J(\theta) = J(\theta, \check{z}) = \|\check{z} - z_{\text{opt}}(\theta)\|^2,$$

где $z_{\text{opt}}(\theta)$ — проекция вектора \check{z} на множество всех процессов системы (1). Величина $J(\theta)$ играет роль квадрата расстояния от «точки» \check{z} до множества процессов (1). Функция $\hat{\theta}(\check{z}) = \arg \min_{\theta} J(\theta, \check{z})$ является примером вариационной (непрямой) оценки. Эффективные алгоритмы ее вычисления были предложены в [9, 10].

В этом разделе определим новый класс «вариационных»⁴ оценок. Он будет включать в себя большинство известных в литературе непрямых оценок: оценки ортогональной регрессии [48], модифицированные оценки М. Аоки, П. Ю [31], оценки по методу вариационной идентификации А.О. Егоршина [9, 10], оценки по модифицированному методу Прони [37, 47], оценки STLS [45] и ряд других, в том числе новых оценок. Общим свойством для всего класса будет *независимость следа матрицы ядра целевой функции от оцениваемого параметра*. В частности, это свойство позволяет доказывать состоятельность в пределе $L \rightarrow \infty$ по одной и той же схеме, предложенной в [15]. Одним из полезных побочных результатов будет также описание различных представителей класса вариационных методов в единых обозначениях (см. разделы 2.1–2.5).

Перейдем к определениям. Введем матричную запись урав-

⁴ *Название не является устоявшимся.*

нения (1):

$$(4) \quad G_\theta z = 0,$$

$$G_\theta = \begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_p & & 0 \\ & \gamma_0 & \gamma_1 & \dots & \gamma_p & \\ & & \ddots & \ddots & & \ddots \\ 0 & & & \gamma_0 & \gamma_1 & \dots & \gamma_p \end{pmatrix}$$

$$\in \mathbb{R}^{(N-p)r \times N(r+m)},$$

$$\gamma_i \doteq \gamma_{i,\theta} = (\alpha_{i,\theta}, -\beta_{i,\theta}) \in \mathbb{R}^{r \times (r+m)}.$$

Тогда $z_{\text{opt}}(\theta) = (I - \Pi_\theta) \tilde{z}$, где $\Pi_\theta \doteq G_\theta^\top (G_\theta G_\theta^\top)^{-1} G_\theta$ есть матрица проектора ($\Pi_\theta^2 = \Pi_\theta$) на ортогональное дополнение к линейному многообразию решений системы (1). С учетом этого обозначения целевая функция (3) принимает вид

$$(5) \quad J(\theta) = \tilde{z}^\top \Pi_\theta \tilde{z}.$$

Матрицу Π_θ будем называть *ядром* целевой функции $J(\theta)$.

Пусть θ_* — некоторое неизвестное «истинное» значение параметра системы (1).

Определение 1. Вариационной оценкой $\hat{\theta}$ параметра θ_* по наблюдению $\tilde{z} = z_* + \varepsilon$, где z_* — истинный процесс, ε — возмущение, называем результат минимизации квадратичной целевой функции

$$(6) \quad J(\theta, \tilde{z}) \doteq \tilde{z}^\top (\Pi_{1,\theta} + \dots + \Pi_{M,\theta}) \tilde{z},$$

$$(7) \quad \hat{\theta} = \hat{\theta}(\tilde{z}) = \arg \min_{\theta} J(\theta, \tilde{z}),$$

где все матрицы $\Pi_j = \Pi_{j,\theta} \geq 0$ являются неотрицательно определенными, симметричными, проективными ($\Pi_j^2 = \Pi_j$) и такими, что верно соотношение $(G_{\theta_*} z = 0) \Rightarrow (\Pi_{j,\theta_*} z = 0)$. Целевую функцию $J(\theta, \tilde{z})$ (6) также называем вариационной.

Можно сказать, что целевые функции, ядра которых есть сумма проекторов, являются вариационными (или *проективными*).

Приведем еще ряд определений, расширяющих понятие вариационной оценки. Пусть возмущения $\varepsilon_{(1)}, \dots, \varepsilon_{(L)} \in \mathbb{R}^{N(r+m)}$

являются независимыми одинаково распределенными случайными векторами с нулевыми математическими ожиданиями и скалярной матрицей вторых моментов:

$$(8) \quad \mathbf{M} \varepsilon_{(i)} = 0, \quad \mathbf{M} \varepsilon_{(i)} \varepsilon_{(i)}^\top = \sigma^2 I.$$

Тогда имеет смысл следующее определение.

Определение 2. Вариационной оценкой $\hat{\theta}_L$ параметра θ_* по множеству наблюдений $\{\check{z}_{(i)} = z_{*(i)} + \varepsilon_{(i)}, i \in \overline{1, L}\}$, где $z_{*(i)}$ — истинные процессы, $\varepsilon_{(i)}$ — возмущения (8), называем результат минимизации квадратичной целевой функции

$$(9) \quad J_L(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}) \doteq L^{-1} \sum_{i=1}^L \check{z}_{(i)}^\top (\Pi_{1,\theta} + \dots + \Pi_{M,\theta}) \check{z}_{(i)},$$

$$(10) \quad \hat{\theta}_L = \hat{\theta}(\check{z}_{(1)}, \dots, \check{z}_{(L)}) = \arg \min_{\theta} J_L(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}),$$

где все матрицы $\Pi_j = \Pi_{j,\theta} \geq 0$ являются неотрицательно определенными, симметричными, проективными ($\Pi_j^2 = \Pi_j$) и такими, что верно соотношение $(G_{\theta_*} z = 0) \Rightarrow (\Pi_{j,\theta_*} z = 0)$.

Пусть теперь возмущения $\varepsilon_{(1)}, \dots, \varepsilon_{(L)} \in \mathbb{R}^{N(r+m)}$ являются независимыми одинаково распределенными случайными векторами со свойствами

$$(11) \quad \mathbf{M} \varepsilon_{(i)} = 0, \quad \mathbf{M} \varepsilon_{(i)} \varepsilon_{(i)}^\top = \Sigma \Sigma^\top,$$

где $\Sigma \in \mathbb{R}^{N(r+m) \times n}$ — матрица с линейно независимыми столбцами.

Определение 3. Вариационной оценкой $\hat{\theta}_L$ параметра θ_* по множеству наблюдений $\{\check{z}_{(i)} = z_{*(i)} + \varepsilon_{(i)}, i \in \overline{1, L}\}$, где $z_{*(i)}$ — истинные процессы, $\varepsilon_{(i)}$ — возмущения (11), называем результат минимизации целевой функции

$$(12) \quad J_{\Sigma,L}(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}) \doteq \\ \doteq L^{-1} \sum_{i=1}^L \check{z}_{(i)}^\top (W_{1,\theta} + \dots + W_{M,\theta}) \check{z}_{(i)},$$

$$(13) \quad \hat{\theta}_L = \hat{\theta}(\check{z}_{(1)}, \dots, \check{z}_{(L)}) = \arg \min_{\theta} J_{\Sigma,L}(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}),$$

где все матрицы $W_{j,\theta} \geq 0$ неотрицательно определены, симметричны и таковы, что верно соотношение $(G_{\theta_*} z = 0) \Rightarrow (W_{j,\theta_*} z = 0)$, а произведения $\Sigma^\top W_j \Sigma = \Pi_j$ являются проективными матрицами со свойством $\Pi_j^2 = \Pi_j$.

В следующем заключительном (наиболее общем) определении уже не будем опираться на проективность матриц, а возьмем за основу более общее свойство независимости следа матрицы от оцениваемого параметра. Полученные таким образом оценки уже не обязательно имеют ядро целевой функции в виде суммы проекторов, но доказательство их состоятельности осуществляется по той же схеме, что и для оценок из определений 2–3 (для оценок с целевой функцией (9) см. в [15]).

Определение 4. *Обобщенной вариационной оценкой $\hat{\theta}_L$ параметра θ_* по множеству наблюдений $\{\check{z}_{(i)} = z_{*(i)} + \varepsilon_{(i)}, i \in \overline{1, L}\}$, где $z_{*(i)}$ — истинные процессы, $\varepsilon_{(i)}$ — возмущения (11), называем результат минимизации целевой функции*

$$J_{\Sigma, L}(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}) \doteq L^{-1} \sum_{i=1}^L \check{z}_{(i)}^{\top} U_{\theta} \check{z}_{(i)},$$

$$\hat{\theta}_L = \hat{\theta}(\check{z}_{(1)}, \dots, \check{z}_{(L)}) = \arg \min_{\theta} J_{\Sigma, L}(\theta, \check{z}_{(1)}, \dots, \check{z}_{(L)}),$$

где матрица $U_{\theta} \geq 0$ неотрицательно определена, симметрична и такова, что верно соотношение $(G_{\theta_*} z = 0) \Rightarrow (U_{\theta_*} z = 0)$, а след произведения $\Sigma^{\top} U_{\theta} \Sigma$ не зависит от θ :

$$\text{Sp } \Sigma^{\top} U_{\theta} \Sigma \equiv \text{const}(\theta).$$

В итоге, определяющим свойством для введенного в этом разделе класса вариационных оценок является независимость следа матрицы ядра целевой функции от оцениваемого параметра. Заметим, что все известные нам примеры вариационных оценок ограничены оценками из определений 2–3, т. е. с проективными ядрами.

Определения 2 и 3 по существу не вносят новых идей в понятие вариационной оценки по сравнению с определением 1. Чтобы продемонстрировать конструктивность данных определений, покажем, что ряд известных в литературе оценок являются частными случаями вариационных из определения 1 ($L = 1$), а также построим новые оценки.

Введем обозначения. Из матриц $\gamma_i \in \mathbb{R}^{r \times (r+m)}$ (4) составим матрицу

$$(14) \quad \gamma_\theta \doteq (\gamma_0 \ \gamma_1 \ \dots \ \gamma_p) \in \mathbb{R}^{r \times (r+m)(p+1)}.$$

Пусть $\text{vect } \gamma$ – вектор, полученный последовательным выстраиванием транспонированных строк матрицы γ . Будем предполагать, что матрица γ_θ (14) зависит от вектора параметров θ аффинным образом:

$$(15) \quad \text{vect } \gamma_\theta = d + D\theta \doteq D\vartheta,$$

где матрица $D \in \mathbb{R}^{r(r+m)(p+1) \times v}$ и вектор $d \in \mathbb{R}^{r(r+m)(p+1)}$ заданы, столбцы составной матрицы $D \doteq (d, D)$ линейно независимы, $\vartheta \doteq \begin{pmatrix} 1 \\ \theta \end{pmatrix}$.

Система уравнений (4) всегда может быть записана в виде

$$(16) \quad G_\theta z \equiv V \text{vect } \gamma_\theta \equiv VD\vartheta = 0,$$

где знак « \equiv » обозначает равенство для всех $z \in \mathbb{R}^{N(r+m)}$, $V = V(z)$ – клеточно-ганкелевая матрица из элементов вектора z :

$$(17) \quad V(z) \doteq \begin{pmatrix} I_r \otimes v_{(1)}^\top \\ \vdots \\ I_r \otimes v_{(N-p)}^\top \end{pmatrix}, \quad v_{(i)}^\top \doteq (z[i]^\top \dots z[i+p]^\top).$$

Символ \otimes обозначает кронекерово произведение матриц, I_r – единичная матрица $r \times r$.

Дополнительно предположим:

(А) для всех $\theta \in \Theta$ каноническая форма многочленной матрицы

$$\gamma_\theta(s) \doteq \gamma_{p,\theta} s^p + \gamma_{p-1,\theta} s^{p-1} + \dots + \gamma_{0,\theta}$$

не имеет тождественно нулевых строк; это условие при $N \geq 2(p+1)$, $p \geq 0$ равносильно линейной независимости строк числовой матрицы G_θ (4) [14, предложение 4.1];

(Б) степень определителя многочленной матрицы знаменателя $\alpha_\theta(s)$ не зависит от параметра θ :

$$\deg \det \alpha_\theta(s) \equiv \text{const},$$

$$\alpha_\theta(s) \doteq \alpha_{p,\theta} s^p + \alpha_{p-1,\theta} s^{p-1} + \dots + \alpha_{0,\theta};$$

(В) степень каждой строки многочленной матрицы знаменателя $\alpha_\theta(s)$ не меньше степени соответствующей строки многочленной матрицы числителя

$$\beta_\theta(s) \doteq \beta_{p,\theta} s^p + \beta_{p-1,\theta} s^{p-1} + \dots + \beta_{0,\theta}$$

(условие причинности); степенью строки многочленной матрицы называется наибольшая степень среди образующих эту строку многочленов.

Заметим, что на систему (4) не налагаются условия управляемости и устойчивости. Последнее возможно благодаря конечности N .

Ввиду равенства $G_\theta z = 0$ условие $\Pi_{j,\theta} z = 0$ из определения 1 позволяет выразить проекторы $\Pi_{j,\theta}$ через линейные комбинации строк матрицы G_θ :

$$(18) \quad \begin{aligned} \Pi_{j,\theta} &= \Phi_{j,\theta}^\top \left(\Phi_{j,\theta} \Phi_{j,\theta}^\top \right)^{-1} \Phi_{j,\theta} \doteq \Phi_{j,\theta}^\top C_{j,\theta} \Phi_{j,\theta}, \\ \Phi_{j,\theta} &= P_j G_\theta, \quad P_j \in \mathbb{R}^{n_j \times (N-p)r}, \end{aligned}$$

где строки матриц P_j (составленные из коэффициентов линейных комбинаций) линейно независимы. Последнее вместе с условием (А) гарантирует строгую положительную определенность матриц $\Phi_{j,\theta} \Phi_{j,\theta}^\top$ и $C_{j,\theta} \doteq \left(\Phi_{j,\theta} \Phi_{j,\theta}^\top \right)^{-1}$.

Проективная целевая функция (6) с учетом параметризации (15) и тождества (16) принимает вид

$$(19) \quad \begin{aligned} J(\theta, \tilde{z}) &\doteq \tilde{z}^\top (\Pi_{1,\theta} + \dots + \Pi_{M,\theta}) \tilde{z} = \\ &= \vartheta^\top D^\top V(\tilde{z})^\top \left(\sum_{j=1}^M P_j^\top C_{j,\theta} P_j \right) V(\tilde{z}) D \vartheta. \end{aligned}$$

Аналогичное выражение получается для целевой функции $J_{\Sigma,L}$ (12): по лемме 2 (из приложения)

$$W_{j,\theta} = G_\theta^\top P_j^\top \underbrace{\left(P_j G_\theta \Sigma \Sigma^\top G_\theta^\top P_j^\top \right)^{-1}}_{C_{\Sigma,j,\theta}} P_j G_\theta,$$

тогда по аналогии с (19) имеем
(20)

$$J_{\Sigma,L} = L^{-1} \sum_{i=1}^L \vartheta^\top D^\top V(\check{z}_{(i)})^\top \left(\sum_{j=1}^M P_j^\top C_{\Sigma,j,\theta} P_j \right) V(\check{z}_{(i)}) D \vartheta.$$

Выражения (19), (20) служат основой для построения вычислительных алгоритмов минимизации (ниже).

2.1. ОЦЕНКИ ОРТОГОНАЛЬНОЙ РЕГРЕССИИ (ОР)

Для идентификации параметров динамических систем (1) М. Левин [42] предложил использовать метод ортогональной регрессии [48]. Покажем, что возникающая здесь целевая функция будет проективной из определения 1. В принятых выше обозначениях оценка М. Левина получается заменой матрицы G_θ системы (1), (4) на матрицу

$$\begin{aligned} G_{\text{OR},\theta} &= \begin{pmatrix} \gamma_0 & \cdots & \gamma_p & & & & 0 \\ & & & \gamma_0 & \cdots & \gamma_p & \\ & & & & & & \ddots \\ 0 & & & & & & \gamma_0 & \cdots & \gamma_p \end{pmatrix} = \\ &= I_M \otimes (\gamma_0 \ \gamma_1 \ \cdots \ \gamma_p). \end{aligned}$$

Матрица $G_{\text{OR},\theta}$ получена из G_θ вычеркиванием части строк (для простоты обозначений положим, что N кратно $p+1$). М. Левин рассматривал случай $r=1$. Ядро целевой функции (6) принимает вид

$$\begin{aligned} (21) \quad \Pi_{1,\theta} + \dots + \Pi_{M,\theta} &= \\ &= \begin{pmatrix} \pi_\theta & & & 0 \\ & 0 & & \\ & & \ddots & \\ 0 & & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & & 0 \\ & \pi_\theta & & \\ & & \ddots & \\ 0 & & & 0 \end{pmatrix} + \dots = \end{aligned}$$

$$\begin{aligned}
 &= \begin{pmatrix} \pi_\theta & & & 0 \\ & \pi_\theta & & \\ & & \ddots & \\ 0 & & & \pi_\theta \end{pmatrix} = \begin{pmatrix} \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} & & & 0 \\ & \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} & & \\ & & \ddots & \\ 0 & & & \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} \end{pmatrix} = \\
 &= G_{\text{OR},\theta}^\top \left(G_{\text{OR},\theta} G_{\text{OR},\theta}^\top \right)^{-1} G_{\text{OR},\theta} \doteq \Pi_\theta.
 \end{aligned}$$

Несложно убедиться, что каждое слагаемое в сумме (21) является проектором ввиду равенств $\pi_\theta^2 = \left(\frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} \right)^2 = \pi_\theta$. Также проектором является вся сумма (21): $\Pi_\theta^2 = \Pi_\theta$. Поэтому оценка М. Левина принадлежит классу вариационных.

В оценке М. Левина каждая матрица P_j из (19) является строкой из нулей и одной единицы на месте с номером $(j-1)(p+1)+1$, $j \in \overline{1, M}$, $M = \frac{N}{p+1}$:

$$P_j = \underbrace{\left(\underbrace{0 \dots 0}_{(j-1)(p+1)} \quad 1 \quad 0 \dots 0 \right)}_{N-p}.$$

Запись (19) принимает вид:

$$\begin{aligned}
 J(\theta, \check{z}) &= \frac{\vartheta^\top D^\top V(\check{z})^\top \left(\sum_{j=1}^M P_j^\top P_j \right) V(\check{z}) D \vartheta}{\gamma_\theta^\top \gamma_\theta} = \\
 (22) \quad &= \frac{\vartheta^\top D^\top V_{\text{OR}}(\check{z})^\top V_{\text{OR}}(\check{z}) D \vartheta}{\vartheta^\top D^\top D \vartheta},
 \end{aligned}$$

где матрица $V_{\text{OR}}(\check{z})$ получена прореживанием ганкелевой матрицы $V(\check{z})$ (17) — удаляются строки, содержащие элементы, имеющиеся в других строках.

Устойчивость оценок с целевой функцией вида (22) исследовали Г. Голуб и Ч. Ван Лоан [39], назвав их оценками TLS (Total Least Squares estimates). Тожественное совпадение оценок TLS Г. Голуба и Ч. Ван Лоана с классическими оценками ортогональной регрессии [48] было отмечено в [36].

2.2. МОДИФИЦИРОВАННЫЕ ОЦЕНКИ ОРТОГОНАЛЬНОЙ РЕГРЕССИИ (ОРМ)

М. Аоки и П. Ю предложили модифицированный вариант (ОРМ) оценки М. Левина, имеющий локально меньший разброс [31]. В наших обозначениях ядро целевой функции модифицированной оценки принимает вид суммы проекторов:

$$\begin{aligned}
 (23) \quad & \Pi_{1,\theta} + \dots + \Pi_{N-p,\theta} = \\
 & = \begin{pmatrix} \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} & & & 0 \\ & 0 & & \\ & & \ddots & \\ 0 & & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & & 0 \\ & \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} & & \\ & & \ddots & \\ 0 & & & 0 \end{pmatrix} + \dots \\
 & \dots + \begin{pmatrix} 0 & & & 0 \\ & 0 & & \\ & & \ddots & \\ 0 & & & \frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top} \end{pmatrix}.
 \end{aligned}$$

Поэтому модифицированная оценка принадлежит классу вариационных. Нули на диагонали есть числа, а клетки $\frac{\gamma_\theta^\top \gamma_\theta}{\gamma_\theta \gamma_\theta^\top}$ имеют размерность $(p+1) \times (p+1)$. Данная сумма уже не является проектором. Число слагаемых $N-p$ здесь больше, чем в методе ОР.

Каждая матрица P_j , $j \in \overline{1, N-p}$ в оценке ОРМ является j -й строкой единичной матрицы I_{N-p} :

$$(24) \quad P_j = \underbrace{(0 \dots 0 \ 1 \ 0 \dots 0)}_{N-p}.$$

Выражение (19) принимает вид

$$(25) \quad J(\theta, \check{z}) = \frac{\vartheta^\top D^\top V(\check{z})^\top V(\check{z}) D \vartheta}{\vartheta^\top D^\top D \vartheta}.$$

Введенные обозначения позволяют достаточно просто построить оценку ОРМ для случая нескольких уравнений $r > 1$.

Такие многомерные оценки в литературе ранее не встречались. Ядро целевой функции:

$$\begin{aligned} & \Pi_{1,\theta} + \dots + \Pi_{N-p,\theta} = \\ & = \begin{pmatrix} \pi_\theta & & 0 \\ & 0 & \\ & & \ddots \\ 0 & & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & 0 \\ & \pi_\theta & \\ & & \ddots \\ 0 & & & 0 \end{pmatrix} + \dots \\ & \dots + \begin{pmatrix} 0 & & 0 \\ & 0 & \\ & & \ddots \\ 0 & & & \pi_\theta \end{pmatrix}, \end{aligned}$$

$$\pi_\theta \doteq \gamma_\theta^\top \left(\gamma_\theta \gamma_\theta^\top \right)^{-1} \gamma_\theta \in \mathbb{R}^{(r+m)(p+1) \times (r+m)(p+1)},$$

$$0 \in \mathbb{R}^{(r+m) \times (r+m)}, \quad \Pi_{j,\theta} \in \mathbb{R}^{(r+m)(N-p) \times (r+m)(N-p)}.$$

Матрицы P_j (24) клеточные:

$$(26) \quad P_j = \underbrace{\left(\underbrace{0 \dots 0}_{(j-1)} \ 1 \ 0 \dots 0 \right)}_{N-p} \otimes I_{r+m}.$$

Целевая функция (19) записывается следующим образом:

$$\begin{aligned} (27) \quad J(\theta, \tilde{z}) &= \vartheta^\top \mathbf{D}^\top \left[\sum_{j=1}^{N-p} V_j(\tilde{z})^\top \left(\gamma_\theta \gamma_\theta^\top \right)^{-1} V_j(\tilde{z}) \right] \mathbf{D} \vartheta = \\ &= \vartheta^\top \mathbf{D}^\top V(\tilde{z})^\top C_{\mathbf{M},\theta} V(\tilde{z}) \mathbf{D} \vartheta, \\ C_{\mathbf{M},\theta} &\doteq I_{N-p} \otimes \left(\gamma_\theta \gamma_\theta^\top \right)^{-1}. \end{aligned}$$

2.3. ОЦЕНКИ ПО МЕТОДУ ВАРИАЦИОННОЙ ИДЕНТИФИКАЦИИ (ВИ) И ПО МЕТОДУ GTLS

Более сложным в вычислительном отношении является метод вариационной идентификации А.О. Егоршина [9]; его многомерный вариант при $r > 1$ изложен в [10]. С точки зрения

разброса оценок при случайных возмущениях этот метод лучше ОР и ОРМ [19]. Целевая функция метода ВИ при $L = 1$ есть (9) с ядром

$$\Pi_{1,\theta} = \dots = \Pi_{M,\theta} \doteq \Pi_\theta = G_\theta^\top \left(G_\theta G_\theta^\top \right)^{-1} G_\theta.$$

Формально $M = 1$,

$$(28) \quad \begin{aligned} P_1 &= I, \\ C_{1,\theta} &\doteq C_\theta = \left(G_\theta G_\theta^\top \right)^{-1}. \end{aligned}$$

Тогда (19) принимает вид

$$(29) \quad J(\theta, \tilde{z}) = \vartheta^\top D^\top V(\tilde{z})^\top C_\theta V(\tilde{z}) D \vartheta.$$

Следует упомянуть также метод идентификации GTLS (Global Total Least Squares), предложенный позже Б.Роордой и К.Хейджем [49]. По целевой функции этот метод совпадает с методом ВИ, поэтому также принадлежит классу вариационных.

2.4. ОЦЕНКИ STLS

Оценки STLS (Structured TLS) были предложены Б.Де Муром [45] как модификация оценок ОР (TLS), учитывающая специальную структуру матрицы наблюдений $V(z)$ (17). В случае $L = 1$, $M = 1$ эта структура клеточно-ганкелевая. Оценки ОР, модифицированные ОР, оценки метода ВИ являются частными случаями STLS. Нас будет интересовать вопрос, совпадает ли класс оценок STLS с классом вариационных оценок из определения 3.

Задача STLS может быть сформулирована следующим образом [40]. Пусть матрица $V(z)$ аффинно зависит от элементов вектора $z \in \mathbb{R}^n$:

$$(30) \quad \begin{aligned} z &= (z_1; \dots; z_n), \\ V(z) &= V_0 + V_1 z_1 + \dots + V_n z_n, \end{aligned}$$

где матрицы V_0, \dots, V_n заранее известны и фиксированы. Пусть $\tilde{z} \in \mathbb{R}^n$ есть вектор наблюдений с матрицей ковариации RR^\top .

Найти минимум

$$(31) \quad \min_{z \in \mathbb{R}^n, \theta \in \mathbb{R}^v} \|R^{-1}(\tilde{z} - z)\|^2 \quad \text{при условии} \quad V(z)g(\theta) = 0,$$

где $g(\theta)$ — матрица, аффинно зависящая от $\theta \in \mathbb{R}^v$:

$$(32) \quad g(\theta) = g_0 + g_1\theta_1 + \dots + g_v\theta_v.$$

С начала 1990-х гг. задачи STLS активно исследуются за рубежом. Широкий круг приложений обсуждался в [45]. В [41] установлена тождественность между оценками STLS и оценками по методу CTLS (Constrained TLS) Т.Абатзоглу и Дж.Менделя [30]. Что касается свойств оценок STLS, то авторами [40] доказана состоятельность многомерных оценок STLS ($g(\theta) \in \mathbb{R}^{r \times v}$) в предельном случае $N \rightarrow \infty$ для клеточно-ганкелевых матриц $V(z)$. Ниже мы установим состоятельность вариационных оценок в случае $L \rightarrow \infty$, и доказав равносильность классов STLS и вариационных, распространим тем самым новое утверждение о состоятельности и на оценки STLS.

Как было отмечено, при $L = 1$, $M = 1$ вариационная оценка (7) является оценкой STLS. Покажем, что в общем случае это отношение сохраняется, более того, имеет место совпадение обоих классов оценок. Этот новый результат устанавливает связь между свойствами оценок вариационных и STLS, исследованных независимо разными авторами [7, 9, 15, 22, 31, 37, 40, 45, 46].

Теорема 1. *Оценка STLS (30)–(32) является вариационной оценкой (13). И наоборот, вариационная оценка (13) всегда может быть получена как решение некоторой задачи STLS (30)–(32).*

Доказательство. См. приложение.

Как следствие этой теоремы, оценки STLS допускают новые проективные формы записи (12), (20) для целевой функции (31), и все полученные ниже результаты о свойствах вариационных оценок — условия единственности, устойчивости, состоятельности при $L \rightarrow \infty$ и конечных N (теоремы 2–6 далее) — распространяются и на оценки STLS. Необходимые переобозначения приведены в приложении в доказательстве теоремы 1.

2.5. СКАЛЯРНЫЕ ОЦЕНКИ ОРТОГОНАЛЬНОЙ РЕГРЕССИИ ДЛЯ СИСТЕМ ИЗ НЕСКОЛЬКИХ УРАВНЕНИЙ (ОРС)

Введем еще один тип вариационных оценок параметров системы (1) из нескольких уравнений ($r > 1$), ранее неизвестный в литературе. По свойствам целевой функции эти оценки наиболее близки к оценкам наиболее простого скалярного ($r = 1$) метода ОРМ. Назовем эти оценки ОРС, обозначим θ_S . Идея состоит в том, чтобы по одному измерению \tilde{z} идентифицировать параметры уравнения с матрицей γ_θ (14) из r строк так, как бы идентифицировалась каждая из этих строк по отдельности скалярным методом ОРМ из раздела 2.2. Существенно, что в оценках ОРС учитывается зависимость коэффициентов разных уравнений от общего параметра θ .

Введем обозначения. В случае $r > 1$ матрица $V(z)$ (17) перестановкой строк приводится к виду

$$V(z) \sim \begin{pmatrix} \bar{V} & & 0 \\ & \ddots & \\ 0 & & \bar{V} \end{pmatrix} = I_r \otimes \bar{V},$$

$$(33) \quad \text{где } \bar{V} \doteq \begin{pmatrix} v_{(1)}^\top \\ \vdots \\ v_{(N-p)}^\top \end{pmatrix} \in \mathbb{R}^{(N-p) \times (p+1)(r+m)}$$

совпадает с матрицей $V(z)$ (17) с $r = 1$. Обозначим $(\gamma_0^i \gamma_1^i \dots \gamma_p^i) \doteq \gamma_\theta^i$ i -ю строку матрицы γ_θ (14). Тогда соотношение (15) запишется в виде

$$\begin{aligned} \text{vect } \gamma_\theta &= \begin{pmatrix} \gamma_\theta^{1\top} \\ \vdots \\ \gamma_\theta^{r\top} \end{pmatrix} = \begin{pmatrix} d_1 + D_1\theta \\ \vdots \\ d_r + D_r\theta \end{pmatrix} = \\ &= \begin{pmatrix} d_1 & D_1 \\ \vdots & \vdots \\ d_r & D_r \end{pmatrix} \begin{pmatrix} 1 \\ \theta \end{pmatrix} \doteq \begin{pmatrix} D_1 \\ \vdots \\ D_r \end{pmatrix} \vartheta \doteq D\vartheta. \end{aligned}$$

Оценки ОРС определяются соотношением

$$(34) \quad \theta_S = \arg \min_{\theta} J_S(\theta, \tilde{z}),$$

$$J_S(\theta, \tilde{z}) \doteq \frac{\vartheta^T D_1^T \bar{V}(\tilde{z})^T \bar{V}(\tilde{z}) D_1 \vartheta}{\vartheta^T D_1^T D_1 \vartheta} + \dots + \frac{\vartheta^T D_r^T \bar{V}(\tilde{z})^T \bar{V}(\tilde{z}) D_r \vartheta}{\vartheta^T D_r^T D_r \vartheta}.$$

Несложно установить равенство

$$(35) \quad J_S(\theta, \tilde{z}) = \vartheta^T D^T V(\tilde{z})^T C_{S,\theta} V(\tilde{z}) D \vartheta,$$

$$C_{S,\theta} \doteq I_{N-p} \otimes \begin{pmatrix} \gamma_{\theta}^1 \gamma_{\theta}^{1T} & & 0 \\ & \ddots & \\ 0 & & \gamma_{\theta}^r \gamma_{\theta}^{rT} \end{pmatrix}^{-1}.$$

Выражение (35) получается из (27) занулением недиагональных элементов матрицы $\gamma_{\theta}^1 \gamma_{\theta}^{1T}$.

Матрицы P_j остаются теми же, что и в методе ОРМ (26).

Убедимся, что оценки ОРС принадлежат классу вариационных. Действительно, ввиду соотношений (23), (25), (34) ядро целевой функции $J_S(\theta, \tilde{z})$ есть сумма проекторов

$$\Pi_{1,\theta}^{(1)} + \dots + \Pi_{N-p,\theta}^{(1)} + \dots + \Pi_{1,\theta}^{(r)} + \dots + \Pi_{N-p,\theta}^{(r)},$$

где проекторы $\Pi_{j,\theta}^{(i)}$ получаются из проекторов $\Pi_{j,\theta}$ (23) заменой γ_{θ} на строку γ_{θ}^i . Далее см. определение 2.

3. Свойства вариационных оценок

Установим условия единственности вариационных оценок (7), (13). Воспользуемся представлениями (18), (19), (20) для целевых функций (6), (12).

Теорема 2. *Необходимым и достаточным условием единственности вариационной оценки (7) параметра θ системы (4) по невозмущенному наблюдению $\tilde{z} = z$, $G_{\theta} z = 0$, является строгая положительная определенность матрицы*

$$Q_M \doteq D^T \left(\sum_{j=1}^M V_j(z)^T V_j(z) \right) D > 0,$$

где $V_j(z) \doteq P_j V(z)$, и матрицы P_j определены в (18).

Доказательство. Единственность равносильна соотношению

$$\forall \Delta\theta \in \mathbb{R}^v \quad J(\theta + \Delta\theta, z) > J(\theta, z).$$

Из (19) следует равенство

$$\begin{aligned} J(\theta + \Delta\theta, z) - J(\theta, z) &= \\ &= \Delta\theta^\top D^\top \left(\sum_{j=1}^M V_j(z)^\top C_{j,\theta+\Delta\theta} V_j(z) \right) D \Delta\theta. \end{aligned}$$

Здесь учтено равенство $V_j(z) (D\theta + d) = 0$. Имеет место строгая положительная определенность матриц $C_{j,\theta+\Delta\theta} > 0$. Поэтому неравенство $J(\theta + \Delta\theta, z) > J(\theta, z)$ имеет место тогда и только тогда, когда строго положительно определена матрица

$$D^\top \left(\sum_{j=1}^M V_j(z)^\top V_j(z) \right) D > 0.$$

Теорема доказана.

Теорема 3. *Необходимым и достаточным условием единственности вариационной оценки (13) параметра θ системы (4) по набору невозмущенных решений $\{\check{z}_{(1)} = z_{(1)}, \dots, \check{z}_{(L)} = z_{(L)}\}$, $G_\theta z_{(i)} = 0$, является строгая положительная определенность матрицы*

$$(36) \quad Q_{(L),M} \doteq L^{-1} \sum_{i=1}^L D^\top \left(\sum_{j=1}^M V_j(z_{(i)})^\top V_j(z_{(i)}) \right) D > 0.$$

Доказательство этой теоремы опирается на соотношение (20) и аналогично доказательству теоремы 2.

3.1. УСЛОВИЯ КОРРЕКТНОСТИ ПАРАМЕТРИЗАЦИИ

Выясним, как зависит выполнение условия единственности (36) от устройства матрицы D . Согласно (15), матрица D используется для задания параметризации $\theta \leftrightarrow G_\theta$.

Определение 5 [43]. Пусть $Z = \{z_{(1)}, \dots, z_{(L)}\}$ – множество наблюдений $z_{(i)}$ таких, что $G_\theta z_{(i)} = 0$. отождествим множество Z с матрицей $(z_{(1)} \dots z_{(L)})$, тогда $G_\theta Z = 0$. Параметризация $\theta \leftrightarrow G_\theta$ называется корректной на множестве наблюдений Z , если для каждого значения параметра $\xi \in \Theta \setminus \{\theta\}$ имеет место неравенство $G_\xi Z \neq 0$.

Дадим обобщение этого определения на случай набора матриц (18) $\Phi_{j,\theta} \doteq P_j G_\theta$, $j \in \overline{1, M}$. Введем обозначения

$$(37) \quad \Phi_\theta \doteq \begin{pmatrix} \Phi_{1,\theta} \\ \vdots \\ \Phi_{M,\theta} \end{pmatrix} = \underbrace{\begin{pmatrix} P_1 \\ \vdots \\ P_M \end{pmatrix}}_{\tilde{P}} G_\theta \doteq \tilde{P} G_\theta.$$

Определение 6. Параметризация $\theta \leftrightarrow \Phi_\theta$ называется корректной на множестве наблюдений Z , если для каждого значения параметра $\xi \in \Theta \setminus \{\theta\}$ имеет место неравенство $\Phi_\xi Z \doteq \tilde{P} G_\xi Z \neq 0$.

Предложение 1. Условие единственности (36) выполнено тогда и только тогда, когда параметризация $\theta \leftrightarrow \Phi_\theta$ корректна на множестве наблюдений $Z = \{z_{(1)}, \dots, z_{(L)}\}$.

Доказательство. Согласно определению (16), неравенство $\Phi_\xi Z \neq 0$ равносильно неравенству

$$\underbrace{\begin{pmatrix} \tilde{P}V(z_{(1)}) \\ \vdots \\ \tilde{P}V(z_{(L)}) \end{pmatrix}}_W (D\xi + d) \doteq W (D\xi + d) \neq 0.$$

Учитывая соотношения $\Phi_\theta Z = 0$, $W (D\theta + d) = 0$, получаем:

$$W (D\xi + d) - W (D\theta + d) = WD (\xi - \theta) \neq 0.$$

Ввиду произвольности ξ последнее означает, что столбцы матрицы WD линейно независимы. Следовательно, строго положи-

тельно определена матрица

$$Q_{(L),M} \doteq L^{-1} \sum_{i=1}^L D^\top \left(\sum_{j=1}^M V_j(z_{(i)})^\top V_j(z_{(i)}) \right) D = \\ = L^{-1} D^\top W^\top W D > 0.$$

Обратное тоже верно. Предложение доказано.

Определение 7. Пусть $Z_\theta \doteq \{z : G_\theta z = 0\}$ — линейное многообразие всех решений системы (4). И пусть H_θ — матрица, столбцы которой образуют фундаментальную систему решений (4) (базис Z_θ). Параметризация $\theta \leftrightarrow G_\theta$ (или $\theta \leftrightarrow \Phi_\theta$) называется корректной, если она корректна на множестве $Z = H_\theta$.

Заметим, что параметризация $\theta \leftrightarrow G_\theta$ полностью определяется системой (4), а параметризация $\theta \leftrightarrow \Phi_\theta$ зависит еще и от метода идентификации (ОР, ОРМ, ВИ, ОРС, ...), т.е. от набора матриц P_j (37).

Из корректности параметризации $\theta \leftrightarrow \Phi_\theta$ следует корректность параметризации $\theta \leftrightarrow G_\theta$. Обратное, вообще говоря, неверно. Установим условия, при которых корректность $\theta \leftrightarrow G_\theta$ равносильна корректности $\theta \leftrightarrow \Phi_\theta$. В этом случае условия единственности оценок не будут зависеть от метода идентификации (от набора матриц P_j).

Определение 8 [15]. Множество наблюдений $Z = \{z_{(1)}, \dots, z_{(L)}\} \subset Z_\theta$ называем полным, если оно содержит базис многообразия Z_θ .

Покажем, что для оценок ОРМ, ВИ, ОРС условия корректности параметризаций $\theta \leftrightarrow G_\theta$ и $\theta \leftrightarrow \Phi_\theta$ совпадают. Другими словами, имеет место следующая теорема.

Теорема 4. При полном множестве наблюдений $Z = \{z_{(1)}, \dots, z_{(L)}\} \subset Z_\theta$ единственность оценок ОРМ, ВИ, ОРС имеет место тогда и только тогда, когда корректна параметризация $\theta \leftrightarrow G_\theta$.

Доказательство. Согласно предложению 1, единственность любой оценки вида (13) равносильна корректности параметриза-

ции $\theta \leftrightarrow \Phi_\theta$, т. е. когда для всех значений параметров $\xi \in \Theta \setminus \{\theta\}$ имеет место неравенство $\Phi_\xi H_\theta \neq 0$. Заметим, что для каждой из оценок ОРМ, ВИ, ОРС соответствующие матрицы $\tilde{P} = \begin{pmatrix} P_1 \\ \vdots \\ P_M \end{pmatrix}$ (26), (28) все имеют линейно независимые столбцы. Поэтому соотношения $\Phi_\xi H_\theta = \tilde{P} G_\xi H_\theta \neq 0$ и $G_\xi H_\theta \neq 0$ для этого случая равносильны. Теорема доказана.

Важно, что условия корректности параметризации $\theta \leftrightarrow G_\theta$ допускают равносильную формулировку в виде ограничений на ранги специальных подматриц из элементов «малой» матрицы γ_θ (14) [16, 17, 18].

Определение 9 [17]. Если все ненулевые строки матрицы D образуют единичную подматрицу, то параметризация называется простой (или правильной).

В литературе это наиболее часто встречающийся тип параметризаций.

Замечание 1. Ввиду того, что свойство максимальности ранга устойчиво по отношению к малым изменениям элементов матрицы, для простых параметризаций корректность оказывается в пространстве параметров свойством общего положения (этот факт следует из теоремы 4 статьи [17]).

К этому замечанию мы вернемся ниже при обсуждении состоятельности.

3.2. УСТОЙЧИВОСТЬ

Устойчивость вариационной оценки $\hat{\theta}(\tilde{z})$ (7) прямо следует из результатов статьи [21]. Сформулируем утверждение.

Теорема 5. Пусть выполнено условие единственности из теоремы 2, $\epsilon \doteq \|\tilde{z} - z\|$, $\Delta\theta \doteq \hat{\theta}(\tilde{z}) - \hat{\theta}(z) = \hat{\theta}(\tilde{z}) - \theta$,

$$(38) \quad R_0 \doteq D^\top \left(\sum_{j=1}^M V_j(z)^\top C_{j,\theta} V_j(z) \right) D$$

и $\lambda_{\min}(R_0)$ — наименьшее собственное число. Тогда

$$\|\Delta\theta\| \leq \frac{1}{\sqrt{\lambda_{\min}(R_0)}} \epsilon + O(\epsilon^2).$$

Верхняя граница для величины остаточного члена $O(\epsilon^2)$ также получена в [21]. Она имеет сложный вид и здесь не приводится.

Для оценки (13) утверждение об устойчивости аналогично теореме 5. Ввиду соотношения (20) нужно наложить условие единственности из теоремы 3 и заменить матрицу R_0 (38) на матрицу

$$(39) \quad R_{\Sigma,0} \doteq L^{-1} \sum_{i=1}^L D^\top \left(\sum_{j=1}^M V_j(z_{(i)})^\top C_{\Sigma,j,\theta} V_j(z_{(i)}) \right) D.$$

На основании теоремы 5 можно предложить количественные критерии идентифицируемости параметров уравнения (1), основанные на вычислении минимального собственного числа матрицы R_0 (38) (или $R_{\Sigma,0}$ (39)). На практике матрица R_0 может быть заменена оценкой \hat{R}_0 , получаемой заменой неизвестных истинных значений θ, z на оценки $\hat{\theta}, z_{\text{opt}}$ [22].

3.3. СОСТОЯТЕЛЬНОСТЬ

Прежде изучения свойств оценки на конечных выборках следует убедиться в ее состоятельности. В статье [40] была установлена состоятельность оценок STLS в пределе $N \rightarrow \infty$ по наблюдению одного процесса ($L = 1$). Здесь рассмотрим предельный случай $L \rightarrow \infty$ наблюдений большого числа процессов конечных длин N с независимыми начальными условиями (подлежащими оцениванию совместно с параметрами уравнения). В этом случае оценка $\hat{\theta}(\tilde{z})$ (7) переходит в оценку $\hat{\theta}(\tilde{z}_{(1)}, \dots, \tilde{z}_{(L)}) \doteq \hat{\theta}_L$ (10) или (13), и сильная состоятельность означает сходимость $\hat{\theta}_L \rightarrow \theta$ п. н. в пределе $L \rightarrow \infty$, где $\hat{\theta}_L$ — одна из точек локального минимума целевой функции. Сформулируем утверждение.

Теорема 6. Пусть выполнено предположение о возмущениях (11). Если (и только если) выполнено условие

$$(40) \quad Q_{(\infty),M} \doteq \lim_{L \rightarrow \infty} Q_{(L),M} > 0,$$

где матрица $Q_{(L),M}$ определена в (36), то вариационная оценка $\hat{\theta}_L$ (13) сильно состоятельна по $L \rightarrow \infty$, где под $\hat{\theta}_L$ понимается одна из точек локального минимума целевой функции (12).

Доказательство близко по идее к доказательству теоремы 1 из [15] и полностью будет опубликовано в другой статье.

В ряде случаев оказывается возможным заменить целевую функцию на более простую. Необходимым условием является сохранение состоятельности оценок.

Определение 10. При возмущениях (11) две оценки вида (13) с различающимися наборами матриц

$$\mathcal{W}_1 \doteq \{W_{1_1}, \dots, W_{M_1}\}, \quad \mathcal{W}_2 \doteq \{W_{1_2}, \dots, W_{M_2}\}$$

называются равносильными по состоятельности, если условия (40) для соответствующих наборов \mathcal{W}_1 и \mathcal{W}_2 равносильны. При отсутствии возмущений оценки вида (13) равносильны по состоятельности в детерминированном смысле, если для них при любых одинаковых множествах невозмущенных наблюдений $\{\check{z}_{(1)} = z_{(1)}, \dots, \check{z}_{(L)} = z_{(L)}\}$ равносильны условия (36).

Заметим, что из равносильности по состоятельности в детерминированном смысле всегда следует равносильность по состоятельности.

Равносильные по состоятельности оценки могут отличаться асимптотическими свойствами.

Теорема 7. Вариационные оценки ВИ, ОРМ, ОРС (с целевыми функциями соответственно (29), (27) и (35)) равносильны по состоятельности в детерминированном смысле.

Доказательство. Достаточно заметить, что перечисленные оценки отличаются только матрицами C_θ , $C_{m,\theta}$, $C_{s,\theta}$ в ядрах целевых функций (29), (27) и (35). Эти матрицы не участвуют в формулировке условий единственности (36).

3.4. СОСТОЯТЕЛЬНОСТЬ И ИДЕНТИФИЦИРУЕМОСТЬ КАК СВОЙСТВА НЕОБЩЕГО ПОЛОЖЕНИЯ

Из теорем 4, 6 следует, что при «естественных» предположениях о распределении процессов $z_{*(i)}$ и возмущений $\varepsilon_{(i)}$ для состоятельности оценок ОРМ, ВИ, ОРС необходимо и достаточно

выполнить условия корректности параметризации $\theta \leftrightarrow G_\theta$. Поэтому, согласно замечанию 1, для простых параметризаций состоятельность оказывается свойством общего положения. Подобный факт был установлен ранее П. Стойкой, который исследовал прямые оценки, получаемые методами типа инструментальных переменных (МИП), и ввел понятие *состоятельности в общем положении* (*generic consistency*) [51, с. 266]. Оно означает отсутствие состоятельности в точках множества меры ноль в пространстве параметров. Причиной является вырождение матрицы ковариации наблюдений [51, с. 272]. В таких точках бесконечно растет и разброс оценок МИП [51, с. 273].

Заметим, что если использовать *количественные* показатели разброса оценок (например, через детерминант или сингулярные числа матрицы ковариации), то понятие состоятельности в общем положении теряет смысл. Например, пусть считается неприемлемым разброс оценок больше заданного уровня E , тогда отсутствие состоятельности (идентифицируемости) будет происходить уже на множестве меры > 0 (эта мера будет зависеть от E).

Нарушение корректности параметризации также происходит на множестве меры ноль (см. замечание 1). Введение количественных показателей идентифицируемости (раздел 3.2) делает множество неидентифицируемых систем множеством меры > 0 , и идентифицируемость перестает быть свойством общего положения.

3.5. МИНИМИЗАЦИЯ ПРОЕКТИВНЫХ ЦЕЛЕВЫХ ФУНКЦИЙ

Минимизация $J(\theta, \tilde{z})$ (22) по θ осуществляется итерациями с обратной матрицей:

$$(41) \quad \vartheta_{(k)} \doteq \begin{pmatrix} 1 \\ \theta_{(k)} \end{pmatrix}, \quad \begin{cases} \tau \doteq A^{-1} B \vartheta_{(k)}, \\ \vartheta_{(k+1)} = \frac{1}{(10\dots 0)\tau} \tau, \end{cases}$$

$$A \doteq D^\top V_{\text{OR}}(\tilde{z})^\top V_{\text{OR}}(\tilde{z}) D, \quad B \doteq D^\top D.$$

Запись $\frac{1}{(10\dots 0)\tau}$ означает нормирование вектора τ с приведением первого элемента к единице. Этот алгоритм является мо-

дификацией стандартного способа поиска минимального (соответствующего минимальному собственному числу) собственного вектора симметричной матрицы [8, с. 421].

Целевая функция метода ВИ (29) имеет сложный характер изоповерхностей [6, 13]. Применение универсальных алгоритмов минимизации типа Ньютона крайне затруднено из-за их малого радиуса сходимости [44]. Тем не менее, в [9, 10] был найден эффективный вычислительный метод, основанный на модификации алгоритма (41):

$$(42) \quad \vartheta_{(k)} \doteq \begin{pmatrix} 1 \\ \theta_{(k)} \end{pmatrix}, \quad \begin{cases} \tau \doteq A_{(k)}^{-1} B \vartheta_{(k)}, \\ \vartheta_{(k+1)} = \frac{1}{(10 \dots 0)_{\tau}} \tau, \end{cases}$$

$$A_{(k)} \doteq D^{\top} V(\tilde{z})^{\top} C_{\theta_{(k)}} V(\tilde{z}) D, \quad B \doteq D^{\top} D.$$

Итерации типа (42) в частном случае однородных систем (1) независимо от А. О. Егоршина использовал М. Осборн [46] и позже переоткрыл Б. Де Мур [45], применяя их для решения задачи STLS (см. раздел 2.4). Особенностью итераций Егоршина—Осборна является высокая скорость сходимости в 2–4 итерации в малую окрестность глобального минимума при слабой зависимости от начального приближения. Примеры расчетов приведены в [15, 20, 22]. Локальную сходимость итерационной процедуры (42) исследовал В. Г. Демиденко [7].

В методе ОРМ минимизация целевой функции (27) может быть осуществлена итерациями (42) с заменой $C_{\theta} = (G_{\theta} G_{\theta}^{\top})^{-1}$ на $C_{m, \theta}$ (27) (результаты экспериментальной проверки этого факта здесь не приводим).

3.6. ЧИСЛО ЭКСТРЕМУМОВ ЦЕЛЕВОЙ ФУНКЦИИ ОРС

Для вариационной идентификации параметров систем (1) оценки ОРС оказываются одними из наиболее простых и в то же время состоятельными (теорема 7), с хорошими экстремальными свойствами.

Теорема 8. Число локальных экстремумов целевой функции ОРС $J_S(\theta, \tilde{z})$ (34) не превосходит $r(p+1)(r+m)$.

Доказательство. Из выражения (34) получаем уравнение на критические точки функции $J_S(\vartheta) \doteq J_S(\theta, \xi)$:

$$\frac{\partial J_S(\vartheta)}{\partial \vartheta} = 2\vartheta^\top \left[\sum_{i=1}^r \frac{D_i^\top \bar{V}^\top \bar{V} D_i}{(\vartheta^\top D_i^\top D_i \vartheta)} - \frac{(\vartheta^\top D_i^\top \bar{V}^\top \bar{V} D_i \vartheta)}{(\vartheta^\top D_i^\top D_i \vartheta)^2} D_i^\top D_i \right] = 0.$$

Обозначим

$$\begin{aligned} a_i(\vartheta) &\doteq \vartheta^\top D_i^\top \bar{V}^\top \bar{V} D_i \vartheta, & A_i &\doteq D_i^\top \bar{V}^\top \bar{V} D_i, \\ b_i(\vartheta) &\doteq \vartheta^\top D_i^\top D_i \vartheta, & B_i &\doteq D_i^\top D_i, \\ J_i(\vartheta) &\doteq \frac{a_i(\vartheta)}{b_i(\vartheta)}. \end{aligned}$$

Уравнение на критические точки запишем в виде

$$\left[\sum_{i=1}^r \frac{1}{b_i(\vartheta)} D_i^\top (\bar{V}^\top \bar{V} - J_i(\vartheta) I) D_i \right] \vartheta = 0.$$

В матричной записи

$$\begin{aligned} (43) \quad & \left(D_1^\top \quad \dots \quad D_r^\top \right) \times \\ & \times \begin{pmatrix} \frac{(\bar{V}^\top \bar{V} - J_1(\vartheta) I)}{b_1(\vartheta)} & & 0 \\ & \ddots & \\ 0 & & \frac{(\bar{V}^\top \bar{V} - J_r(\vartheta) I)}{b_r(\vartheta)} \end{pmatrix} \begin{pmatrix} D_1 \\ \vdots \\ D_r \end{pmatrix} \vartheta \doteq \\ & \doteq D^\top \begin{pmatrix} W_1 & & 0 \\ & \ddots & \\ 0 & & W_r \end{pmatrix} D \vartheta \doteq D^\top W D \vartheta = 0. \end{aligned}$$

Несложно проверить тождество $\forall \vartheta \quad \vartheta^\top D^\top W D \vartheta = 0$. Оно показывает, что производная $\frac{\partial J_S(\vartheta)}{\partial \vartheta}$ перпендикулярна направлению ϑ , т. е. $J_S(\vartheta)$ не зависит от нормы ϑ .

Как следует из принятых обозначений, матрица W симметричная. Ввиду линейной независимости столбцов матрицы D условие (43) при $\vartheta \neq 0$ означает, что матрица W вырождена. Отсюда следует необходимое условие существования экстремальной точки:

$$\exists i \in \overline{1, r} \quad \det \left(\bar{V}^\top \bar{V} - J_i(\vartheta) I \right) = 0.$$

Следовательно, число экстремальных точек не превосходит числа rt , где $t = (p + 1)(r + m)$ есть размерность матрицы $\bar{V}^T \bar{V}$ (см. (33)). Теорема доказана.

Будем говорить, что разные уравнения системы (1) параметризованы аффинно и независимо, если матрица D из условия (15) имеет клеточно-диагональный вид

$$(44) \quad D = \begin{pmatrix} D_{11} & & 0 \\ & \ddots & \\ 0 & & D_{rr} \end{pmatrix},$$

$$D_{ii} \in \mathbb{R}^{(p+1)(r+m) \times v_i}, \quad \sum_{i=1}^r v_i = v + 1,$$

а вектор параметров $\vartheta = (1; \theta)$ распадается на r подвекторов $\vartheta^i = (1; \theta^i)$, $i \in \overline{1, r}$, так что каждая из строк γ^i в матрице γ_θ (14) системы (1), (4) зависит только от i -го подвектора параметров θ^i .

Теорема 9. Пусть разные уравнения системы (1) параметризованы аффинно и независимо. Тогда число экстремальных точек целевой функции ОРС (34) не превосходит $r(v + 1)$, а оценка ОРС ϑ_S , минимизирующая (34), имеет вид $\vartheta_S = (\vartheta_S^1; \dots; \vartheta_S^r)$, где каждый из подвекторов ϑ_S^i является решением скалярной ($r = 1$) задачи ОРМ вида (25):

$$(45) \quad \vartheta_S^i = \arg \min_{\theta} J_i(\vartheta),$$

$$J_i(\vartheta) \doteq \frac{\vartheta^T D_{ii}^T \bar{V}^T \bar{V} D_{ii} \vartheta}{\vartheta^T D_{ii}^T D_{ii} \vartheta}.$$

Доказательство. В случае независимой параметризации уравнений из выражения (35) ввиду (44) и (33) получаем

$$J_S(\vartheta) = J_S(\vartheta^1, \dots, \vartheta^r) =$$

$$= \frac{\vartheta^1 T D_{11}^T \bar{V}^T \bar{V} D_{11} \vartheta^1}{\vartheta^1 T D_{11}^T D_{11} \vartheta^1} + \dots + \frac{\vartheta^r T D_{rr}^T \bar{V}^T \bar{V} D_{rr} \vartheta^r}{\vartheta^r T D_{rr}^T D_{rr} \vartheta^r}.$$

Целевая функция (45) имеет ровно $v + 1$ критических точек, где $v + 1$ — размерность вектора ϑ , поскольку все такие точки являются собственными векторами регулярного пучка матриц $D_{ii}^T (\bar{V}^T \bar{V} - \mu I) D_{ii}$ [4, гл.Х, п.6]. Отсюда сразу следует утверждение теоремы.

4. Заключение

После работ А.Н. Колмогорова и Н. Винера [12, 54] в области приложений методов идентификации возобладали подход, основанный на вычислении корреляционных функций наблюдаемых сигналов. Это направление со временем стало магистральным [1, 23, 29, 32, 35, 51], во многом благодаря ранним работам К. Острема с коллегами [25]. Несмотря на простоту, корреляционные («прямые») методы идентификации на конечных выборках наблюдений нередко приводили к трудностям в интерпретации результатов ввиду большого числа локальных экстремумов у целевой функции.

Известная работа Я. Виллемса [2] послужила толчком для развития направления, основанного более на идеях аппроксимации наблюдений, чем на вычислении коэффициентов корреляций. Эту группу методов мы называем вариационными («непрямыми»). Первые существенные результаты в этой области были получены значительно раньше А.О. Егоршиным [9, 10] и М. Осборном [46]. Развитие этого направления связано с именами Г. Голуба, Ч. Ван Лоана [39], Б. Де Мура [45], С. Ван Хуффель, Дж. Вандевалле [53] и мн. др. Непрямые методы идентификации более сложны с точки зрения числа арифметических операций.

В данной статье впервые предложено точное определение для всего класса вариационных методов и рассмотрены их экстремальные свойства. Показано, что в ряде существенных случаев увеличение числа арифметических операций приводит не к усложнению процедуры идентификации, а в конечном счете, наоборот, к ее упрощению, ввиду значительного уменьшения числа локальных экстремумов, и как следствие, улучшения свойств сходимости итерационных процедур.

Автор благодарит рецензентов за конструктивные замечания, способствовавшие улучшению текста статьи.

Приложение. Доказательство теоремы 1

Первое утверждение теоремы сформулируем более конкретно. Пусть θ_S, z_S — решение задачи (30)–(32). Тогда можно указать целевую функцию $J_S(\theta, \check{x}) = \check{x}^\top W_{S,\theta} \check{x}$, $\check{x} \doteq (1; \check{z})$, с ядром таким, что произведение $\Sigma^\top W_{S,\theta} \Sigma = \Pi_{S,\theta}$, $\Sigma = \begin{pmatrix} 0 & 0 \\ 0 & R \end{pmatrix}$, есть проективная матрица $\Pi_{S,\theta}^2 = \Pi_{S,\theta}$, и при этом $\theta_S = \arg \min_{\theta} J_S(\theta, \check{z})$. Докажем это утверждение.

Условие $V(z)g(\theta) = 0$ равносильно уравнению

$$[I \otimes g(\theta)] \text{vect } V(z) = 0.$$

Ввиду соотношения (30) имеем $\text{vect } V(z) = f + Fz$, где вектор f и матрица F вычисляются по матрицам V_0, \dots, V_n . Также из (32) следует представление

$$I \otimes g(\theta) = \Gamma_{\theta}, \quad \Gamma_{\theta} = \Gamma_0 + \Gamma_1 \theta_1 + \dots + \Gamma_v \theta_v,$$

где матрицы $\Gamma_0, \dots, \Gamma_v$ вычисляются по матрицам g_0, \dots, g_v .

Задачу (31) запишем в виде

$$\min_{z \in \mathbb{R}^n, \theta \in \mathbb{R}^v} \|R^{-1}(\check{z} - z)\|^2 \quad \text{при условии} \quad \Gamma_{\theta}(f + Fz) = 0.$$

Определим векторы $x(z) \doteq \begin{pmatrix} 1 \\ z \end{pmatrix} \in \mathbb{R}^{n+1}$, $\check{x} \doteq x(\check{z})$ и матрицы $\Sigma \doteq \begin{pmatrix} 0 & 0 \\ 0 & R \end{pmatrix}$, $G_{\theta} \doteq (\Gamma_{\theta} f \quad \Gamma_{\theta} F)$. Тогда $G_{\theta} x = 0$, $\|R^{-1}(\check{z} - z)\| = \|\Sigma^+ (\check{x} - x)\|$, где $\Sigma^+ = \begin{pmatrix} 0 & 0 \\ 0 & R^{-1} \end{pmatrix}$ — обобщенная обратная Мура–Пенроуза [27, 1.b.5.(IV)]. В новых обозначениях задача (31) принимает вид

$$\min_{x \in \mathbb{R}^{n+1}, \theta \in \mathbb{R}^v} \|\Sigma^+ (\check{x} - x)\|^2$$

при условиях

$$G_{\theta} x = 0, \quad (1, 0, \dots, 0)x = 1.$$

Минимизация по x приводит к безусловной задаче минимизации

$$\min_{\theta \in \mathbb{R}^v} \check{x}^\top G_\theta^\top \underbrace{\left(G_\theta \Sigma \Sigma^\top G_\theta^\top \right)^{-1}}_{W_{S,\theta}} G_\theta \check{x},$$

$$x_{\text{opt}} = \left[I - \Sigma \Sigma^\top G_\theta^\top \left(G_\theta \Sigma \Sigma^\top G_\theta^\top \right)^{-1} G_\theta \right] \check{x}.$$

Матрица $\Sigma^\top W_{S,\theta} \Sigma$ является проектором. Поэтому оценка STLS

$$\hat{\theta} = \arg \min_{\theta} \check{x}^\top W_{S,\theta} \check{x}$$

есть вариационная согласно определению 3. Первая часть теоремы доказана.

Прежде доказательства второй части установим три леммы.

Лемма 1. Пусть $FXF^\top \doteq \Pi = \Pi^\top \geq 0$ — неотрицательно определенная симметричная матрица со свойством проектора: $\Pi^2 = \Pi$. И пусть матрица X неособенная, и столбцы матрицы F линейно независимы. Тогда $X = (F^\top F)^{-1}$.

Доказательство. Из равенства $(FXF^\top)(FXF^\top) = FXF^\top$ следует $F^\top(FXF^\top)(FXF^\top)F = F^\top FXF^\top F$, откуда сразу получаем $XF^\top FX = X$. Учитывая неособенность X , приходим к соотношению $F^\top FX = I$, т. е. $X = (F^\top F)^{-1}$.

Лемма 2. Пусть $W^\top = W \geq 0$ — неотрицательно определенная симметричная матрица со свойством $(Gz = 0) \Rightarrow (Wz = 0)$, где G — заданная матрица с линейно независимыми строками. И пусть для заданной матрицы Σ с линейно независимыми столбцами выполняется соотношение $\Sigma^\top W \Sigma \doteq \Pi$, $\Pi^2 = \Pi$. Тогда $W = G^\top P^\top (PG\Sigma\Sigma^\top G^\top P^\top)^{-1} PG$, где P — некоторая матрица с линейно независимыми строками.

Доказательство. Из симметричности W и свойства $(Gz = 0) \Rightarrow (Wz = 0)$ следует $W = G^\top P^\top XPG$, где X — некоторая неособенная матрица. По условию,

$$\left(\Sigma^\top W \Sigma \right) \left(\Sigma^\top W \Sigma \right) = \Sigma^\top W \Sigma.$$

После переобозначения $PG\Sigma \doteq F^\top$ получаем равенство $(FXF^\top)(FXF^\top) = FXF^\top$. По лемме 1 $X = (F^\top F)^{-1}$, тогда

$$W = G^\top P^\top \left(PG\Sigma\Sigma^\top G^\top P^\top \right)^{-1} PG.$$

Лемма доказана.

Лемма 3. *Задача минимизации*

$$J = \min_{z \in \mathbb{R}^n} \|\Sigma^{-1}(\check{z} - z)\|^2 \quad \text{при условии} \quad Hz = 0$$

имеет единственное решение

$$J = \check{z}^\top H^\top \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H\check{z},$$

$$z_{\text{opt}} = \left[I - \Sigma\Sigma^\top H^\top \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H \right] \check{z}.$$

Доказательство. Введем функцию Лагранжа

$$J^*(z, \lambda) = \|\Sigma^{-1}(\check{z} - z)\|^2 + \lambda^\top Hz.$$

Необходимое условие минимума $J_\lambda^{*'} = 0$, $J_z^{*'} = 0$:

$$Hz = 0, \quad H^\top \lambda - 2\Sigma^{-\top} \Sigma^{-1}(\check{z} - z) = 0.$$

Левым умножением последнего уравнения на $H\Sigma\Sigma^\top$ получаем

$$\lambda = 2 \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H\check{z}.$$

Подставляя λ в условие минимума, вычисляем

$$z_{\text{opt}} = \left[I - \Sigma\Sigma^\top H^\top \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H \right] \check{z}.$$

Подстановка z_{opt} в J приводит к выражению

$$J = \check{z}^\top H^\top \left(H\Sigma\Sigma^\top H^\top \right)^{-1} H\check{z}.$$

Лемма доказана.

Докажем вторую часть теоремы. Для упрощения обозначений положим $L = 1$, $M = 1$. В случае $L > 1$, $M > 1$ ход рассуждений остается тем же.

Пусть (12), т. е. $J(\theta, \check{z}) = \check{z}^\top W_\theta \check{z}$. Матрица W_θ удовлетворяет условиям леммы 2, поэтому имеет место выражение

$$W_\theta = G_\theta^\top P^\top \left(P G_\theta \Sigma \Sigma^\top G_\theta^\top P^\top \right)^{-1} P G_\theta$$

для некоторой матрицы P с линейно независимыми строками. Обозначим $H_\theta \doteq P G_\theta$, тогда

$$J(\theta, \check{z}) = \check{z}^\top H_\theta^\top \left(H_\theta \Sigma \Sigma^\top H_\theta^\top \right)^{-1} H_\theta \check{z}.$$

По лемме 3

$$J(\theta, \check{z}) = \min_{z \in \mathbb{R}^n} \|\Sigma^{-1}(\check{z} - z)\|^2 \quad \text{при условии} \quad H_\theta z = 0.$$

Используя определения (16), (17), всегда можно построить матрицу $V(z)$ с аффинной параметризацией (30) такую, что верно соотношение $H_\theta z \equiv V(z) \text{vect } \gamma_\theta \equiv V(z) D \vartheta = 0$. Осталось заметить, что для матрицы (из одного столбца)

$$g(\theta) \doteq \text{vect } \gamma_\theta = D \vartheta = d + D\theta$$

выполнено условие (32). Мы получили, что оценка (13) есть оценка STLS (30)–(32). Теорема доказана.

Литература

1. БУНИЧ А. Л., БАХТАДЗЕ Н. Н. *Синтез и применение дискретных систем управления с идентификатором*. – М.: Наука, 2003. – 232 с.
2. ВИЛЛЕМС Я. *От временного ряда к линейной системе // Теория систем. Математические методы и моделирование*. – М.: Мир, 1989. – С. 8–191.

3. ВОРЧИК Б. Г. *Единственность оценок максимального правдоподобия параметров стохастических систем (проблема локальных экстремумов)* // Автоматика и телемеханика. – 1984. – № 6. – С. 47–55.
4. ГАНТМАХЕР Ф. Р. *Теория матриц*. – М.: Наука, 1966. – 576 с.
5. ГАУСС К. Ф. *Избранные геодезические сочинения*. – Т. 1. М.: Геодезиздат, 1957. – С. 59–87.
6. ДЕМИДЕНКО В. Г. *Разработка и программная реализация метода оценки параметров линейных моделей генных сетей*. – Магистерская диссертация. Новосибирск: Новосибирский гос. университет, 2007.
7. ДЕМИДЕНКО В. Г. *Восстановление коэффициентов систем линейных разностных уравнений* // Вестник НГУ (Серия: математика, механика, информатика). – 2010. – Т. 10. Вып. 2. – С. 45–53.
8. ДЕМИДОВИЧ Б. П., МАРОН И. А. *Основы вычислительной математики*. – М.: Наука, 1966. – 664 с.
9. ЕГОРШИН А. О. *Вычислительные замкнутые методы идентификации линейных объектов* // Оптимальные и самонастраивающиеся системы. – Новосибирск, 1971. – С. 40–53.
10. ЕГОРШИН А. О. *Метод наименьших квадратов и «быстрые» алгоритмы в вариационных задачах идентификации и фильтрации (метод ВИ)* // Автотметрия. – 1988. – № 1. – С. 30–42.
11. КАШЬЯП Р. Л., РАО А. Р. *Построение динамических стохастических моделей по экспериментальным данным*. – М.: Наука, 1983. – 384 с.
12. КОЛМОГОРОВ А. Н. *Интерполирование и экстраполирование стационарных случайных последовательностей* // Изв. АН СССР. Сер. мат. – 1941. – Т. 5. – С. 3–14. (В кн.: Колмогоров А. Н. *Теория вероятностей и математическая статистика*. – М.: Наука, 1986. – С. 255–263.)
13. КОСТИН В. И. *О точках экстремума одной функции* //

- Управляемые системы. – Новосибирск: ИМ СО АН СССР, 1984. Т. 24. – С. 35–42.
14. ЛОМОВ А. А. *Минимальные описания стационарных линейных моделей* // Труды Института математики СО РАН. Модели и методы оптимизации. – Новосибирск: ИМ СО РАН, 1994. Т. 28, – С. 91–117.
 15. ЛОМОВ А. А. *Идентификация линейных динамических систем по коротким участкам переходных процессов при аддитивных измерительных возмущениях* // Известия РАН ТСУ. – 1997. – № 3. – С. 20–26.
 16. ЛОМОВ А. А. *Параметрическая идентифицируемость линейных стохастических систем по наблюдениям коротких отрезков траекторий* // Известия РАН ТСУ. – 2002. – № 2. – С. 53–58.
 17. ЛОМОВ А. А. *Условия различимости стационарных линейных систем* // Дифференц. уравнения. – 2003. – Т. 39. № 2. – С. 261–266.
 18. ЛОМОВ А. А. *О различимости стационарных линейных систем с коэффициентами, зависящими от параметра* // Сибирский журнал индустриальной математики. – 2003. – Т. 6. № 4(16). – С. 60–66.
 19. ЛОМОВ А. А. *Сравнение методов оценивания параметров линейных динамических систем по измерениям коротких участков переходных процессов* // Автоматика и телемеханика. – 2005. – № 3. – С. 39–47.
 20. ЛОМОВ А. А. *Оценка трендов и идентификация динамики временных рядов на коротких интервалах наблюдения* // Известия РАН ТСУ. – 2009. – № 1. – С. 25–37.
 21. ЛОМОВ А. А. *О локальной устойчивости в задаче идентификации коэффициентов линейного разностного уравнения* // Вестник НГУ. Серия: Математика, механика, информатика. – 2010. – Т. 10, вып. 4. – С. 81–103.
 22. ЛОМОВ А. А. *О количественных априорных показателях идентифицируемости коэффициентов линейных динамических систем* // Известия РАН ТСУ. – 2011. – № 1. – С. 3–

- 15.
23. ЛЬЮНГ Л. *Идентификация систем*. – М.: Наука, 1991. – 432 с.
24. НОРКИН К.Б. *Поисковые методы настройки управляемых моделей в задачах настройки параметров объектов* // Автоматика и телемеханика. – 1968. – № 11. – С. 61–67.
25. ОСТРЕМ К., БОЛИН Т. *Цифровая идентификация линейных динамических систем на основе данных о нормальном режиме работы* // Теория самонастраивающихся систем управления: Труды II Международного конгресса IFAC. М.: Наука, 1969. – С. 99–116.
26. ПЕРЕЛЬМАН И. И. *Методы состоятельного оценивания параметров линейных динамических объектов и проблематичность их реализации на конечных выборках* // Автоматика и телемеханика. – 1981. – № 3. – С. 49–55.
27. РАО С. Р. *Линейные статистические методы и их применение*. – М.: Наука, 1968. – 548 с.
28. *Справочник по теории автоматического управления* / Под. ред. А. А. Красовского. – М.: Наука, 1987. – 712 с.
29. ЦЫПКИН Я. З. *Основы информационной теории идентификации*. – М.: Наука, 1984. – 336 с.
30. AVATZOGLU T. J., MENDEL J. M. *Constrained Total Least Squares* // Proc. 1987 IEEE ICASSP (Dallas). 1987. – P. 1485–1488.
31. АОКИ М., YUE P. C. *On A Priori Error Estimates of Some Identification Methods* // IEEE Trans. on Automat. Control. V. AC-15. 1970. – P. 541–548.
32. ÅSTROM K. J., WITTENMARK B. *Adaptive Control*. – 2nd ed. Mineola, New York: Dover Publ. Inc. 2008. – 580 p.
33. BOHLIN T. *On the Problem of Ambiguities in Maximum Likelihood Identification* // Automatica. – 1971. – V. 7. – P. 137–146.
34. BROCKETT R. *Finite Dimensional Linear Systems*. – New York: Wiley, 1970. – 129 p.
35. CHEN H.-F. *Recursive Identification of EIV ARMA Processes*

- // Proceedings of the 17th World Congress IFAC. Seoul, Korea, July 6–11, 2008. – P. 1366–1371.
36. CHENG C.-L., VAN NESS J. W. *Statistical Regression with Measurement Error* // Kendall's Library of Statistics, 6, Arnold, – London. 1999. – 262 p.
37. DUPUIS P., SELS T., DRIESEN T., BELMANS R. *Exponential Parameters Measurement Using a Modified Prony Method* // Proc. Instrumentation and Measurement Technology Conference. Como, Italy, 18–20 May. 2004. – P. 1590–1594.
38. GLESER L. J. *Improvements of the Naive Approach to Estimation in Nonlinear Errors-in-Variables Regression Models* // Contemporary Mathematics. V. 112. Statistical Analysis of Measurement Error Models and Applications. Providence, Rhode Island: AMS, 1990. – P. 99–114.
39. GOLUB G. H., VAN LOAN C. F. *An Analysis of the Total Least Squares Problem* // SIAM J. Numer. Anal. – 1980. – V. 17, – P. 883–893.
40. KUKUSH A., MARKOVSKY I., VAN HUFFEL S. *Consistency of the Structured Total Least Squares Estimator in a Multivariate Errors-in-variables Model* // Journal of Statistical Planning and Inference. – 2005. – V. 133. No. 2. – P. 315–358.
41. LEMMERLING PH., MOOR DE B., VANHUFFEL S. *On the Equivalence of Constrained Total Least Squares and Structured Total Least Squares* // IEEE Trans. on Signal Processing. – 1996. V. 44. No. 11. – P. 2908–2911.
42. LEVIN M. J. *Estimation of a System Pulse Transfer Function in the Presence of Noise* // IEEE Trans. on Automatic Control. – 1964. – V. AC-9. – P. 229–235.
43. LOMOV A. A. *Correct Parametrizations of Linear Models* // Siberian Advances in Mathematics. – 1994. V. 4. – P. 95–113.
44. MAINE R. E., ILIFF K. W. *Formulation and Implementation of a Practical Algorithm for Parameter Estimation with Process and Measurement Noise* // SIAM Journal on Applied

- Mathematics. – 1981. V. 41. No. 3. – P. 558–579.
45. MOOR DE B. *Structured Total Least Squares and L_2 Approximation Problems* // Linear Algebra Appl. – 1993. V. 188–189. – P. 163–207.
 46. OSBORNE M.R. *A Class of Nonlinear Regression Problems* // Data Representation / Eds. R. S. Anderssen and M. R. Osborne. St. Lucia: University of Queensland Press, 1970. – P. 94–101.
 47. OSBORNE M.R., SMYTH G.K. *A modified Prony algorithm for fitting functions defined by difference equations* // SIAM J. Sci. Statist. Comput. – 1991. – V. 12. – P. 362–382.
 48. PEARSON K. *On Lines and Planes of Closest Fit to Systems of Points in Space* // Phil. Mag. – 1901. – VI. No. 2. – P. 559–572.
 49. ROORDA B., HEIJ C. *Global Total Least Squares Modelling of Multivariable Time Series* // IEEE Trans. on Automatic Control. – 1995. – V. AC-40. – P. 50–63.
 50. SÖDERSTRÖM T. *On the Uniqueness of Maximum Likelihood Identification* // Automatica. – 1975. – V. 11. – P. 193–197.
 51. SÖDERSTRÖM T., STOICAP. *System Identification*. – London: Prentice-Hall, 1989. – 613 p.
 52. SÖDERSTRÖM T. *Errors-in-variables Methods in System Identification* // 14th IFAC Symposium on System Identification, Newcastle, Australia, 2006. – P. 1–19.
 53. VAN HUFFEL S., VANDEWALLE J. *The total least squares problem*. – SIAM, Philadelphia, 1991. – 300 p.
 54. WIENER N. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. NDRC Report to the Services 370, February 1, 1942. – Cambridge: MIT Press, 1949. – 176 p.

VARIATIONAL IDENTIFICATION METHODS FOR LINEAR DYNAMIC SYSTEMS AND THE LOCAL EXTREMA PROBLEM

Andrei Lomov, Sobolev Institute of Mathematics of the Siberian Branch of RAS, Novosibirsk State University, Novosibirsk;
Cand.Sc., assistant professor (lomov@math.nsc.ru).

Abstract: The problem of a large number of local extrema is considered. This problem arises when using «direct» methods to identify parameters of linear dynamical systems with finite-sample observations. A new class of variational («indirect») parameter estimators is defined by the projectivity property of matrix kernels in the objective function. The variational objective functions are constructed having the number of local extrema not greater than the number of elements in system matrices. We obtain conditions for consistency of variational estimates in the limit of large number of observations of independent finite-length trajectories.

Keywords: parameter identification, difference equations, dynamic systems.

*Статья представлена к публикации
членом редакционной коллегии Н. Н. Бахтадзе*

УДК 681.5.013; 681.514; 681.511.22; 681.516.42; 681.516.74
ББК 32.965.4.32.965.6

СИНТЕЗ АНИЗОТРОПИЙНЫХ СУБОПТИМАЛЬНЫХ РЕГУЛЯТОРОВ ЗАДАННОГО ПОРЯДКА НА ОСНОВЕ ПОЛУОПРЕДЕЛЕННОГО ПРОГРАММИРОВАНИЯ И АЛГОРИТМА ПОИСКА ВЗАИМНООБРАТНЫХ МАТРИЦ¹

Чайковский М.М.²

(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Рассматривается задача подавления внешних возмущений для линейной дискретной стационарной системы под воздействием случайных возмущений с неточно известными распределениями. Синтезируемый анизотропийный субоптимальный регулятор представляет собой динамический компенсатор заданного порядка, стабилизирующий замкнутую систему и сохраняющий ее анизотропийную норму ниже заданного порогового значения. Предлагаемый подход к синтезу анизотропийных регуляторов является новым.

Ключевые слова: дискретные линейные стационарные системы, случайные возмущения, статистическая неопределенность, норма, анизотропия, полуопределенное программирование, линейные матричные неравенства, взаимнообратные матрицы.

Введение

Статистическая неопределенность случайных возмущений, рассматриваемая как различие между неточно известным распределением реального шума измерений и его номинальной модели,

¹ Работа выполнена при поддержке РФФИ (грант №11-08-00714-а) и Программы №14 ОЭММПУ РАН.

² Чайковский Михаил Михайлович, кандидат физико-математических наук (mmtchaikovsky@hotmail.com).

может значительно ухудшить качество работы системы управления, если применяемая процедура синтеза регулятора основана на определенном законе распределения возмущения и предположении, что этот закон известен точно. Подобные ситуации могут также возникать из природного непостоянства условий рабочей среды системы управления. Так, \mathcal{H}_2 - и \mathcal{H}_∞ -регуляторы являются полностью эффективными лишь при достаточно точном выполнении базовых гипотез о природе внешних возмущений. Известно, что \mathcal{H}_2 - (или линейно-квадратичный гауссовский) регулятор может оказаться недостаточно эффективным в случае, если внешнее возмущение представляет собой сильно коррелированный шум [23], в то время как \mathcal{H}_∞ -регулятор, проектируемый для наихудшего случая детерминированного возмущения [24], проявляет излишний консерватизм и требует избыточных энергетических затрат на управление, если внешнее возмущение представляет собой белый или слабо коррелированный случайный сигнал.

Одна из первых идей, направленных на преодоление указанного недостатка линейно-квадратичного гауссовского регулятора в случае, когда внешнее возмущение не является гауссовским белым шумом, была представлена в работе [39], посвященной некоторой модификации критерия качества. Эта идея привела к развитию целого класса задач в теории управления — управление системами, чувствительными к рискам [68, 69].

Идеи построения регуляторов, которые сочтали бы положительные качества линейно-квадратичных гауссовских (\mathcal{H}_2 -) и \mathcal{H}_∞ -регуляторов (т.е. минимизировали линейно-квадратичный критерий качества и были бы достаточно робастны) возникли в начале 1990-х годов. В частности, можно выделить подход, связанный с минимизацией \mathcal{H}_2 -нормы замкнутой системы при ограничениях на ее \mathcal{H}_∞ -норму [17], и подход, связанный с минимизацией функционала \mathcal{H}_∞ -энтропии при ограничениях на \mathcal{H}_∞ -норму замкнутой системы [49].

Как показано в [29], задача синтеза регулятора, минимизирующего функционал \mathcal{H}_∞ -энтропии, до известной степени эк-

вивалентна задаче синтеза оптимального регулятора, чувствительного к риску. Множество работ посвящено задачам, связанным с минимизацией функционала \mathcal{H}_∞ -энтропии, см. например [26, 34, 36, 50, 70].

Идеи смешанного $\mathcal{H}_2/\mathcal{H}_\infty$ управления, впервые представленные в [17], были расширены в [57, 71] на основе разделения внешних возмущений на сигналы с ограниченным спектром и ограниченной мощностью и применения смешанного $\mathcal{H}_2/\mathcal{H}_\infty$ критерия качества. Решение задачи стохастического смешанного $\mathcal{H}_2/\mathcal{H}_\infty$ управления для дискретных систем получено в [48].

Во всех перечисленных выше работах применяются методики, основанные на решении уравнений Риккати определенного вида, иногда перекрестно связанных. В [40] смешанная $\mathcal{H}_2/\mathcal{H}_\infty$ задача была рассмотрена в терминах неравенств (а не уравнений) Риккати и решена с помощью выпуклой оптимизации. С тех пор как были разработаны эффективные алгоритмы внутренней точки для решения задач выпуклой оптимизации [18, 51, 52], выпуклая оптимизация стала стандартной стратегией анализа и синтеза систем управления. Методы линейных матричных неравенств (ЛМН) зарекомендовали себя, как мощная и гибкая методика формулирования проектных требований к разрабатываемой системе и синтеза регуляторов, применимая к широкому спектру линейных задач теории управления [19]. После того как решение задачи синтеза \mathcal{H}_∞ -регулятора с помощью ЛМН было получено в [27, 35], полуопределенное программирование успешно применяется для получения эффективных решений многокритериальных $\mathcal{H}_2/\mathcal{H}_\infty$ задач управления [13, 15, 20, 33, 46, 54, 58, 60, 62].

Подход к подавлению неопределенных случайных возмущений на основе минимаксного управления был предложен в середине 1990-х годов С.В. Гусевым в [30]–[32] и впоследствии распространен на случай многомерных систем и синтез регуляторов с заданной структурой методами ЛМН в [59]. Вместо точного знания коэффициентов ковариации возмущения, при применении данного подхода требуется лишь, чтобы коэффициенты ковариации принадлежали известному множеству. Синтезируемый ре-

гулятор минимизирует наихудшую возможную асимптотическую дисперсию выхода для всех таких возмущений. Рассматриваемая задача является промежуточной между экстремальными \mathcal{H}_2 - и \mathcal{H}_∞ -сценариями синтеза и сводится к задаче робастного управления с неопределенностью в сигнале внешнего возмущения [59].

В то же время, другой многообещающий подход на основе стохастического минимакса возник из идей И.Г. Владимиров, разработавшего анизотропийную теорию стохастического робастного управления, представленную в ряде работ [5, 6, 61, 66]. В свете этого подхода, робастность в стохастическом управлении достигается с помощью явного включения различных сценариев распределения шума в единый показатель качества, подлежащий оптимизации; статистическая неопределенность измеряется в терминах энтропии, и показатель робастного качества можно выбрать так, чтобы количественно охарактеризовать возможности системы по подавлению наихудшего внешнего возмущения. Главными понятиями анизотропийной теории стохастического робастного управления являются анизотропия случайного вектора и анизотропийная норма системы.

Функционал анизотропии, введенный И.Г. Владимировым, является энтропийной мерой отклонения вероятностного распределения в евклидовом пространстве от гауссовских распределений с нулевым средним и скалярными ковариационными матрицами. Средняя анизотропия стационарной случайной последовательности определяется как интенсивность анизотропии на единицу времени для достаточно длинных сегментов последовательности. Применительно к случайным возмущениям, действующим на систему, средняя анизотропия характеризует величину статистической неопределенности, понимаемой как несоответствие между неточно известным фактическим распределением шума и семейством номинальных моделей возмущения в виде стационарного дискретного гауссовского белого шума со скалярной ковариационной матрицей [22, 66].

Вторым базовым понятием теории И.Г. Владимирова является a -анизотропийная норма дискретной линейной стационар-

ной системы (ДЛСС), количественно определяющая возможности системы по подавлению возмущений наибольшим отношением мощностной нормы выхода системы к мощностной норме ее входа при условии, что средняя анизотропия входного сигнала не превышает заданного неотрицательного уровня a [22, 66]. Обобщение анизотропийного анализа робастного качества на конечный интервал времени было сделано в [4].

В контексте стохастического робастного управления, направленного на подавление потенциально неблагоприятного воздействия статистической неопределенности, анизотропийный подход предлагает важную альтернативу методам синтеза управления, основанным на точном знании закона распределения случайного внешнего возмущения. Минимизация анизотропийной нормы замкнутой системы в критерии качества приводит к стабилизирующему регулятору по выходу, который проявляет меньший консерватизм по сравнению с \mathcal{H}_∞ -регулятором и является более эффективным при подавлении коррелированных возмущений, чем \mathcal{H}_2 -регулятор [22]. Решение задачи синтеза анизотропийного оптимального регулятора в пространстве состояний, полученное И.Г. Владимировым в [67], основано на решении трех перекрестно связанных алгебраических уравнений Риккати, алгебраического уравнения Ляпунова и уравнения относительно логарифма детерминанта положительно определенной матрицы. Получаемый в результате решения задачи синтеза оценивающий регулятор полного порядка (центральный регулятор) является единственным. Расширение этих результатов на класс объектов с параметрической неопределенностью было получено в [7, 42]. Но решение сложных систем перекрестно связанных уравнений требует разработки и применения специальных вычислительных алгоритмов на основе метода гомотопий [21]. Вместе с тем, применяемая процедура синтеза на основе решения уравнений не направлена на синтез анизотропийных регуляторов пониженного или заданного порядка (а также децентрализованных и многокритериальных регуляторов, регуляторов с заданной структурой), задачи синтеза которых до недавнего времени были

открыты.

Синтез анизотропийных субоптимальных регуляторов является естественным продолжением подхода, предложенного И.Г. Владимировым в [67]. Вместо минимизации анизотропийной нормы системы субоптимальный регулятор стабилизирует замкнутую систему и обеспечивает ограниченность ее анизотропийной нормы заданным значением, т.е. гарантирует подавление случайных внешних возмущений, средняя анизотропия которых не превосходит известного уровня, с качеством не хуже заданного. В отличие от синтеза оптимального анизотропийного регулятора, решение субоптимальных задач синтеза приводит к некоторому семейству регуляторов, таким образом предоставляя дополнительные степени свободы для определения некоторых дополнительных требований к замкнутой системе с целью достижения желаемого качества управления, например, требования заданного расположения полюсов замкнутой системы для достижения желаемого качества переходных процессов. Для решения задачи синтеза анизотропийного субоптимального регулятора требуется критерий проверки ограниченности анизотропийной нормы системы заданным значением. Частотная теорема для анизотропийной нормы, представленная в [43], является стохастическим аналогом известной частотной теоремы для \mathcal{H}_∞ -нормы ДЛСС под воздействием статистически неопределенных стационарных гауссовских возмущений с ограниченной средней анизотропией. Полученный критерий сформулирован в виде неравенства относительно логарифма детерминанта матрицы, выраженной из решения алгебраического уравнения Риккати, зависящего от скалярного параметра. Аналогичный критерий для дискретных линейных нестационарных систем (ДЛНС), сформулированный в виде неравенства, зависящего от дискретного времени, и разностного уравнения Риккати, получен в [47]. Достаточная версия частотной теоремы для анизотропийной нормы была сформулирована в [11, 64, 65] как задача выпуклой оптимизации при ограничениях в виде строгого неравенства относительно детерминанта положительно определенной матрицы и ЛМН. Бы-

ло показано, что ограничение на детерминант линейно зависит от квадрата порогового значения анизотропийной нормы, минимизация которого на выпуклом множестве позволяет вычислять α -анизотропийную норму ДЛСС из решения задачи выпуклой оптимизации [11, 65]. Разработанная процедура анализа является привлекательной с вычислительной точки зрения и легко реализуется средствами некоммерческого программного обеспечения со свободным доступом для выпуклой оптимизации [45, 63]. Предлагаемая вниманию читателя работа направлена на применение современной методологии выпуклой оптимизации и полуопределенного программирования к синтезу анизотропийных субоптимальных регуляторов заданного порядка. Анизотропийные регуляторы являются привлекательной и перспективной альтернативой \mathcal{H}_2 - и \mathcal{H}_∞ -регуляторам в задачах подавления случайных внешних возмущений с неточно известными распределениями. В сравнении с решением задачи синтеза анизотропийного оптимального регулятора в пространстве состояний, полученным ранее в [67], новый подход на основе численной оптимизации является не требует разработки и применения специальных вычислительных алгоритмов на основе метода гомотопий [21].

Структура работы следующая. В разделе 1 изложена постановка общей задачи синтеза анизотропийного субоптимального регулятора заданного порядка. В разделе 2 представлено общее решение задачи синтеза. В разделе 3 рассматривается ряд вычислительных примеров. Заключительные замечания и выводы даны в разделе 4.

ОБОЗНАЧЕНИЯ

Множество вещественных чисел обозначается \mathbb{R} , множество $(n \times m)$ -матриц — $\mathbb{R}^{n \times m}$. Для комплексной матрицы $M = [m_{ij}]$, M^* обозначает эрмитово сопряжение этой матрицы $M^* := [m_{ji}^*]$. Для вещественной матрицы $M = [m_{ij}]$, M^T обозначает транспонирование: $M^T := [m_{ji}]$. Для вещественных симметричных матриц $M \succ N$ означает, что матрица $M - N$ положительно определена. В блочно-симметричных матрицах символ $*$ заменяет блоки, вид которых определяется симметрией. Спектраль-

ный радиус матрицы M обозначается $\rho(M) := \max_k |\lambda_k(M)|$, где $\lambda_k(M)$ — k -е собственное значение матрицы M . Максимальное сингулярное значение комплексной матрицы M обозначается $\bar{\sigma}(M) := \sqrt{\lambda_{\max}(M^*M)}$. I_n обозначает единичную матрицу размерности $(n \times n)$, $0_{n \times m}$ — нулевую $(n \times m)$ -матрицу. Размерности нулевых матриц в случаях, когда их нетрудно понять из контекста, указываться не будут.

Угловое граничное значение матричной передаточной функции $F(z)$, аналитической в единичном диске комплексной плоскости $|z| < 1$, обозначается $\widehat{F}(\omega) := F(e^{i\omega})$.

$\mathcal{H}_2^{p \times m}$ и $\mathcal{H}_\infty^{p \times m}$ обозначают пространства Харди $(p \times m)$ -матричных передаточных функций $F(z)$ комплексной переменной z , аналитических в единичном диске $|z| < 1$, с ограниченной \mathcal{H}_2 - и \mathcal{H}_∞ -нормой соответственно:

$$\|F\|_2 := \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \text{tr}(\widehat{F}(\omega)\widehat{F}^*(\omega))d\omega \right)^{1/2} < +\infty,$$

$$\|F\|_\infty := \sup_{|z| \leq 1} \bar{\sigma}(F(z)) = \text{ess sup}_{-\pi \leq \omega \leq \pi} \bar{\sigma}(\widehat{F}(\omega)) < +\infty.$$

1. Постановка задачи синтеза

Объект управления представлен дискретной линейной стационарной моделью $P(z)$ с n_x -мерным состоянием X , m_w -мерным входом возмущения W , m_u -мерным входом управления U , p_z -мерным управляемым выходом Z и p_y -мерным измеряемым выходом Y . Предполагается, что все сигналы представляют собой двусторонние бесконечные векторные дискретные последовательности, связанные уравнениями

$$(1) \quad P(z) : \begin{bmatrix} x_{k+1} \\ z_k \\ y_k \end{bmatrix} = \begin{bmatrix} A & B_w & B_u \\ C_z & D_{zw} & D_{zu} \\ C_y & D_{yw} & 0 \end{bmatrix} \begin{bmatrix} x_k \\ w_k \\ u_k \end{bmatrix},$$

где размерности всех матриц согласованы, $p_z \leq m_w$, пара матриц (A, B_u) является стабилизируемой, а пара (A, C_y) — детектируемой.

Предполагается, что внешнее возмущение $W = (w_k)_{-\infty < k < +\infty}$ является стационарной последовательностью случайных векторов w_k с нулевым средним $\mathbf{E}w_k = 0$, неизвестной ковариационной матрицей $\mathbf{E}w_k w_k^T = \Sigma_W \succ 0$ и гауссовской плотностью распределения вероятности

$$p(w_k) := (2\pi)^{-m_1/2} (\det \Sigma_W)^{-1/2} \exp\left(-\frac{1}{2} \|w_k\|_{\Sigma_W^{-1}}^2\right),$$

где $\|w_k\|_{\Sigma_W^{-1}} = \sqrt{w_k^T \Sigma_W^{-1} w_k}$ и \mathbf{E} обозначает математическое ожидание. Обозначим пространство стационарных в узком смысле последовательностей интегрируемых с квадратом случайных векторов [7] через

$$\ell_{\mathcal{P}}^m = \{W = (w_k)_{-\infty < k < +\infty} : w_k \in \mathbb{L}_2^m \wedge \|W\|_{\mathcal{P}} < +\infty\},$$

где мощностная норма последовательности случайных векторов $W = (w_k)_{-\infty < k < +\infty}$ определяется как [7, 71]

$$\|W\|_{\mathcal{P}} := \left(\lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{k=-N}^N \mathbf{E}|w_k|^2 \right)^{1/2}$$

и может быть вычислена через спектральную плотность $S(\omega)$ этой последовательности:

$$\|W\|_{\mathcal{P}} = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \text{tr} S(\omega) d\omega \right)^{1/2}.$$

Обозначим множество входных сигналов с ограниченной средней анизотропией через

$$\mathcal{W}_a := \{W \in \ell_{\mathcal{P}}^m : \overline{\mathbf{A}}(W) \leq a\},$$

где

$$\overline{\mathbf{A}}(W) = -\frac{1}{4\pi} \int_{-\pi}^{\pi} \ln \det \left(\frac{m_{\omega} S(\omega)}{\|W\|_{\mathcal{P}}^2} \right) d\omega$$

— функционал средней анизотропии [22, 66].

Задача синтеза — найти регулятор заданного порядка по измеряемому выходу в форме динамического компенсатора

$$(2) \quad K(z) : \begin{bmatrix} \xi_{k+1} \\ u_k \end{bmatrix} = \begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix} \begin{bmatrix} \xi_k \\ y_k \end{bmatrix}$$

с n_ξ -мерным состоянием $\Xi = (\xi_k)_{-\infty < k < +\infty}$, стабилизирующий замкнутую систему (рис. 1) и гарантирующий некоторый желаемый уровень робастного качества подавления внешних возмущений.

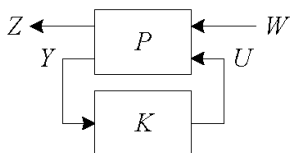


Рис. 1. Замкнутая система

Известно (см. например [35]), что задачу синтеза динамического регулятора заданного порядка можно представить в виде задачи синтеза статического регулятора по выходу, дополнив вектор состояния объекта управления состояниями регулятора:

$$(3) \quad \begin{bmatrix} A & B_w & B_u \\ C_z & D_{zw} & D_{zu} \\ C_y & D_{yw} & 0 \end{bmatrix} := \left[\begin{array}{cc|cc} A & 0 & B_w & 0 & B_u \\ 0 & 0 & 0 & I_{n_\xi} & 0 \\ \hline C_z & 0 & D_{zw} & 0 & D_{zu} \\ 0 & I_{n_\xi} & 0 & 0 & 0 \\ \hline C_y & 0 & D_{yw} & 0 & 0 \end{array} \right].$$

Реализация замкнутой системы с расширенным объектом управления (3) имеет вид

$$\begin{aligned} \begin{bmatrix} A & B \\ C & D \end{bmatrix} &= \begin{bmatrix} A & B_w \\ C_z & D_{zw} \end{bmatrix} + \begin{bmatrix} B_u \\ D_{zu} \end{bmatrix} K \begin{bmatrix} C_y & D_{yw} \end{bmatrix} = \\ &= \begin{bmatrix} A + B_u K C_y & B_w + B_u K D_{zw} \\ C_z + D_{zw} K C_y & D_{zw} + D_{zu} K D_{yw} \end{bmatrix}, \end{aligned}$$

где матрица K включает матрицы параметров регулятора:

$$(4) \quad K := \begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix}.$$

Поэтому в дальнейшем мы будем предполагать, что модель объекта управления (1) представляет собой расширенную реализацию (3) с n -мерным вектором состояния, и искать анизотропийный субоптимальный регулятор в виде статической обратной связи по выходу (4).

Предполагается, что для расширенного объекта управления (3) и регулятора (4) выполняется условие Кимуры [41] нулевого порядка

$$n - m_u - p_y < 0,$$

где в общем случае $n = n_x + n_\xi$. Выполнение этого условия гарантирует существование стабилизирующего статического регулятора по выходу для расширенной реализации (3).

Пусть $T_{zw}(z)$ — матричная передаточная функция замкнутой системы от возмущения W к управляемому выходу Z , заданная нижним дробно-линейным преобразованием пары (P, K) :

$$(5) \quad T_{zw}(z) = \mathcal{F}_l(P, K) = P_{zw} + P_{zu}K(I_{p_y} - P_{yu}K)^{-1}P_{yw},$$

где

$$(6) \quad P_{zw}(z) \sim \left[\begin{array}{c|c} A & B_w \\ \hline C_z & D_{zw} \end{array} \right], \quad P_{zu}(z) \sim \left[\begin{array}{c|c} A & B_u \\ \hline C_z & D_{zu} \end{array} \right],$$

$$P_{yw}(z) \sim \left[\begin{array}{c|c} A & B_w \\ \hline C_y & D_{yw} \end{array} \right], \quad P_{yu}(z) \sim \left[\begin{array}{c|c} A & B_u \\ \hline C_y & 0 \end{array} \right].$$

Напомним, что a -анизотропийная норма передаточной функции $T_{zw}(z) \in \mathcal{H}_\infty^{p_z \times m_w}$ количественно характеризует возможности замкнутой системы по подавлению внешних возмущений наибольшим отношением мощностной нормы выхода этой системы к мощностной норме входа при условии, что средняя анизотропия возмущения не превосходит уровня a [7, 22, 66]:

$$(7) \quad \|T_{zw}\|_a := \sup_{W \in \mathcal{W}_a} \frac{\|Z\|_{\mathcal{P}}}{\|W\|_{\mathcal{P}}}.$$

Из [22, 66] известно, что a -анизотропийная норма заданной системы $F \in \mathcal{H}_\infty^{p_z \times m_w}$ является неубывающей функцией уровня средней анизотропии a , удовлетворяющей неравенству

$$(8) \quad \frac{1}{\sqrt{m_w}} \|T_{zw}\|_2 = \|T_{zw}\|_0 \leq \lim_{a \rightarrow +\infty} \|T_{zw}\|_a = \|T_{zw}\|_\infty.$$

Эти выражения показывают, что \mathcal{H}_2 - и \mathcal{H}_∞ -нормы являются предельными случаями a -анизотропийной нормы при $a \rightarrow 0, +\infty$, соответственно.

Общая постановка задачи синтеза анизотропийного субоптимального регулятора заданного порядка следующая.

Задача синтеза. Для заданных объекта управления P с моделью в пространстве состояний (1), уровня средней анизотропии $a \geq 0$ входного возмущения W и некоторого желаемого порогового значения $\gamma > 0$ найти регулятор в виде статической обратной связи по измеряемому выходу

$$(9) \quad u_k = Ky_k,$$

стабилизирующий замкнутую систему и гарантирующий, что ее a -анизотропийная норма не превосходит порогового значения γ :

$$(10) \quad \|T_{zw}\|_a < \gamma.$$

2. Решение задачи синтеза

В этом разделе приводится решение общей задачи синтеза регулятора заданного порядка. Для решения задачи синтеза применяется критерий проверки условия ограниченности анизотропийной нормы системы заданным пороговым значением для модели в пространстве состояний. Этот критерий, называемый частотной теоремой для анизотропийной нормы, был недавно представлен в [11, 64]. Чтобы применить данный критерий к задаче синтеза, его требуется переформулировать.

2.1. ЧАСТОТНАЯ ТЕОРЕМА ДЛЯ АНИЗОТРОПИЙНОЙ НОРМЫ В ЗАДАЧЕ СИНТЕЗА

Для объекта управления P и регулятора K , определенных выше, реализация замкнутой системы имеет вид

$$(11) \quad T_{zw}(z) : \begin{bmatrix} \chi_{k+1} \\ z_k \end{bmatrix} = \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \begin{bmatrix} \chi_k \\ w_k \end{bmatrix},$$

где $\chi_k \in \mathbb{R}^n$, $n = n_x + n_\xi$,

$$(12) \quad \left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C} & \mathcal{D} \end{array} \right] = \left[\begin{array}{c|c} A + B_u K C_y & B_w + B_u K D_{yw} \\ \hline C_z + D_{zu} K C_y & D_{zw} + D_{zu} K D_{yw} \end{array} \right].$$

В [64, 65] показано, что для заданных $a \geq 0$, $\gamma > 0$ неравенство (10) выполняется, если существует $\eta > \gamma^2$, такое что неравенство

$$(13) \quad \eta - (e^{-2a} \det(\eta I_{m_w} - \mathcal{B}^T \Phi \mathcal{B} - \mathcal{D}^T \mathcal{D}))^{1/m_w} < \gamma^2$$

выполняется для некоторой вещественной $(n \times n)$ -матрицы $\Phi = \Phi^T \succ 0$, удовлетворяющей ЛМН

$$(14) \quad \begin{bmatrix} \mathcal{A}^T \Phi \mathcal{A} - \Phi + \mathcal{C}^T \mathcal{C} & \mathcal{A}^T \Phi \mathcal{B} + \mathcal{C}^T \mathcal{D} \\ \mathcal{B}^T \Phi \mathcal{A} + \mathcal{D}^T \mathcal{C} & \mathcal{B}^T \Phi \mathcal{B} + \mathcal{D}^T \mathcal{D} - \eta I_{m_w} \end{bmatrix} \prec 0.$$

Условия (13), (14) частотной теоремы для анизотропийной нормы [64, 65] невозможно непосредственно применить для решения поставленной задачи синтеза из-за перекрестных произведений неизвестной матрицы Ляпунова Φ и матриц реализации замкнутой системы $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$, аффинно зависящих от K , которые также возникают в (13).

Преодолеть указанную трудность позволяет введение вспомогательной переменной, вещественной $(m_w \times m_w)$ -матрицы $\Psi = \Psi^T \succ 0$, удовлетворяющей неравенствам

$$(15) \quad \eta - (e^{-2a} \det \Psi)^{1/m_w} < \gamma^2, \quad \Psi \prec \eta I_{m_w} - \mathcal{B}^T \Phi \mathcal{B} - \mathcal{D}^T \mathcal{D},$$

что эквивалентно (13). Чтобы избавиться от произведений матриц Φ , \mathcal{B} и \mathcal{D} , перепишем последнее неравенство в (15) в виде

$$\Psi - \eta I_{m_w} - \begin{bmatrix} \mathcal{B}^T & \mathcal{D}^T \end{bmatrix} \begin{bmatrix} -\Phi^{-1} & 0 \\ 0 & -I_{p_z} \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{B} \\ \mathcal{D} \end{bmatrix} \prec 0,$$

где $\begin{bmatrix} -\Phi^{-1} & 0 \\ 0 & -I_{p_z} \end{bmatrix} \prec 0$, что эквивалентно

$$(16) \quad \begin{bmatrix} \Psi - \eta I_{m_w} & \mathcal{B}^T & \mathcal{D}^T \\ \mathcal{B} & -\Phi^{-1} & 0 \\ \mathcal{D} & 0 & -I_{p_z} \end{bmatrix} \prec 0$$

в силу леммы Шура (см. например [19]).

Чтобы избавиться от перекрестных произведений матриц Φ , \mathcal{A} , and \mathcal{B} в (14), представим это неравенство в виде

$$\begin{bmatrix} -\Phi + \mathcal{C}^T \mathcal{C} & \mathcal{C}^T \mathcal{D} \\ \mathcal{D}^T \mathcal{C} & \eta I_{m_w} + \mathcal{D}^T \mathcal{D} \end{bmatrix} - \begin{bmatrix} \mathcal{A}^T \\ \mathcal{B}^T \end{bmatrix} (-\Phi^{-1})^{-1} \begin{bmatrix} \mathcal{A} & \mathcal{B} \end{bmatrix} \prec 0,$$

где $-\Phi^{-1} \prec 0$. В силу леммы Шура, предыдущее неравенство эквивалентно

$$(17) \quad \begin{bmatrix} -\Phi + \mathcal{C}^T \mathcal{C} & \mathcal{C}^T \mathcal{D} & \mathcal{A}^T \\ \mathcal{D}^T \mathcal{C} & \mathcal{D}^T \mathcal{D} - \eta I_{m_w} & \mathcal{B}^T \\ \mathcal{A} & \mathcal{B} & -\Phi^{-1} \end{bmatrix} \prec 0.$$

Чтобы избавиться от перекрестных произведений матриц \mathcal{C} и \mathcal{D} , представим неравенство (17) в виде

$$\begin{bmatrix} -\Phi & 0 & \mathcal{A}^T \\ 0 & -\eta I_{m_w} & \mathcal{B}^T \\ \mathcal{A} & \mathcal{B} & -\Phi^{-1} \end{bmatrix} - \begin{bmatrix} \mathcal{C}^T \\ \mathcal{D}^T \\ 0 \end{bmatrix} (-I_{p_z})^{-1} \begin{bmatrix} \mathcal{C} & \mathcal{D} & 0 \end{bmatrix} \prec 0$$

где $-I_{p_z} \prec 0$. Повторное применение леммы Шура к предыдущему неравенству и обозначение $\Pi := \Phi^{-1}$ приводят к следующей формулировке частотной теоремы для анизотропийной нормы в терминах взаимнообратных матриц.

Лемма 1. Пусть $T_{zw} \in \mathcal{H}_\infty^{p_z \times m_w}$ — матричная передаточная функция системы с реализацией (11), где $\rho(\mathcal{A}) < 1$. Анизотропийная норма (7) системы T_{zw} строго ограничена заданным пороговым значением $\gamma > 0$, т.е. $\|T_{zw}\|_a < \gamma$, если система неравенств

$$(18) \quad \eta - (e^{-2a} \det \Psi)^{1/m_w} < \gamma^2,$$

$$(19) \quad \begin{bmatrix} \Psi - \eta I_{m_w} & \mathcal{B}^T & \mathcal{D}^T \\ \mathcal{B} & -\Pi & 0 \\ \mathcal{D} & 0 & -I_{p_z} \end{bmatrix} \prec 0,$$

$$(20) \quad \begin{bmatrix} -\Phi & 0 & \mathcal{A}^T & \mathcal{C}^T \\ 0 & -\eta I_{m_w} & \mathcal{B}^T & \mathcal{D}^T \\ \mathcal{A} & \mathcal{B} & -\Pi & 0 \\ \mathcal{C} & \mathcal{D} & 0 & -I_{p_z} \end{bmatrix} \prec 0,$$

$$(21) \quad \eta > \gamma^2, \quad \Psi \succ 0, \quad \Phi \succ 0, \quad \Pi \succ 0$$

разрешима относительно скалярной переменной η , вещественных $(m_w \times m_w)$ -матрицы Ψ и двух взаимнообратных $(n \times n)$ -матриц Φ , Π , удовлетворяющих условию

$$(22) \quad \Phi \Pi = I_n.$$

Замечание 1. Неравенство (18) представляет собой выпуклое ограничение относительно вспомогательной переменной Ψ . Известно, что

1. Функция $(\det \Psi)^p$ ($m \times m$)-матрицы $\Psi = \Psi^T \succcurlyeq 0$ является вогнутой по своему аргументу для любого $0 \leq p \leq \frac{1}{m}$ [16].

2. Функция $(\det \Psi)^{1/m}$ ($m \times m$)-матрицы $\Psi = \Psi^T \succcurlyeq 0$ есть не что иное как среднее геометрическое собственных значений этой матрицы $\sqrt[m]{\lambda_1(\Psi) \dots \lambda_m(\Psi)}$.

3. В [16, с. 105] было показано, что подграфик геометрического среднего двух неотрицательных величин, множество

$$\{(\lambda_1, \lambda_2, t) \in \mathbb{R}^3 \mid x_1, x_2 \geq 0, t \leq \sqrt{\lambda_1 \lambda_2}\}$$

представимо в виде конуса второго порядка

$$\left\{ (\lambda_1, \lambda_2, t) \mid \exists \tau : t \leq \tau; \tau \geq 0, \left\| \begin{bmatrix} \tau \\ \frac{\lambda_1 - \lambda_2}{2} \end{bmatrix} \right\|_2 \leq \frac{\lambda_1 + \lambda_2}{2} \right\},$$

а подграфик геометрического среднего 2^l неотрицательных величин, множество

$$\left\{ (\lambda_1, \dots, \lambda_{2^l}, t) \in \mathbb{R}^{2^l+1} \mid \lambda_i \geq 0, i = 1, \dots, 2^l, t \leq (\lambda_1 \lambda_2 \dots \lambda_{2^l})^{1/2^l} \right\}$$

также представимо в виде пересечения конечного числа конусов второго порядка.

4. В [16, с. 147] было показано, что если p — рациональное число, $0 \leq p \leq \frac{1}{m}$, то выпуклая функция $-(\det \Psi)^p$ ($m \times m$)-матрицы $\Psi = \Psi^T \succcurlyeq 0$ представима в виде ЛМН. А именно, множество

$$\{(\Psi, t) \mid \Psi = \Psi^T \succcurlyeq 0, t \leq (\det \Psi)^p\}$$

представимо в виде

$$\left\{ (\Psi, t) \mid \Psi = \Psi^T \succcurlyeq 0, \begin{bmatrix} \Psi & \Delta \\ \Delta^T & \text{diag } \Delta \end{bmatrix} \succcurlyeq 0, t \leq (\delta_1 \dots \delta_m)^p \right\},$$

где Δ — нижняя треугольная $(m \times m)$ -матрица, составленная из вспомогательных переменных с диагональными элементами δ_i . Подграфик вогнутого одночлена $t \leq (\delta_1 \dots \delta_m)^p$ представим в виде конуса второго порядка [16, с. 108] и, следовательно, в виде ЛМН.

Замечание 2. Систему неравенств (18)–(21) леммы 1 можно решить с помощью доступных свободно распространяемых программных пакетов для решения задач выпуклой оптимизации, позволяющих использовать выпуклую функцию $-(\det(\Psi))^{1/m}$ $(m \times m)$ -матрицы $\Psi \succcurlyeq 0$ не только в качестве целевой функции, но и в качестве ограничения. Такими программными средствами являются, например, интерфейс YALMIP [45] в сочетании с программой-решателем SeDuMi [63] для систем Matlab и Scilab. В интерфейсе YALMIP функции $-(\det(\Psi))^{1/m}$ соответствует команда `geomean`, возвращающая среднее геометрическое собственных чисел положительно-определенной матрицы [45].

Замечание 3. Проверка выполнения условия $\|T_{zw}\|_a < \gamma$ сводится к поиску положительной скалярной величины η и двух взаимнообратных матриц $\Phi \succ 0$, $\Pi \succ 0$, $\Phi\Pi = I_n$, удовлетворяющих ЛМН (19), (20) и выпуклому ограничению (18). Для численного решения такой задачи могут применяться известные алгоритмы, разработанные в [1, 12, 14, 25, 28, 38, 53, 55, 56] для поиска взаимнообратных матриц, удовлетворяющих выпуклым ограничениям.

2.2. АЛГОРИТМ ПОИСКА ВЗАИМНООБРАТНЫХ МАТРИЦ НА ОСНОВЕ МЕТОДА УСЛОВНОГО ГРАДИЕНТА

Взаимнообратные матрицы $\Phi = \Phi^T \succ 0$, $\Pi = \Pi^T \succ 0$, $\Phi\Pi = I$, удовлетворяющие ЛМН (19), (20) и выпуклому ограничению (18), можно найти с помощью алгоритма на основе метода условного градиента [8, 14, 55], согласно которому задача (18)–(22) рассматривается как задача минимизации вогнутой целевой

функции

$$(23) \quad \begin{aligned} & \operatorname{tr}(\Pi - \Phi^{-1}) \rightarrow \min \\ & \text{на множестве } \Psi, \Phi, \Pi, \eta, \\ & \text{удовлетворяющих ограничениям (18)–(21) и} \end{aligned}$$

$$(24) \quad \begin{bmatrix} \Pi & I \\ I & \Phi \end{bmatrix} \succcurlyeq 0.$$

Целевая функция $f(\Pi, \Phi) = \operatorname{tr}(\Pi - \Phi^{-1})$ является вогнутой по Π и Φ , ее градиент определяется выражениями [55]

$$\nabla_{\Pi} f(\Pi, \Phi) = I, \quad \nabla_{\Phi} f(\Pi, \Phi) = \Phi^{-2}.$$

Алгоритм состоит из следующих шагов:

Алгоритм 1.

1) $k = 0$. Выбираются начальные условия $\Pi_0 > 0$, $\Phi_0 > 0$, удовлетворяющие неравенствам (18)–(21) для некоторых значений Ψ, η .

2) Решается локальная задача выпуклого программирования

$$(25) \quad \begin{aligned} & \operatorname{tr}(\Pi + \Phi_k^{-1} \Phi \Phi_k^{-1}) \rightarrow \min \\ & \text{на множествах } \Psi, \Phi, \Pi, \eta, \\ & \text{удовлетворяющих ограничениям (18)–(21) и} \end{aligned}$$

$$\begin{bmatrix} \Pi & I \\ I & \Phi \end{bmatrix} \succcurlyeq 0.$$

3) Если задача (25) разрешима и неизвестные переменные вычислены, проверяются условия остановки

$$(26) \quad |\operatorname{tr}(\Pi + \Phi_k^{-1} \Phi \Phi_k^{-1})| < \varepsilon, \quad k_N \leq k,$$

где ε — заданная точность, k_N — заданное максимальное число итераций. Если одно из условий (26) выполняется, алгоритм останавливается.

4) $k := k + 1$; $\Phi_k := \Phi$; осуществляется переход к шагу 2.

Известно [8, 14, 55], что целевая функция $f(\Pi, \Phi)$ в алгоритме 1 монотонно убывает и итерации сходятся к точке локального минимума (Π_*, Φ_*) , удовлетворяющей необходимому условию оптимальности [8]

$$\begin{aligned} \langle \nabla f(\Pi_*, \Phi_*), \text{blockdiag}(\Pi - \Pi_*, \Phi - \Phi_*) \rangle = \\ = \text{tr}(\nabla f(\Pi_*, \Phi_*) \text{blockdiag}(\Pi - \Pi_*, \Phi - \Phi_*)) = \\ = \text{tr}(\Pi - \Pi_* + \Phi_*^{-2}\Phi - \Phi_*^{-1}) \geq 0 \\ \forall \Pi, \Phi, \text{удовлетворяющих (18)–(21), (24)}. \end{aligned}$$

В [8] показано, что скорость сходимости метода условного градиента зависит от многих факторов, включая свойства целевой функции и множества решений (гладкость, выпуклость, сильная выпуклость и т.п.). Например, для задачи минимизации вогнутой целевой функции при наличии линейных ограничений вида неравенств метод условного градиента конечен [55].

2.3. СТАТИЧЕСКАЯ ОБРАТНАЯ СВЯЗЬ ПО ВЫХОДУ

Прямое применение достаточных условий (18)–(20) леммы 1 к реализации замкнутой системы (12) приводит к следующему прямому решению задачи синтеза.

Следствие 1. Для заданных $a \geq 0$, $\gamma > 0$ статический регулятор по выходу (9), являющийся решением задачи синтеза, существует, если система неравенств

$$(27) \quad \eta - (e^{-2a} \det \Psi)^{1/m_w} < \gamma^2,$$

$$(28) \quad \begin{bmatrix} \Psi - \eta I_{m_w} & * & * \\ B_w + B_u K D_{yw} & -\Pi & * \\ D_{zw} + D_{zu} K D_{yw} & 0 & -I_{p_z} \end{bmatrix} \prec 0,$$

$$(29) \quad \begin{bmatrix} -\Phi & * & * & * \\ 0 & -\eta I_{m_w} & * & * \\ A + B_u K C_y & B_w + B_u K D_{yw} & -\Pi & * \\ C_z + D_{zu} K C_y & D_{zw} + D_{zu} K D_{yw} & 0 & -I_{p_z} \end{bmatrix} \prec 0,$$

$$(30) \quad \eta > \gamma^2, \quad \Psi \succ 0, \quad \Phi \succ 0, \quad \Pi \succ 0$$

разрешима относительно скалярной переменной η , вещественных $(m_w \times m_w)$ -матрицы Ψ , $(m_u \times p_y)$ -матрицы K и двух взаимнообратных $(n \times n)$ -матриц Φ , Π , удовлетворяющих условию

$$(31) \quad \Phi\Pi = I_n.$$

Замечание 4. Матрица параметров регулятора K непосредственно входит в неравенства синтеза (28), (29), что позволяет накладывать на нее дополнительные структурные требования для синтеза, например, децентрализованного управления с блочно-диагональной матрицей K или регулятора заданной структуры из решения задачи (27)–(31).

Задача вычисления матрицы статической обратной связи по выходу K , являющейся решением задачи синтеза анизотропийного субоптимального регулятора, сводится к задаче поиска взаимнообратных матриц, удовлетворяющих ограничениям (27)–(30), которая может быть решена с помощью алгоритма 1.

3. Вычислительный пример: управление самолетом при заходе на посадку

В качестве вычислительного примера рассмотрим решение задачи управления продольным движением самолета при заходе на посадку по глиссаде с заданным углом наклона в условиях сдвига ветра при наличии шума измерений. Эта задача решена в [44] с помощью анизотропийного оптимального регулятора полного порядка. Полученный регулятор стабилизирует линеаризованную модель объекта управления в отклонениях от желаемых значений переменных состояния при движении по глиссаде с заданным углом наклона в присутствии детерминированного (сдвиг ветра) и стохастических возмущений.

3.1. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ПРОДОЛЬНОГО ДВИЖЕНИЯ САМОЛЕТА. ПОСТАНОВКА ЗАДАЧИ УПРАВЛЕНИЯ

Продольное движение самолета с учетом ветровых возмущений в скоростной системе координат (касательная и нормаль к

траектории полета) описывается следующей системой нелинейных дифференциальных уравнений [2, 3]

$$(32) \quad \begin{cases} m\dot{V} &= T \cos \alpha - D - mg \sin \theta - m(\dot{w}_x \cos \theta + \dot{w}_y \sin \theta), \\ mV\dot{\theta} &= T \sin \alpha + L - mg \cos \theta + m(\dot{w}_x \sin \theta - \dot{w}_y \cos \theta), \\ J_z \dot{\omega}_z &= M_z, \\ \dot{\vartheta} &= \omega_z, \end{cases}$$

где m — масса самолета; V — воздушная скорость; T — сила тяги; α — угол атаки; D — сила лобового сопротивления; g — ускорение свободного падения; θ — угол наклона траектории полета; \dot{w}_x и \dot{w}_y — полные градиенты горизонтальной и вертикальной составляющих скорости ветра в инерциальной системе отсчета соответственно; L — подъемная сила; J_z — момент инерции относительно поперечной оси z самолета; ω_z — угловая скорость относительно поперечной оси z самолета; M_z — момент тангажа и $\vartheta = \alpha + \theta$ — угол тангажа (см. рис. 2).

Эти уравнения справедливы в предположении, что самолет жесткий, направление силы тяги совпадает с осью самолета, масса самолета постоянна, Земля плоская, ветер стационарный [3]. Модель (32) также не содержит аналитических зависимостей для силы лобового сопротивления D , подъемной силы L , момента инерции самолета J_z и момента тангажа M_z . Предполагается, что значения этих переменных являются табличными значениями, полученными в результате экспериментов, и выбираются из соответствующих таблиц при линеаризации нелинейной модели (32) [9, 10].

Сила тяги T и угол атаки α являются переменными управления в уравнениях (32) и в свою очередь зависят от отклонения сектора газа δ_t и обобщенного руля высоты самолета δ_e соответственно. Таким образом, управление самолетом в продольной плоскости реализуется с помощью обобщенного руля высоты δ_e и сектора газа δ_t .

Дифференциальное уравнение для высоты центра масс самолета имеет вид

$$(33) \quad \dot{h} = V \sin \theta + w_y.$$

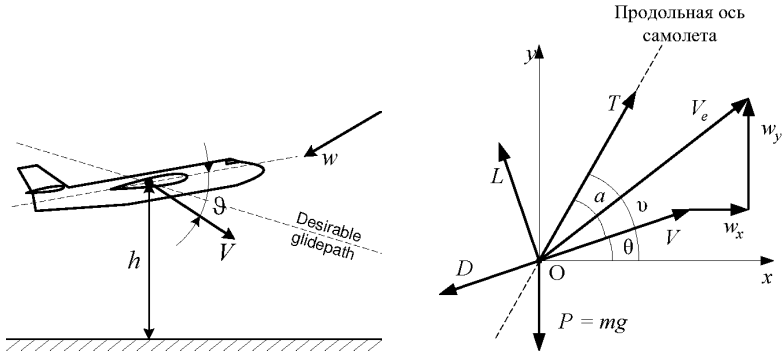


Рис. 2. К задаче управления продольным движением самолета.
Система координат и переменные

Динамика двигателя описывается следующим уравнением

$$(34) \quad \Delta \dot{T} = \frac{1}{T_e} (-\Delta T + K_e \Delta \delta_t),$$

где T_e — постоянная времени двигателя; K_e — некоторый заданный числовой коэффициент и $\Delta \delta_t$ — отклонение сектора газа от предписанного значения.

Отклонение обобщенного руля высоты $\Delta \delta_t$ с учетом контура короткопериодического движения формируется следующим образом

$$\Delta \delta_e = K_{\omega_z} \Delta \omega_z + K_{\vartheta} \Delta \vartheta + K_{c\vartheta} \Delta \vartheta_{c\vartheta},$$

где K_{ω_z} , K_{ϑ} , и $K_{c\vartheta}$ — некоторые заданные числовые коэффициенты; $\Delta \vartheta_{c\vartheta}$ — сигнал управления, формируемый регулятором.

Анизотропный, линейно-квадратичный гауссовский и \mathcal{H}_∞ -регуляторы синтезированы для модели самолета Ту-154 при заходе на посадку по глиссаде с углом наклона $\theta_0 = -2,7$ град. Нелинейные уравнения (32)–(34), описывающие продольное движение самолета, линеаризованы в точке траектории

$$\begin{aligned} V_0 &= 71,375 \text{ м/сек}, & \theta_0 &= -2,7 \text{ град}, & \omega_{z0} &= 0 \text{ сек}^{-1} \\ \vartheta_0 &= 0 \text{ град}, & h_0 &= 600 \text{ м}, & T_0 &= 52540 \text{ Н}. \end{aligned}$$

Стандартная линеаризованная дискретная стационарная модель объекта управления (1) была получена для значения шага дис-

кретизации 0,01 с, имеет порядок $n_x = 6$. В окрестности заданной глissады продольное движение самолета аппроксимируется дискретной линеаризованной моделью в отклонениях (1), где

$$\begin{aligned} x_k &= [\Delta V_k \quad \Delta \theta_k \quad \Delta \omega_{z,k} \quad \Delta \vartheta_k \quad \Delta h_k \quad \Delta T_k]^T, \\ w_k &= [n_{u_1,k} \quad n_{u_2,k} \quad n_{y_1,k} \quad n_{y_2,k}]^T, \\ u_k &= [\Delta \theta_{cy,k} \quad \Delta \delta_{t,k}]^T, \\ z_k &= [\Delta V_k \quad \Delta h_k \quad \Delta \vartheta_{cy,k} \quad \Delta \delta_{t,k}]^T, \\ y_k &= [\Delta V_k + n_{y_1,k} \quad \Delta h_k + n_{y_2,k}]^T, \end{aligned}$$

где V_k — воздушная скорость самолета; θ_k — угол наклона траектории; $\omega_{z,k}$ — угловая скорость тангажа; ϑ_k — угол тангажа; h_k — высота центра масс; T_k — тяга двигателей; $\theta_{cy,k}$ — управление обобщенными рулями высоты; $\delta_{t,k}$ — управление сектором газа; $n_{u_1,k}, n_{u_2,k}, n_{y_1,k}, n_{y_2,k}$ — шумы приводов и измерений. Матрицы реализации модели в пространстве состояний имеют вид

$$\begin{aligned} A &= \begin{bmatrix} 0,9994 & -0,0008 & 0 & -0,0009 & 0 & 0,0009 \\ 0,0022 & 0,9938 & 0,0011 & 0,0072 & 0 & 0 \\ 0,0001 & 0,0052 & 0,9842 & -0,0154 & 0 & 0 \\ 0 & 0 & 0,0099 & 0,9999 & 0 & 0 \\ -0,0005 & 0,0124 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,9960 \end{bmatrix}, \\ B_u &= \begin{bmatrix} 0 & 0 \\ -0,0012 & 0 \\ 0,0117 & 0 \\ 0,0001 & 0 \\ 0 & 0 \\ 0 & 0,004 \end{bmatrix}, \quad C_y = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \\ B_w &= \begin{bmatrix} 0 & -0,01 & 0,0005 & 0 & 0 \\ 0 & -0,0004 & -0,008 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0,01 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad C_z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \\ D_{yw} &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad D_{zu} = \begin{bmatrix} 0 \\ I_2 \end{bmatrix}. \end{aligned}$$

3.2. СТАТИЧЕСКАЯ ОБРАТНАЯ СВЯЗЬ ПО ВЫХОДУ

Для задачи управления посадкой самолета Ту-154 в условиях сдвига ветра был синтезирован анизотропийный субоптимальный регулятор в виде статической обратной связи по измеряемому выходу. Матрица коэффициента усиления регулятора вычислялась из решения системы неравенств (27)–(30) при ограничении (37) в соответствии с алгоритмом 1 для заданного уровня средней анизотропии внешнего возмущения $a = 0,1$ и фиксированного порогового значения анизотропийной нормы замкнутой системы $\gamma = 20$. При синтезе регулятора учитывалось дополнительное требование расположения полюсов замкнутой системы в выпуклой ЛМН-области комплексной плоскости — в диске с центром в начале координат $D_r = \{z \in \mathbb{C}: |z| < r\}$ радиусом $r = 0,9997$. С этой целью неравенства синтеза (27)–(30) были дополнены неравенствами

$$(35) \quad \begin{bmatrix} -rQ & A + B_u K C_y \\ A^T + C_y^T K^T B_u^T & -rP \end{bmatrix} \prec 0,$$

$$(36) \quad P \succ 0, \quad Q \succ 0,$$

а невыпуклое ограничение (31) — ограничением

$$(37) \quad PQ = I_{n_x}.$$

Заданная точность для условий (26) остановки алгоритма принималась равной $\varepsilon = 1 \cdot 10^{-7}$, максимальное число итераций $k_N = 1500$. Начальные условия $\Pi_0 \succ 0$, $\Phi_0 \succ 0$, $P_0 \succ 0$, $Q_0 \succ 0$ алгоритма 1 выбирались из решения системы неравенств (27)–(30), (35), (36) без учета невыпуклых ограничений (31), (37).

Матрица коэффициента усиления анизотропийного субоптимального статического регулятора по выходу K_a представлена ниже вместе с матрицами K_2 - и K_∞ -статических \mathcal{H}_2 - и \mathcal{H}_∞ -регуляторов:

$$K_2 = \begin{bmatrix} -0,3394 & -0,1779 \\ -0,9803 & -0,1322 \end{bmatrix}, \quad K_\infty = \begin{bmatrix} -0,2051 & -0,1793 \\ -0,2716 & -0,01814 \end{bmatrix},$$

$$K_a = \begin{bmatrix} -0,3265 & -0,08319 \\ -1,075 & -0,09644 \end{bmatrix}.$$

Результаты моделирования замкнутых систем с расположением полюсов в заданной ЛМН-области в условиях сдвига ветра и шумов измерений представлены вместе с результатами решения задачи в таблице 1 и проиллюстрированы на рис. 3–8. При моделировании применялся типичный профиль ветра, описываемый моделью в форме вихревого кольца [37].

Таблица 1. Посадка самолета Ту-154: статическая обратная связь по измеряемому выходу. Сравнение замкнутых систем

	Регулятор в цепи обратной связи		
	K_2	K_a	K_∞
Результаты решения:			
γ	20	20	20
$\ T_{zw}\ _2$	1,2241	1,3582	0,81519
$\ T_{zw}\ _{0,1}$	3,3179	5,7172	7,2877
$\ T_{zw}\ _\infty$	16,453	28,652	36,734
Число итераций:	654	782	639
Результаты моделирования:			
max $ \Delta V $, м/с	7,063	3,089	11,69
max $ \Delta h $, м	81,9	93,56	80,51
max $ \Delta \theta $, град	12,4	7,081	11,45
max $ \Delta \omega_z $, град/с	2,767	1,369	2,751
max $ \Delta \vartheta $, град	15,03	8,168	14,77
max $ \Delta T $, кН	11,91	9,267	3,226
max $ \Delta \vartheta_{cy} $, град	15,81	8,405	15,51
max $ \Delta \delta_t $, град	14,47	11,25	3,845

Из результатов решения задачи в таблице 1 и на рис. 8 можно заключить, что

- a -анизотропийная норма замкнутой системы с анизотропийным γ -оптимальным регулятором удовлетворяет условию $\|T_{zw}\|_{0,1} < \gamma$; регулятор является субоптимальным;
- \mathcal{H}_2 - и \mathcal{H}_∞ -нормы замкнутых систем с \mathcal{H}_2 - и \mathcal{H}_∞ -регуляторами удовлетворяют условиям $\|T_{zw}\|_2 < \gamma_2$, $\|T_{zw}\|_\infty < \gamma_\infty$;
- для всех замкнутых систем выполняется требование размещения полюсов в диске \mathcal{D}_r заданного радиуса $r = 0,9997$.

Анализ результатов моделирования, представленных в таблице 1 и на рис. 3–6 показывает, что

- анизотропный субоптимальный регулятор приводит к наименьшей амплитуде отклонения воздушной скорости и наибольшей амплитуде отклонения высоты;
- время установления переходных процессов в системе с анизотропным регулятором меньше, чем в системах с \mathcal{H}_2 - и \mathcal{H}_∞ -регуляторами;
- в замкнутой системе с анизотропным регулятором амплитуды отклонения угла наклона траектории, угловой скорости тангажа и угла тангажа значительно меньше, чем в замкнутых системах с \mathcal{H}_2 - и \mathcal{H}_∞ -регуляторами;
- наименьшая дополнительная тяга двигателя наблюдается в системе с \mathcal{H}_∞ -регулятором;
- амплитуды сигналов управления анизотропного статического регулятора по выходу меньше, чем у \mathcal{H}_2 -регулятора и больше, чем у \mathcal{H}_∞ -регулятора.

3.3. РЕГУЛЯТОРЫ ЗАДАННОГО ПОРЯДКА

Для решения задачи управления самолетом в режиме посадки были также построены анизотропные субоптимальные регуляторы заданного порядка $n_\xi = 1, \dots, 6$, которые вычислялись из решения системы неравенств (27)–(30) при ограничении (31) в соответствии с алгоритмом 1 для расширенного объекта управления вида (3). Для нахождения взаимнообратных матриц $\Pi \succ 0$, $\Phi \succ 0$, $\Pi\Phi = I$, удовлетворяющих ЛМН (28), (29) и выпуклому ограничению (27) применялся алгоритм 1 на основе метода условного градиента.

Анизотропные субоптимальные регуляторы заданных порядков $n_\xi = 1, \dots, 6$, вычислялись для заданного уровня средней

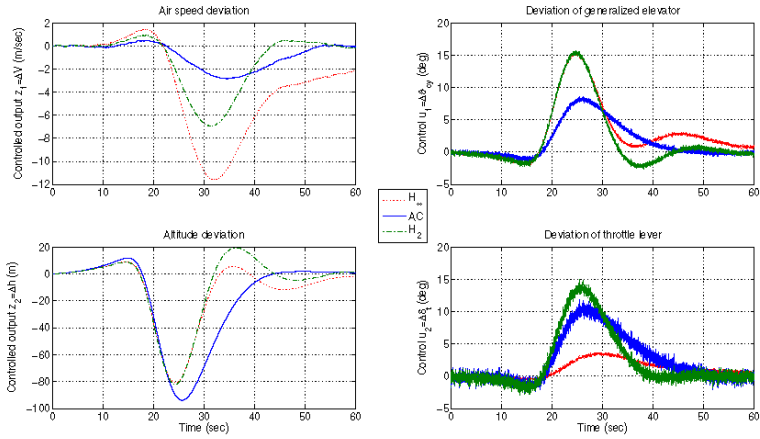


Рис. 3. Посадка самолета Ту-154: статический регулятор по выходу. Воздушная скорость ΔV , высота Δh (диаграммы слева) и сигналы управления $\Delta\vartheta_{cy}$, $\Delta\delta_t$ (диаграммы справа)

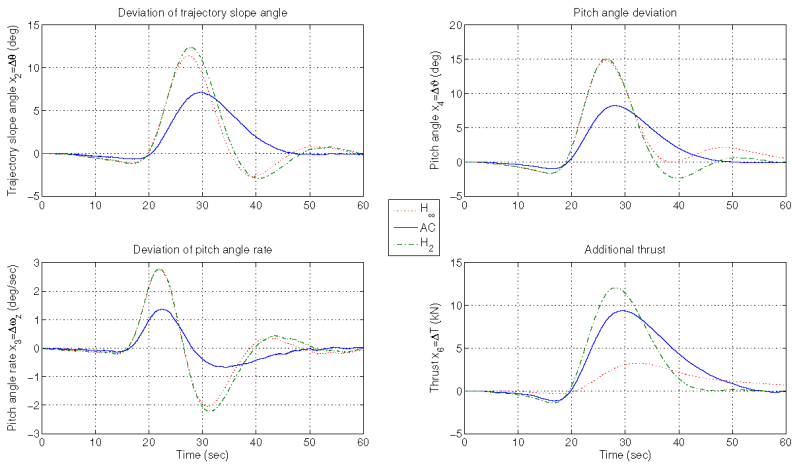


Рис. 4. Посадка самолета Ту-154: статический регулятор по выходу. Угол наклона траектории $\Delta\theta$, угловая скорость тангажа $\Delta\omega_z$ (диаграммы слева), угол тангажа $\Delta\vartheta$, тяга двигателя ΔT (диаграммы справа)

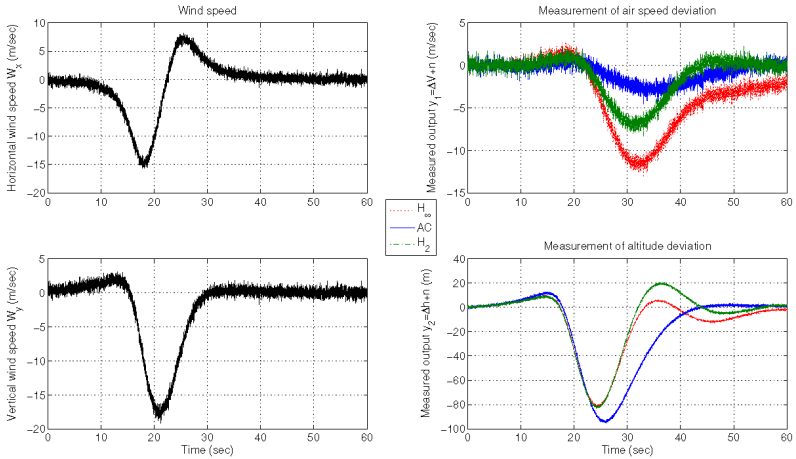


Рис. 5. Посадка самолета Ту-154: статический регулятор по выходу. Профиль ветра (диаграммы слева) и измерения с шумами (диаграммы справа)

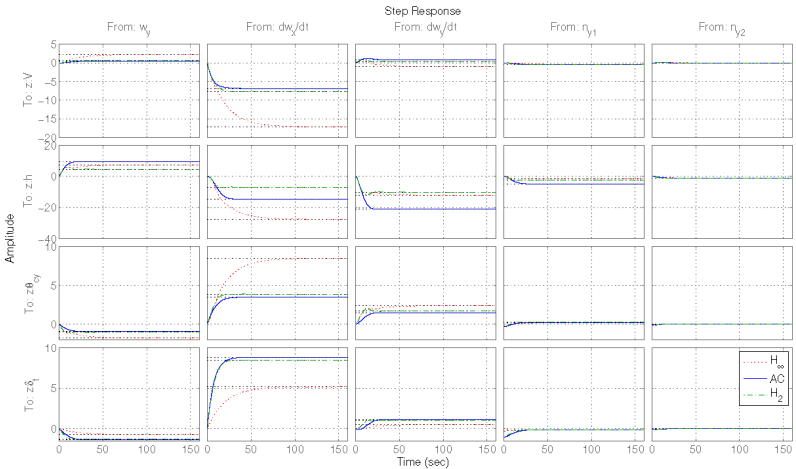


Рис. 6. Посадка самолета Ту-154: статический регулятор по выходу. Переходные характеристики

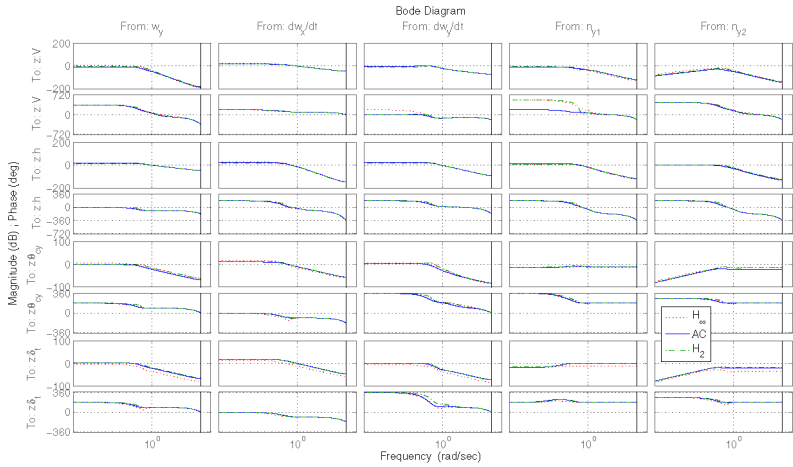


Рис. 7. Посадка самолета Ту-154: статический регулятор по выходу. Диаграмма Бode

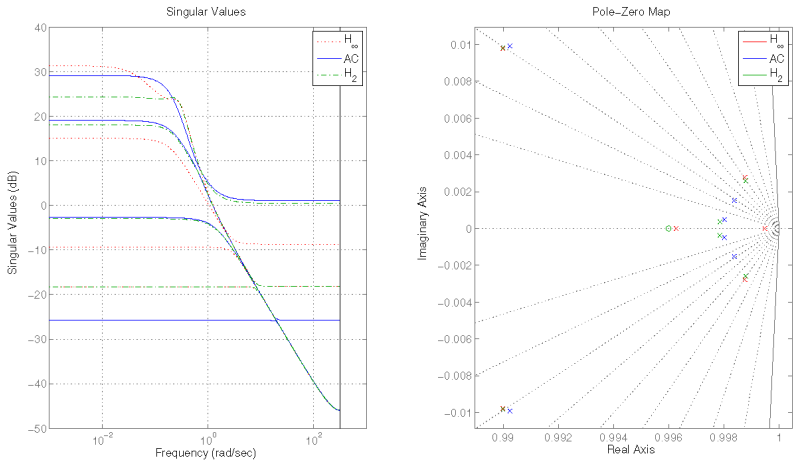


Рис. 8. Посадка самолета Ту-154: статический регулятор по выходу. Сингулярные значения (диаграмма слева), расположение нулей и полюсов (диаграмма справа)

анизотропии внешнего возмущения $a = 0,7$ и фиксированного порогового значения анизотропийной нормы замкнутой системы $\gamma = 20$. При синтезе регуляторов учитывалось дополнительное требование расположения полюсов замкнутой системы в выпуклой ЛМН-области на комплексной плоскости — в диске с центром в начале координат $\mathcal{D}_r = \{z \in \mathbb{C}: |z| < r\}$ радиусом $r = 0,9997$. С этой целью неравенства синтеза (27)–(30) были дополнены неравенствами (35), (36), а невыпуклое ограничение (31) — ограничением (37), и фактически осуществлялся поиск двух пар взаимнообратных матриц, удовлетворяющих условию

$$\begin{bmatrix} \Phi & 0 \\ 0 & \Pi \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & P \end{bmatrix} = I_{2n},$$

$n = n_x + n_\xi$. Заданная точность для условий (26) остановки алгоритма принималась равной $\varepsilon = 1 \cdot 10^{-7}$, максимальное число итераций $k_N = 1500$. Начальные условия $\Pi_0 \succ 0$, $\Phi_0 \succ 0$, $P_0 \succ 0$, $Q_0 \succ 0$ выбирались из решения системы неравенств (27)–(30), (35), (36) без учета невыпуклых ограничений (31), (37). Для регуляторов $K_a^{n_\xi}$ порядков $n_\xi = 5, 3$ заданная точность ε не была достигнута за k_N шагов. Реализации синтезированных регуляторов $K_a^{n_\xi}$ порядков $n_\xi = 6, 4, 2, 1$ приводятся ниже:

$$K_a^6 = \begin{bmatrix} A_c^6 & B_c^6 \\ C_c^6 & D_c^6 \end{bmatrix},$$

$$A_c^6 = \begin{bmatrix} 0,0024 & -0,0002 & -0,0015 & -0,0013 & -0,00084 & 0,00024 \\ -0,00014 & 0,0021 & -0,00029 & 0,0022 & -0,00052 & -0,00042 \\ -0,0014 & -0,0002 & 0,0011 & 0,00047 & 0,00069 & -5,96 \cdot 10^{-5} \\ -0,0015 & 0,0025 & 0,00057 & 0,0035 & -6,2 \cdot 10^{-5} & -0,00055 \\ -0,00083 & -0,00047 & 0,00072 & -8,27 \cdot 10^{-5} & 0,00054 & 3,77 \cdot 10^{-5} \\ 0,00028 & -0,00041 & -8,9 \cdot 10^{-5} & -0,00049 & 1,69 \cdot 10^{-5} & 0,00013 \end{bmatrix},$$

$$C_c^6 = \begin{bmatrix} 1,006 & 0,2312 & -0,7747 & -0,3115 & -0,5271 & 0,01095 \\ -0,8721 & -0,1261 & 0,2615 & 0,2941 & 0,1298 & -0,05475 \end{bmatrix},$$

$$B_c^6 = \begin{bmatrix} -0,002336 & 0,009808 \\ 0,001071 & -0,01202 \\ 0,002105 & -0,005306 \\ -0,003142 & -0,01961 \\ 0,0009146 & -0,00228 \\ -0,001945 & 0,002905 \end{bmatrix}, \quad D_c^6 = \begin{bmatrix} -0,2988 & -0,1826 \\ -0,2011 & -0,1891 \end{bmatrix},$$

$$K_a^4 = \begin{bmatrix} 0,001124 & 0,0001895 & -0,0009076 & -0,0003107 & 0,00445 & -0,001455 \\ 0,000213 & 0,001551 & 0,0008611 & -0,0003292 & 0,004879 & -0,008758 \\ -0,0009223 & 0,0008697 & 0,003372 & -0,0003911 & -0,002122 & -0,01597 \\ -0,000304 & -0,0003404 & -0,0004638 & 0,0002996 & -0,0001242 & 0,00594 \\ -1,009 & -0,1889 & 0,6695 & 0,3422 & -0,286 & -0,187 \\ -0,326 & 0,8818 & 0,1112 & 0,0734 & -0,1897 & -0,198 \end{bmatrix},$$

$$K_a^2 = \begin{bmatrix} 0,0004513 & -0,0003284 & 0,003121 & -0,001795 \\ -0,0003007 & 0,001849 & 0,001835 & -0,008197 \\ 0,536 & -0,4517 & -0,2988 & -0,1636 \\ -0,4541 & 1,084 & -0,02619 & -0,007835 \end{bmatrix},$$

$$K_a^1 = \begin{bmatrix} 0,3851 & 0,01642 & -0,0256 \\ -6,05 & -0,1921 & -0,4216 \\ 7,858 & -0,2948 & 0,1986 \end{bmatrix}.$$

Таблица 2. Посадка самолета Ту-154: регуляторы заданного порядка. Сравнение замкнутых систем

	Порядок анизотропийного субоптимального регулятора			
	$K_a^6, n_\xi = 6$	$K_a^4, n_\xi = 4$	$K_a^2, n_\xi = 2$	$K_a^1, n_\xi = 1$
Результаты решения:				
γ	20	20	20	20
$\ T_{zw}\ _2$	0,79969	0,77843	0,83191	0,93857
$\ T_{zw}\ _{0,7}$	12,458	12,144	19,224	14,885
$\ T_{zw}\ _\infty$	25,197	24,564	49,011	30,109
Число итераций:	1321	1173	1069	972
Результаты моделирования:				
$\max \Delta V $, м/с	4,861	6,349	12,13	8,27
$\max \Delta h $, м	82,89	79,63	82,78	82,15
$\max \Delta \theta $, град	11,93	13,18	10,8	11,88
$\max \Delta \omega_z $, град/с	2,561	2,946	2,565	2,672
$\max \Delta \vartheta $, град	13,73	15,45	14	14,39
$\max \Delta T $, кН	14,66	14,2	1,218	9,205
$\max \Delta \vartheta_{cy} $, град	14,13	16	14,59	15,15
$\max \Delta \delta_t $, град	17,27	16,96	1,423	11,25

Результаты моделирования замкнутых систем в условиях сдвига ветра и шумов измерений представлены вместе с результатами решения задачи в таблице 2 и показаны на диаграммах рис. 9–14. Из результатов решения задачи в таблице 2 и на рис. 14 можно заключить, что

- для всех замкнутых систем с анизотропийными регуляторами $K_a^{n_\xi}$ заданных порядков $n_\xi = 6, 4, 2, 1$, а анизотропийная норма удовлетворяет условию $\|T_{zw}\|_{0,7} < \gamma$; все регуляторы являются субоптимальными;
- для всех замкнутых систем выполняется требование размещения полюсов в диске \mathcal{D}_r заданного радиуса $r = 0,9997$.

Анализ результатов моделирования, представленных в таблице 2 и на рис. 9–14 показывает, что

- наименьшая амплитуда отклонения воздушной скорости наблюдается в замкнутой системе с регулятором 6-го порядка, наибольшая — в системе с регулятором 2-го порядка;
- наименьшая амплитуда отклонения высоты наблюдается в замкнутой системе с регулятором 4-го порядка, наибольшая — в системе с регулятором 6-го порядка;
- в замкнутой системе с регулятором 4-го порядка наблюдается наибольшее отклонение рулей высоты, наименьшее — в системе с регулятором 6-го порядка;
- в замкнутой системе с регулятором 6-го порядка наблюдается наибольшее отклонение сектора газа, наименьшее — в системе с регулятором 2-го порядка.

В сравнении с субоптимальными регуляторами заданного порядка (таблица 2), качество работы замкнутых систем со статическими анизотропийными регуляторами по выходу (таблица 1) хуже. В то же время статические регуляторы характеризуются меньшими амплитудами управляющих сигналов и амплитудами отклонений угла наклона траектории, угловой скорости тангажа, угла тангажа и тяги двигателя.

Алгоритм 1 на основе метода условного градиента в данной задаче показал низкую скорость сходимости (см. таблицы 1, 2,

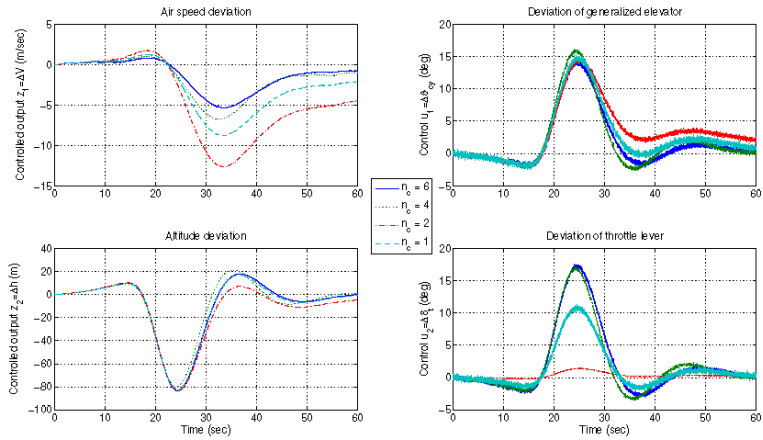


Рис. 9. Посадка самолета Ту-154: регуляторы заданного порядка. Воздушная скорость ΔV , высота Δh (диаграммы слева) и сигналы управления $\Delta \delta_{cy}$, $\Delta \delta_t$ (диаграммы справа)

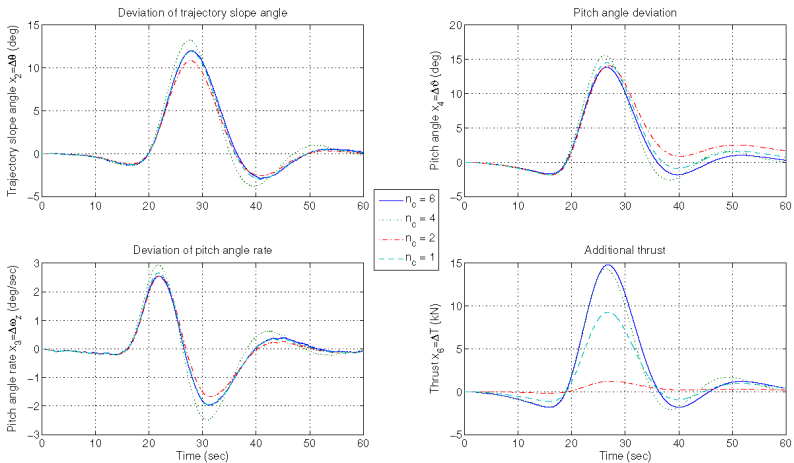


Рис. 10. Посадка самолета Ту-154: регуляторы заданного порядка. Угол наклона траектории $\Delta \theta$, угловая скорость тангажа $\Delta \omega_z$ (диаграммы слева), угол тангажа $\Delta \vartheta$, тяга двигателя ΔT (диаграммы справа)

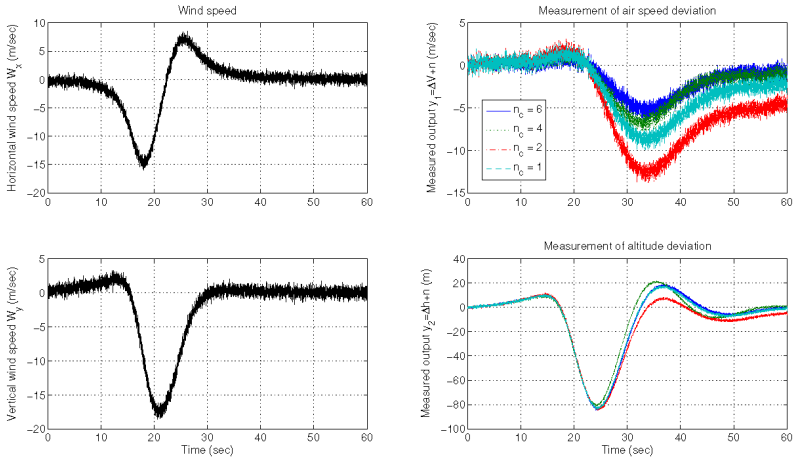


Рис. 11. Посадка самолета Ту-154: регуляторы заданного порядка. Профиль ветра (диаграммы слева) и измерения с шумами (диаграммы справа)

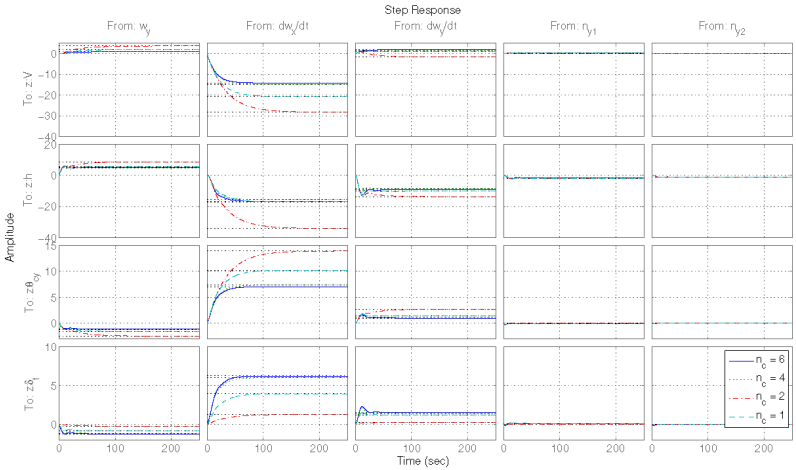


Рис. 12. Посадка самолета Ту-154: регуляторы заданного порядка. Переходные характеристики

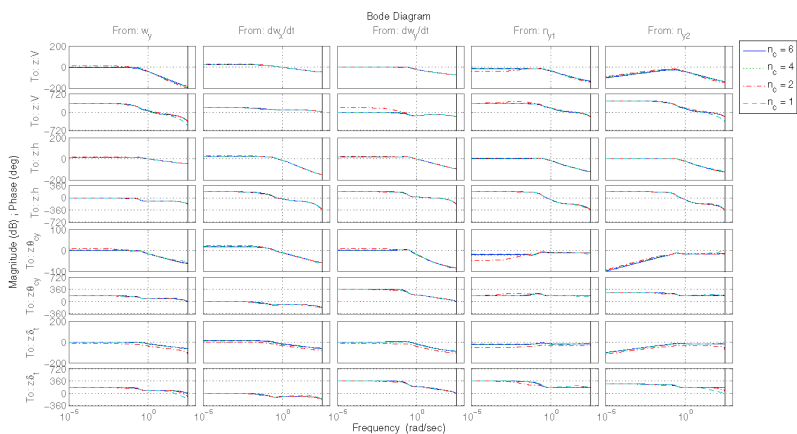


Рис. 13. Посадка самолета Ту-154: регуляторы заданного порядка. Диаграмма Бode

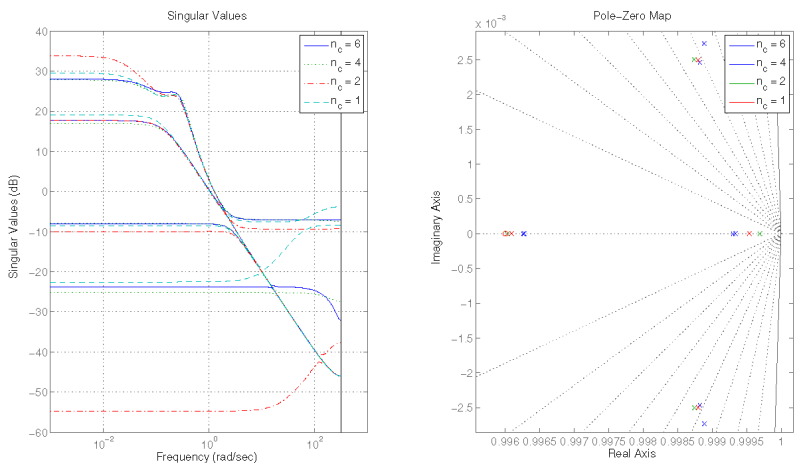


Рис. 14. Посадка самолета Ту-154: регуляторы заданного порядка. Сингулярные значения (диаграмма слева), расположение нулей и полюсов (диаграмма справа)

число итераций). Для алгоритма поиска взаимнообратных матриц, разработанного в [1], не удалось подобрать начальных условий, которые приводили бы к его сходимости к глобальному минимуму и к нахождению решений. Для генерации точек начальных условий алгоритма 1, равномерно распределенных на множестве решений системы неравенств (27)–(30), (35), (36), можно применить алгоритм “Hit&Run” [55].

4. Заключение

В работе предлагается метод решения задач синтеза анизотропийных субоптимальных регуляторов, основанный на полуопределенном программировании и выпуклой оптимизации. Анизотропийный субоптимальный регулятор стабилизирует замкнутую систему и гарантирует, что ее анизотропийная норма не превосходит заданного порогового значения. Общая процедура синтеза регулятора заданного порядка сводится к решению неравенства относительно детерминанта положительно определенной матрицы и двух ЛМН относительно взаимнообратных матриц; задача оптимизации не является выпуклой. Для решения этой задачи применяется локально-оптимальный алгоритм поиска взаимнообратных матриц на основе метода условного градиента. В сравнении с решением задачи синтеза анизотропийного оптимального регулятора [67], предлагаемый подход на основе полуопределенного программирования является новым и позволяет учитывать при проектировании ряд дополнительных требований к замкнутой системе, например, заданную структуру регулятора.

Литература

1. БАЛАНДИН Д.В., КОГАН М.М. *Синтез регуляторов на основе решения линейных матричных неравенств и алгоритма поиска взаимнообратных матриц* // АиТ. – 2005, №1. – С. 82-99.
2. БОДНЕР В.А., КОЗЛОВ М.С. *Стабилизация летательных аппаратов и автопилоты*. М: Оборонгиз, 1961.

3. БУКОВ В.Н. *Адаптивные прогнозирующие системы управления полетом*. М.: Наука, 1987.
4. ВЛАДИМИРОВ И.Г., ДАЙМОНД П., КЛОЕДЕН П. *Анизотропный анализ робастного качества линейных нестационарных дискретных систем на конечном временном интервале* // *АиТ*. – 2006, №8. – С. 92-111.
5. ВЛАДИМИРОВ И.Г., КУРДЮКОВ А.П., СЕМЕНОВ А.В. *Анизотропия сигналов и энтропия линейных стационарных систем* // *ДАН*. – 1995. №3. – С. 583-585.
6. ВЛАДИМИРОВ И.Г., КУРДЮКОВ А.П., СЕМЕНОВ А.В. *Стохастическая проблема \mathcal{H}_∞ -оптимизации* // *ДАН*. – 1995. – Т. 343, №5. – С. 607-609.
7. КУРДЮКОВ А.П., МАКСИМОВ Е.А. *Решение задачи стохастической \mathcal{H}_∞ -оптимизации для линейной системы с неопределенностью* // *АиТ*. – 2006, №8. – С. 112-142.
8. ПОЛЯК Б.Т. *Введение в оптимизацию*. М.: Наука, 1983.
9. *Разработка основ теории нетрадиционных подходов и исследование алгоритмов управления полетом в сложных условиях. Отчет о научно-исследовательской работе по теме №074-95/01*. М: Институт проблем управления РАН, 1995.
10. *Разработка принципов автоматизации полета и исследования новых алгоритмов управления на этапах захода на посадку и приземления. Отчет о научно-исследовательской работе по теме №053-93/01*. М: Институт проблем управления РАН, 1993.
11. ЧАЙКОВСКИЙ М.М., КУРДЮКОВ А.П. *Критерий строгой ограниченности анизотропной нормы заданным значением в терминах матричных неравенств* // *ДАН*. – 2011, Т.441, №3. – С. 318-321.
12. APKARIAN, P., NOLL, D., TUAN, H.D. *Fixed-order \mathcal{H}_∞ control design via a partially augmented Lagrangian method* // *Int. J. of Nonlinear and Robust Contr.* – 2003. – Vol. 13. – P. 1137-1148.
13. APKARIAN, P., PELLANDA, P.C., TUAN, H.D. *Mixed*

- $\mathcal{H}_2/\mathcal{H}_\infty$ multi-channel linear parameter-varying control in discrete time // Syst. & Contr. Lett. – 2000. – Vol. 41. – P. 333-346.
14. APKARIAN, P., TUAN, H.D. *Concave programming in control theory* // J. of Glob. Opt. – 1999. – Vol. 15. – P. 343-370.
 15. ARZELIER, D., PEAUCELLE, D. *An iterative method for mixed $\mathcal{H}_2/\mathcal{H}_\infty$ synthesis via static output feedback* // Proc. IEEE Conf. Dec. Contr. – 2002. – P. 3464-3469.
 16. BEN-TAL, A., NEMIROVSKII, A. *Lectures on Modern Convex Optimization*. Technion, Haifa, Israel, 2000.
 17. BERNSTEIN, D.S., HADDAD, W.M. *LQG control with an \mathcal{H}_∞ performance bound: a Riccati equation approach* // IEEE Trans. AC. – 1989. – Vol. 34. – P. 293-305.
 18. BOYD, S.P., GHAOUI, L.EL. *Method of centers for minimizing generalized eigenvalues* // Lin. Alg. Appl. – 1993. – Vol. 188. – P. 63-111.
 19. BOYD, S., GHAOUI, L.EL., FERON, E., BALAKRISHNAN, V. *Linear Matrix Inequalities in Systems and Control theory*. SIAM, Philadelphia, PA, 1994.
 20. CHEN, X., WEN, J.T. *A linear matrix inequality approach to the general mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control problem* // Proc. American Control Conf. – 1995. – P. 1443-1447.
 21. DIAMOND, P., KURDJUKOV, A.P., SEMYONOV, A.V., VLADIMIROV, I.G. *Homotopy methods and anisotropy-based stochastic \mathcal{H}_∞ optimization of control systems* // Report 97-14 of The University of Queensland, Australia. – 1997. – P. 1-22.
 22. DIAMOND, P., VLADIMIROV, I.G., KURDYUKOV, A.P., SEMYONOV, A.V. *Anisotropy-based performance analysis of linear discrete time invariant control systems* // Int. J. of Contr. – 2001. – Vol. 74. – P. 28-42.
 23. DOYLE, J.C. *Guaranteed margins for LQG regulators* // IEEE Trans. AC. – 1978. – Vol. 23. – P. 756-757.
 24. DOYLE, J.C., GLOVER, K., KHARGONEKAR, P.P.,

- FRANCIS, B.A. *State-space solutions to standard \mathcal{H}_2 and \mathcal{H}_∞ control problems* // IEEE Trans. AC. – 1989. – Vol. 34. – P. 831-848.
25. FARES, B., APKARIAN, P., NOLL, D. *An augmented Lagrangian method for a class of LMI-constrained problems in robust control theory* // Int. J. of Contr. – 2001. – Vol. 74. – P. 348-360.
26. FRIDMAN, E., SHAKED, U. *Robust \mathcal{H}_∞ minimum entropy static output-feedback control of singularly perturbed systems* // Automatica. – 2000. – Vol. 36. – P. 1181-1188.
27. GAHINET, P., APKARIAN, P. *A linear matrix inequality approach to \mathcal{H}_∞ control* // Int. J. of Robust and Nonlinear Contr. – 1994. – Vol. 4. – P. 421-448.
28. GHAOUI, L.EL., OUSTRY, F., RAMI, M.A. *A cone complementarity linearization algorithm for static output-feedback and related problems* // IEEE Trans. AC. – 1997. – Vol. 42. – P. 1171-1176.
29. GLOVER, K., DOYLE, J.C. *State-space formulae for all stabilizing controllers that satisfy an \mathcal{H}_∞ -norm bound and relations to risk sensitivity* // Syst. & Contr. Lett. – 1988. – Vol. 11. – P. 167-172.
30. GUSEV, S.V. *Minimax control under a bound on the partial covariance sequence of the disturbance* // Automatica. – 1995. – Vol. 31. – P. 1287-1301.
31. GUSEV, S.V. *Minimax control under a restriction on the moments of disturbance* // Proc. 34th IEEE Conf. on Decision and Control, New Orleans, USA. – 1995. – P. 1195-1200.
32. GUSEV, S.V. *Method of moment restrictions in robust control and filtering* // Proc. 13th IFAC World Congress, San Francisco, USA. – 1996. – P. 415-420.
33. HINDI, H.A., HASSIBI, B., BOYD, S.P. *Multiobjective $\mathcal{H}_2/\mathcal{H}_\infty$ -optimal control via finite dimensional Q-parametrization and linear matrix inequalities* // Proc. American Control Conf. – 1998. – P. 3244-3248.
34. IGLESIAS, P.A., MUSTAFA, D. *State-space solution of*

- the discrete-time minimum entropy control problem via separation* // IEEE Trans. AC. – 1993. – Vol. 38. – P. 1525-1530.
35. IWASAKI, T., SKELTON, R.E. *All controllers for the general \mathcal{H}_∞ control problem: LMI existence conditions and state space formulas* // Automatica. – 1994. – Vol. 30. – P. 1307-1317.
36. IGLESIAS, P.A., MUSTAFA, D., GLOVER, K. *Discrete time \mathcal{H}_∞ controllers satisfying a minimum entropy criterion* // Syst. & Contr. Lett. – Vol. 14. – P. 275-286.
37. IVAN, M. *A ring vortex downburst model for flight simulation* // J. Aircraft. – 1996. – Vol. 23. – P. 232-236.
38. IWASAKI, T., SKELTON, R.E. *The XY-centering algorithm for the dual LMI Problem: A new approach to fixed order design* // Int. J. of Contr. – 1995. – Vol. 62. – P. 1257-1272.
39. JACOBSON, D.H. *Extensions of Linear-Quadratic Control, Optimization and Matrix Theory*. Academic Press, NY, 1977.
40. KHARGONEKAR, P.P., ROTEA, M.A. *Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control: a convex optimization approach* // IEEE Trans. AC. – 1991. – Vol. 36. P. 824-837.
41. KIMURA, H. *Pole assignment by gain output feedback* // IEEE Trans. AC. – 1975. – Vol. AC-20. – P. 509-516.
42. KURDYUKOV, A.P., MAXIMOV, E.A. *State-space solution to stochastic \mathcal{H}_∞ -optimization problem with uncertainty* // Proc. 16th IFAC World Congr., Prague, Czechia. – 2005.
43. KURDYUKOV, A.P., MAXIMOV, E.A., TCHAIKOVSKY, M.M. *Anisotropy-based bounded real lemma* // Proc. 19th Int. Symp. on Mathematical Theory of Networks and Systems, Budapest, Hungary. – 2010. – P. 2391-2397.
44. KURDYUKOV, A.P., PAVLOV, B.V., TIMIN, V.N., VLADIMIROV, I.G. *Longitudinal anisotropy-based flight control in a wind shear* // Proc. 16th IFAC Symposium on Automatic Control in Aerospace, Saint-Petersburg, Russia. – 2004.
45. L"OFBERG, J. *YALMIP: A toolbox for modeling*

- and optimization in MATLAB* // Proc. of the CACSD Conference, Taipei, Taiwan. – 2004. Available from <http://users.isy.liu.se/johanl/yalmip/>.
46. MASUBUCHI, I., OHARA, A., SUDA, N. *LMI-based controller synthesis: A unified formulation and solution* // Int. J. of Robust and Nonlinear Contr. – 1998. – Vol. 8. – P. 669-686.
 47. MAXIMOV, E.A., KURDYUKOV, A.P., VLADIMIROV, I.G. *Anisotropic norm bounded real lemma for linear discrete time varying systems* // Proc. 18th IFAC World Congr., Milano, Italy. – 2011.
 48. MIRADORE, R., RICCI, G. *Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control: the discrete-time case* // Syst. & Contr. Lett. – 2005, – Vol. 54. – P. 1-13.
 49. MUSTAFA, D., GLOVER, K. *Minimum Entropy \mathcal{H}_∞ Control*. Springer-Verlag, NY, 1991.
 50. MUSTAFA, D., GLOVER, K., LIMEBEER, D. *Solutions to the \mathcal{H}_∞ general distance problem which minimize an entropy integral* // Automatica. – 1991. – Vol. 27. – P. 193-199.
 51. NEMIROVSKII, A., GAHINET, P. *The projective method for solving linear matrix inequalities* // Math. Programming Series B. – 1997. – Vol. 77. – P. 163-190.
 52. NESTEROV, YU., NEMIROVSKII, A. *Interior-point Polynomial Methods in Convex Programming*, volume 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.
 53. NOLL, D., TORKI, M., APKARIAN, P. *Partially augmented Lagrangian method for matrix inequality constraints* // SIAM J. on Opt. – 2004. – Vol. 15. – P. 161-184.
 54. OLIVEIRA, M.C., GEROMEL, J.C., BERBNUSSOU, J. *An LMI optimization approach to multiobjective controller design for discrete-time systems* // Proc. IEEE Conf. on Decision and Control. – 1999. – P. 27-38.
 55. POLYAK, B.T., GRYAZINA, E.N. *Hit-and-Run: Randomized technique for control problems recasted as concave*

- programming* // Proc. 18th IFAC World Congr., Milano, Italy. – 2011.
56. POLYAK, B.T., GRYAZINA, E.N. *Markov chain Monte Carlo method exploiting barrier functions with applications to control and optimization* // Proc. IEEE Multi-Conf. on Systems and Control. – 2010. – P. 1553-1557.
 57. ROTSTEIN, H., SZNAIER, M. *An exact solution to general four-block discrete-time mixed $\mathcal{H}_2/\mathcal{H}_\infty$ problems via convex optimization* // IEEE Trans. AC. – 1998. – Vol. 43. – P. 1475-1481.
 58. SCHERER, C.W. *Multiobjective $\mathcal{H}_2/\mathcal{H}_\infty$ control* // IEEE Trans. AC. – 1995. – Vol. 40. – P. 1054-1062.
 59. SCHERER, C.W. *Robust controller design by output feedback against uncertain stochastic disturbances* // Proc. 3rd IFAC Symp. on Robust Control Design, Prague, Czechia. – 2000.
 60. SCHERER, C.W. *Multi-objective control without Youla parametrization* // In Perspectives in Robust Control. Lecture Notes on Control and Information Sciences. – 2001. – Vol. 268. – P. 311-325.
 61. SEMYONOV, A.V., VLADIMIROV, I.G., KURDJUKOV, A.P. *Stochastic approach to \mathcal{H}_∞ -optimization* // Proc. 33rd IEEE Conf. on Decision and Control, Florida, USA. – 1994. – P. 2249-2250.
 62. SCHERER, C.W., GAHINET, P., CHILALI, M. *Multiobjective output-feedback control via LMI optimization* // IEEE Trans. AC. – 1997. – Vol. 42. – P. 896-911.
 63. STURM, J.F. *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones* // Optimization Methods and Software. – 1999. – Vol. 11. – P. 625-653.
 64. TCHAIKOVSKY, M.M., KURDYUKOV, A.P., TIMIN, V.N. *Strict anisotropic norm bounded real lemma in terms of inequalities* // Proc. 18th IFAC World Congr., Milano, Italy. – 2011. – P. 2332-2337.
 65. TCHAIKOVSKY, M.M., KURDYUKOV, A.P., TIMIN, V.N. *A convex formulation of strict anisotropic norm*

- bounded real lemma* // Preprint. Available from <http://arxiv.org/abs/1108.5140>. – 2011.
66. VLADIMIROV, I.G., KURDJUKOV, A.P., SEMYONOV, A.V. *On computing the anisotropic norm of linear discrete-time-invariant systems* // Proc. 13th IFAC World Congress, San-Francisco, USA. – 1996. – P. 179-184.
67. VLADIMIROV, I.G., KURDJUKOV, A.P., SEMYONOV, A.V. *State-space solution to anisotropy-based stochastic \mathcal{H}_∞ -optimization problem* // Proc. 13th IFAC World Congress, San Francisco, USA. – 1996. – P. 427-432.
68. WHITTLE, P. *Risk-sensitive linear/quadratic/Gaussian control* // Adv. Appl. Prob. – 1981. Vol. 13. – P. 764-777.
69. WHITTLE, P. *Entropy-minimizing and risk-sensitive control rules* // Syst. & Contr. Lett. – 1989. – Vol. 13. – P. 1-7.
70. YAESH, I., SHAKED, U. *Minimum entropy static output-feedback control with an \mathcal{H}_∞ -norm performance bound* // IEEE Trans. AC. – 1997. – Vol. 42. – P. 853-858.
71. ZHOU, K., GLOVER, K., BODENHEIMER, B.A., DOYLE, J.C. *Mixed \mathcal{H}_2 and \mathcal{H}_∞ performance objectives I: Robust performance analysis, II: Optimal control* // IEEE Trans. AC. – 1994. – Vol. 39. – P. 1564-1574, 1575-1587.

**SYNTHESIS OF ANISOTROPIC SUBOPTIMAL
FIXED-ORDER CONTROLLERS VIA SEMIDEFINITE
PROGRAMMING AND ALGORITHM FOR
SEARCHING RECIPROCAL MATRICES**

Michael Tchaikovsky, Institute of Control Sciences of RAS,
Moscow, Candidate of Science (mmtchaikovsky@hotmail.com).

Abstract: A disturbance attenuation problem is considered for a linear discrete time invariant system under random disturbances with imprecisely known distributions. The designed anisotropic suboptimal controller is a dynamic fixed-order output-feedback compensator, which is required to stabilize the closed-loop system and to keep its anisotropic norm below a prescribed threshold value. The proposed optimization-based approach to synthesis of the anisotropic controllers is novel.

Keywords: discrete linear time invariant systems, random disturbances, statistical uncertainty, norm, anisotropy, semidefinite programming, linear matrix inequalities, reciprocal matrices.

*Статья представлена к публикации
членом редакционной коллегии А. П. Курдюковым*

УДК 681.51
ББК 3.30.10

ИССЛЕДОВАНИЕ СИСТЕМЫ УПРАВЛЕНИЯ ДВИГАТЕЛЕМ ПОСТОЯННОГО ТОКА С МНОГОУРОВНЕВЫМ ПРЕОБРАЗОВАТЕЛЕМ НАПРЯЖЕНИЯ

Гордеев А. А.¹, Юркевич В. Д.², Зиновьев Г. С.³
(Новосибирский государственный
технический университет, Новосибирск)

Обсуждается задача синтеза системы управления для двигателя постоянного тока независимого возбуждения с многоуровневым преобразователем напряжения. Рассматривается двухконтурная система подчиненного регулирования, которая содержит контур стабилизации тока цепи якоря и контур стабилизации скорости вращения двигателя. В качестве закона управления для каждого контура используется ПИ-регулятор, где расчет параметров регулятора выполнен на основе метода разделения движения.

Ключевые слова: двигатель постоянного тока, DC–DC конвертор, широтно-импульсный модулятор, ПИ-регулятор, метод разделения движений.

1. Введение

Повышение мощности электровозов ведет к увеличению потерь в сетях питающего напряжения. Современные электровозы

¹ Артем Александрович Гордеев, магистр техники и технологий (artem.a.gordeev@gmail.com).

² Валерий Дмитриевич Юркевич, доктор технических наук, профессор (тел. (383) 346-49-35, yurkev@ac.cs.nstu.ru).

³ Геннадий Степанович Зиновьев, доктор технических наук, профессор (тел. (383) 346-11-82, genstep@mail.ru).

постоянного тока рассчитаны на работу в сетях с напряжением 1,5 кВ и 3 кВ. Развитие силовых полупроводниковых устройств позволяет создавать экономичные и эффективные понижающие преобразователи напряжения, например, на основе применения многоуровневых преобразователей напряжения для контактных сетей с напряжением в 12–18 кВ и более [3]. Переход к контактным сетям с повышенным напряжением требует разработки методов синтеза систем управления для двигателей электровоза с многоуровневым преобразователем напряжения.

2. Постановка задачи

Обсуждается проблема синтеза системы управления для многоуровневого DC–DC конвертора контактного напряжения 12 кВ в напряжение электрооборудования электровоза 3 кВ. Данный конвертор рассматривался в работе [3], его принципиальная схема представлена на рис. 1. В работе рассматривается система подчиненного регулирования, двухконтурная структура которой содержит контур стабилизации тока цепи якоря и контур стабилизации скорости вращения двигателя (рис. 2). В качестве привода колесной пары рассматриваются два последовательно соединенных в электрической цепи тяговых двигателя постоянного тока, например, типа НБ-511, где $U_{я,ном} = 1475$ В; $I_{я,ном} = 510$ А; $\omega_{ном} = 70$ рад/с.

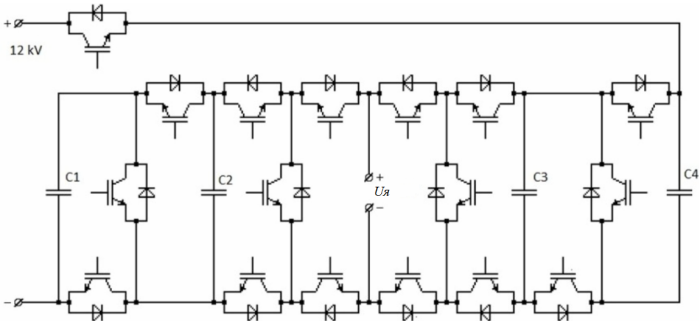


Рис. 1. Схема многоуровневого DC–DC конвертора

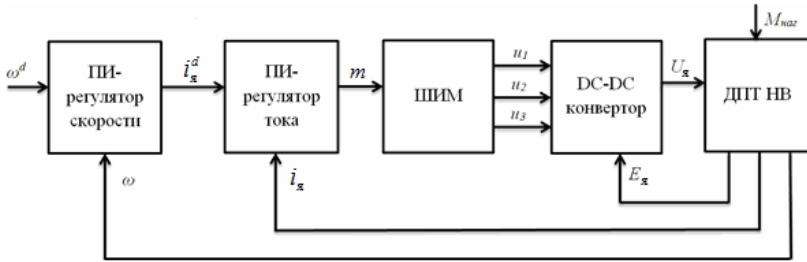


Рис. 2. Структура системы управления

В контуре стабилизации тока в цепи якоря электродвигателя необходимо обеспечить выполнение условия

$$(1) \lim_{t \rightarrow \infty} I_{\text{я}}(t) = i_{\text{я}}^d$$

и заданное время переходных процессов, где $I_{\text{я}}$ – средняя величина тока в цепи якоря двигателя; $i_{\text{я}}^d$ – желаемое значение тока. В контуре стабилизации скорости вращения двигателя требуется обеспечить условие

$$(2) \lim_{t \rightarrow \infty} \omega(t) = \omega^d$$

с заданной величиной времени переходных процессов, где ω – скорость вращения двигателя; ω^d – желаемая скорость вращения двигателя.

3. Математическая модель DC–DC конвертора

Принцип работы конвертора основан на периодическом переключении конденсаторов с последовательного соединения на параллельное. На первом этапе все конденсаторы подключены к источнику питания с напряжением E_1 и внутренним сопротивлением R_{in} , где $E_1 = 12$ кВ. На втором этапе конденсаторы C_1 и C_2 разряжаются на нагрузку, представленную как последовательное соединение активного сопротивления $R_{\text{я}}$ в цепи якоря, индуктивности $L_{\text{я}}$ обмотки якоря и противо-ЭДС $E_{\text{я}}$ в цепи якоря двигателя постоянного тока с независимым возбуждением, где $U_{\text{я}} = L_{\text{я}} \dot{i}_{\text{я}}^{(1)} + R_{\text{я}} i_{\text{я}} + E_{\text{я}}$ и $E_{\text{я}} = k_1 \omega$. На третьем этапе на нагрузку разряжаются конденсаторы C_3 и C_4 .

Процессы, протекающие в преобразователе при условии идеальных ключей, могут быть описаны следующей системой дифференциальных уравнений [9]:

$$\begin{aligned}
 \frac{di_{\text{я}}}{dt} &= -\frac{R_{\text{я}}}{L_{\text{я}}}i_{\text{я}} - \frac{E_{\text{я}}}{L_{\text{я}}} + \frac{u_{C1}}{L_{\text{я}}}u_2 + \frac{u_{C3}}{L_{\text{я}}}u_3, \\
 \frac{du_{C1}}{dt} &= \frac{1}{R_{\text{in}}C_1}(E_1 - u_{C1} - u_{C2} - u_{C3} - u_{C4})u_1 - \frac{i_{\text{я}}}{C_1 + C_2}u_2, \\
 \frac{du_{C2}}{dt} &= \frac{1}{R_{\text{in}}C_2}(E_1 - u_{C1} - u_{C2} - u_{C3} - u_{C4})u_1 - \frac{i_{\text{я}}}{C_1 + C_2}u_2, \\
 \frac{du_{C3}}{dt} &= \frac{1}{R_{\text{in}}C_3}(E_1 - u_{C1} - u_{C2} - u_{C3} - u_{C4})u_1 - \frac{i_{\text{я}}}{C_3 + C_4}u_3, \\
 \frac{du_{C4}}{dt} &= \frac{1}{R_{\text{in}}C_4}(E_1 - u_{C1} - u_{C2} - u_{C3} - u_{C4})u_1 - \frac{i_{\text{я}}}{C_3 + C_4}u_3,
 \end{aligned}
 \tag{3}$$

где $i_{\text{я}}$ – мгновенное значение тока в цепи якоря; u_{Ci} – мгновенное значение напряжения на i -м конденсаторе; u_1, u_2, u_3 – функции переключения, с помощью которых осуществляется переключение между этапами работы преобразователя. Для этапа 1 $u_1 = 1, u_2 = 0, u_3 = 0$; для этапа 2 $u_1 = 0, u_2 = 1, u_3 = 0$; для этапа 3 $u_1 = 0, u_2 = 0, u_3 = 1$.

Вид функций переключения u_1, u_2 и u_3 на выходе ШИМ показан на рис. 3.

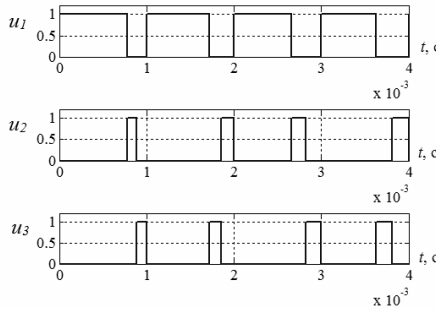


Рис. 3. Управляющие сигналы на выходе ШИМ

С целью устранения различий в условиях заряда и разряда конденсаторов использована следующая последовательность переключений между этапами: этап 1, этап 2, этап 3, этап 1, этап 3, этап 2 и т.д.

Управление преобразователем осуществляется с помощью широтно-импульсного модулятора (ШИМ), где входным сигналом модулятора является переменная m , которая определяет длительность этапа заряда конденсаторов, $m \in (0, 1)$. Длительности этапов разряда первой и второй пары конденсаторов предполагаются одинаковыми. Правило формирования функций переключения u_1 , u_2 и u_3 на выходе ШИМ, показанных на рис. 3, задано следующими условиями:

$$(4) \quad u_1 = \begin{cases} 1 & \text{for } t_k < t \leq t_k + m(t_k)T_s \\ 0 & \text{for } t_k + m(t_k)T_s < t \leq t_k + T_s; \end{cases}$$

$$(5) \quad u_2 = \begin{cases} 0 & \text{for } t_k < t \leq t_k + m(t_k)T_s, \\ 1 & \text{for } t_k + m(t_k)T_s < t \leq t_k + m(t_k)T_s + [1 - m(t_k)]T_s / 2, \\ 0 & \text{for } t_k + m(t_k)T_s + [1 - m(t_k)]T_s / 2 < t \leq t_k + T_s; \end{cases}$$

$$(6) \quad u_3 = \begin{cases} 0 & \text{for } t_k < t \leq t_k + m(t_k)T_s, \\ 0 & \text{for } t_k + m(t_k)T_s < t \leq t_k + m(t_k)T_s + [1 - m(t_k)]T_s / 2, \\ 1 & \text{for } t_k + m(t_k)T_s + [1 - m(t_k)]T_s / 2 < t \leq t_k + T_s; \end{cases}$$

где T_s – период дискретизации ШИМ; $m(t_k)$ – величина коэффициента заполнения импульса при $t = t_k$, $t_k = kT_s$, $k = 0, 1, 2, \dots$

Предполагая, что период дискретизации T_s является малой величиной и отсутствует насыщение в широтно-импульсном модуляторе, т.е. $m \in (0, 1)$, рассмотрим усредненную модель для процессов в преобразователе [8, 9]:

$$\frac{dI_{\text{я}}}{dt} = -\frac{R_{\text{я}}}{L_{\text{я}}} I_{\text{я}} - \frac{E_{\text{я}}}{L_{\text{я}}} + \frac{U_{C1} + U_{C3}}{2L_{\text{я}}} - \frac{U_{C1} + U_{C3}}{2L_{\text{я}}} m,$$

$$\frac{dU_{C1}}{dt} = -\frac{I_{\text{я}}}{2(C_1 + C_2)} + \left[\frac{E_1 - \sum U_{C_i}}{R_m C_1} + \frac{I_{\text{я}}}{2(C_1 + C_2)} \right] m,$$

$$\frac{dU_{C2}}{dt} = -\frac{I_{я}}{2(C_1 + C_2)} + \left[\frac{E_1 - \Sigma U_{Ci}}{R_{in} C_2} + \frac{I_{я}}{2(C_1 + C_2)} \right] m,$$

$$\frac{dU_{C3}}{dt} = -\frac{I_{я}}{2(C_3 + C_4)} + \left[\frac{E_1 - \Sigma U_{Ci}}{R_{in} C_3} + \frac{I_{я}}{2(C_3 + C_4)} \right] m,$$

$$\frac{dU_{C4}}{dt} = -\frac{I_{я}}{2(C_3 + C_4)} + \left[\frac{E_1 - \Sigma U_{Ci}}{R_{in} C_4} + \frac{I_{я}}{2(C_3 + C_4)} \right] m,$$

где $I_{я}$ – среднее значение тока в цепи якоря; U_{Ci} – среднее значение напряжения на i -м конденсаторе; $\Sigma U_{Ci} = U_{C1} + U_{C2} + U_{C3} + U_{C4}$. Принимая $C_1 = C_2 = C_3 = C_4 = C$ и при условии, что выполняется равенство $U_C = U_{C1} = U_{C2} = U_{C3} = U_{C4}$, где $U_C \approx E_1/4$, а также учитывая, что процессы по напряжению протекают значительно быстрее процессов по току, получена следующая упрощенная модель для средней величины тока якоря:

$$(7) \quad \frac{dI_{я}}{dt} = -\frac{R_{я}}{L_{я}} I_{я} - \frac{E_{я}}{L_{я}} + \frac{E_1}{4L_{я}} - \frac{E_1}{4L_{я}} m.$$

4. Синтез регулятора тока

Регулятор контура стабилизации тока якоря должен обеспечивать условие (1). Рассмотрим алгоритм управления в виде следующего дифференциального уравнения:

$$(8) \quad \mu_{я}^2 m^{(2)} + d_{я} \mu_{я} m^{(1)} = k_{я} [(i_{я}^d - i_{я}) / T_{я} - i_{я}^{(1)}],$$

где $\mu_{я}$ – малый положительный параметр; $i_{я}$ – мгновенное значение измеряемого тока в обмотке якоря, $\mu_{я} > 0$; $d_{я} > 0$ и $T_{я} > 0$. Выполняя преобразование Лапласа для уравнения (8) при нулевых начальных условиях, получим

$$m(s) = \frac{k_{я}}{\mu_{я}(\mu_{я}s + d_{я})} \left\{ \frac{1}{sT_{я}} [i_{я}^d(s) - i_{я}(s)] - i_{я}(s) \right\}.$$

Из данного выражения следует, что алгоритм управления (8) является пропорционально-интегральным (ПИ) регулятором с дополнительной фильтрацией и, соответственно, может быть реализован без применения операции дифференцирования. При практической реализации алгоритм управления (8) можно пред-

ставить в виде системы двух дифференциальных уравнений в форме Коши [7].

Заметим, что свойство фильтрации высокочастотной колебательной составляющей в процессах по току $i_{\text{я}}(t)$ является важным для уменьшения амплитуды высокочастотных колебаний в управляющем сигнале $m(t)$ на входе ШИМ.

Анализ свойств замкнутой системы выполним на основе рассмотрения усредненной модели (7) с алгоритмом управления (8), где выполним замену $i_{\text{я}} = I_{\text{я}}$. Полагаем, что $\mu_{\text{я}} \rightarrow 0$. Наличие малого параметра приводит к возникновению быстрых и медленных процессов в замкнутой системе (7), (8), для анализа которых используем метод разделения движений. Применяя процедуру разделения движений [5, 6] к системе (7), (8), получим уравнения для подсистемы быстрых движений (ПБД) контура стабилизации тока

$$(9) \quad \begin{aligned} \mu_{\text{я}} \frac{dm_1}{dt} &= m_2, \\ \mu_{\text{я}} \frac{dm_2}{dt} &= \frac{k_{\text{я}} E_1}{4L_{\text{я}}} m_1 - d_{\text{я}} m_2 + k_{\text{я}} \left[\frac{i_{\text{я}}^d - I_{\text{я}}}{T_{\text{я}}} + \frac{R_{\text{я}}}{L_{\text{я}}} I_{\text{я}} + \frac{E_{\text{я}}}{L_{\text{я}}} - \frac{E_1}{4L_{\text{я}}} \right], \end{aligned}$$

где $m_1 = m$, $m_2 = \mu_{\text{я}} \dot{m}$. Величины E_1 , $E_{\text{я}}$, $I_{\text{я}}$ в данной системе рассматриваются как замороженные переменные на интервале времени переходных процессов в ПБД (9). Характеристический полином для ПБД (9) имеет вид

$$(10) \quad \mu_{\text{я}}^2 s^2 + d_{\text{я}} \mu_{\text{я}} s - k_{\text{я}} E_1 / (4L_{\text{я}}).$$

Выбирая коэффициент усиления $k_{\text{я}}$ регулятора (8), например, $k_{\text{я}} = -4L_{\text{я}}/E_1$, получим характеристический полином ПБД контура стабилизации тока вида $\mu_{\text{я}}^2 s^2 + d_{\text{я}} \mu_{\text{я}} s + 1$, где время переходных процессов в (9) задается выбором параметра $\mu_{\text{я}}$, например, $\mu_{\text{я}} = 0,0013$ с, а величина демпфирования быстрых процессов задается выбором параметра $d_{\text{я}}$, например, $d_{\text{я}} = 2$.

Таким образом, в силу данного выбора параметров регулятора (8) обеспечивается устойчивость процессов в ПБД (9). Тогда для квазиравновесного режима ПБД (9) после затухания переходных процессов получим $m_1 = m_1^s$, где

$$(11) \quad m_1^s = \frac{4L_{я}}{E_1} \left[\frac{i_{я}^d - I_{я}}{T_{я}} + \frac{R_{я}}{L_{я}} I_{я} + \frac{E_{я}}{L_{я}} - \frac{E_1}{4L_{я}} \right].$$

Подстановка $m = m_1 = m_1^s$ вида (11) в выражение (7) приводит к уравнению подсистемы медленных движений (ПМД) для замкнутой системы (7), (8) в виде дифференциального уравнения

$$(12) \quad \dot{I}_{я} = (i_{я}^d - I_{я}) / T_{я},$$

где параметр $T_{я}$ регулятора тока выбирается в зависимости от желаемого времени $t_{я}$ для переходных процессов по току $I_{я}$ в соответствии с условием $T_{я} \approx t_{я}/3$, например, полагаем $t_{я} \approx 0,03$ с. Разделение темпов быстрых и медленных процессов в системе (7), (8) обеспечивается выбором параметра $\mu_{я}$ таким образом, чтобы выполнялось условие $\mu_{я} \approx T_{я}/\eta_{я}$, где $\eta_{я}$ – степень разделения быстрых и медленных движений, например, $\eta_{я} > 8$.

Из выражения (10) следует, что отклонения величины индуктивности $L_{я}$ и величины питающего напряжения контактной сети E_1 от номинальных (расчетных) значений приводят к нарушению условию $k_{я}E_1/(4L_{я}) = -1$. При этом величина $k_{я}E_1/(4L_{я})$ может изменяться в широком диапазоне без нарушения устойчивости процессов в ПБД (9) и степени разделения темпов быстрых и медленных процессов в контуре стабилизации тока. Тем самым имеет место грубость свойства устойчивости процессов в контуре стабилизации тока по отношению к изменениям величины индуктивности $L_{я}$ и величины питающего напряжения контактной сети E_1 .

В результате, в силу свойств решений уравнения ПМД (12), обеспечивается требование (1) для средней величины тока $I_{я}$ в цепи якоря двигателя для равновесного режима контура стабилизации тока (в силу астатизма системы) при отсутствии переменных возмущающих воздействий в системе. Очевидно, что имеется отличие между мгновенной величиной тока $i_{я}$ и его желаемой величиной $i_{я}^d$, заданной на входе регулятора тока (рис. 2), обусловленное высокочастотной колебательной составляющей в процессах по току $i_{я}(t)$, что является следствием импульсного режима работы ШИМ. Однако влияние данной высо-

кочастотной колебательной составляющей в процессах по току на механические процессы по скорости двигателя пренебрежимо мало в силу относительно большой инерционности двигателя.

В тоже время переходные процессы по скорости вращения двигателя порождают изменение противо-ЭДС $E_{\text{я}}$ в цепи якоря двигателя и изменение задающего воздействия $i_{\text{я}}^d$ на входе регулятора тока (рис. 2), что порождает дополнительную ошибку в реализации условия (1). Тогда величину относительной ошибки реализации условия (1) в режиме линейной заводки для тока, $i_{\text{я}}^d = c_i t$, можно оценить соотношением

$$(13) \lim_{t \rightarrow \infty} [i_{\text{я}}^d(t) - I_{\text{я}}(t)] / c_i = T_{\text{я}} + 4T_{\text{я}}\mu_{\text{я}}d_{\text{я}}R_{\text{я}} / (k_{\text{я}}E_1).$$

Величину относительной ошибки реализации условия (1) в режиме линейной заводки для противо-ЭДС, $E_{\text{я}} = c_E t$, можно оценить соотношением

$$(14) \lim_{t \rightarrow \infty} [i_{\text{я}}^d(t) - I_{\text{я}}(t)] / c_E = 4T_{\text{я}}\mu_{\text{я}}d_{\text{я}} / (k_{\text{я}}E_1).$$

Соотношения (13), (14) являются следствием выражений (3.8) и (3.10) приведенных на стр.70–71 в [7].

Численное моделирование процессов, протекающих в контуре стабилизации тока для преобразователя (3) с ШИМ (4)–(6) и регулятором (8) было выполнено при следующих параметрах:

$$C_1 = C_2 = C_3 = C_4 = 0,002 \text{ Ф}; L_{\text{я}} = 0,0015 \text{ Гн}; \\ R_{\text{ин}} = 0,1 \text{ Ом}; R_{\text{я}} = 0,16 \text{ Ом}; T_s = 0,001 \text{ с}; T_{\text{я}} = 0,01 \text{ с}; \\ \mu_{\text{я}} = 0,0013 \text{ с}; d_{\text{я}} = 2; k_{\text{я}} = -4L_{\text{я}}/E_1 = -5 \cdot 10^{-7}; E_1 = 12 \text{ кВ}; E_{\text{я}} = 0; \\ k_1 = k_2 = 27,56.$$

На рис. 4 представлен график для мгновенной величины тока в цепи якоря двигателя. Задающее воздействие принимает значения 1 кА и 3 кА. На рис. 5 представлен график для мгновенной величины напряжения на конденсаторе C_1 . Из графика на рис. 4 видно, что поведение средней величины тока якоря соответствует решениям уравнения (12), где время переходных процессов по току примерно равно 0,03 с.

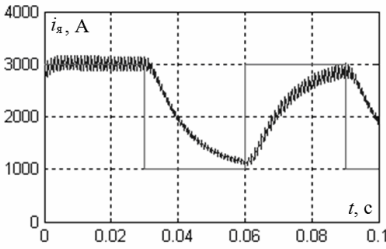


Рис. 4. Ток $i_{я}$ в цепи якоря

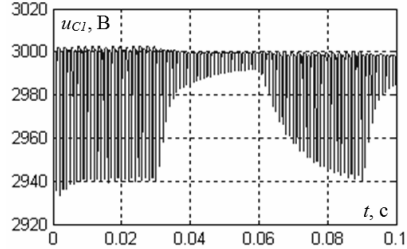


Рис. 5. Напряжение u_{C1} на конденсаторе C_1

5. Синтез регулятора скорости

Рассмотрим модель двух последовательно соединенных по электрической цепи тяговых двигателей колесной пары, как один эквивалентный двигатель постоянного тока. Дополним уравнение (7), описывающее поведение тока якоря двигателя, уравнением вращения ротора. Полагая, что $E_{я} = k_1 \omega$, рассмотрим модель эквивалентного двигателя постоянного тока независимо возбуждения при условии, что ток возбуждения постоянен [4]:

$$(15) \quad \begin{aligned} \frac{dI_{я}}{dt} &= -\frac{R_{я}}{L_{я}} I_{я} - \frac{k_1}{L_{я}} \omega + \frac{E_1}{4L_{я}} - \frac{E_1}{4L_{я}} m, \\ \frac{d\omega}{dt} &= \frac{k_2}{J} I_{я} - \frac{M_{наг}}{J}, \end{aligned}$$

где $k_1 = c_e c_\phi I_b$; $k_2 = c_m c_\phi I_b$; I_b – ток возбуждения; J – приведенный к валу двигателя суммарный момент инерции вращающихся частей; ω – угловая скорость вала двигателя; $M_{наг}$ – момент сопротивления нагрузки; c_e , c_m , c_ϕ – конструктивные коэффициенты.

Рассмотрим уравнения замкнутой системы (8), (15), где выполним замену $i_{я} = I_{я}$. Из уравнений (8), (15) получим уравнение ПБД вида (9) для контура стабилизации тока якоря, а соответствующее уравнение ПМД приобретет следующий вид:

$$(16) \quad \begin{aligned} \frac{dI_{я}}{dt} &= \frac{i_{я}^d - I_{я}}{T_{я}}, \\ \frac{d\omega}{dt} &= \frac{k_2}{J} I_{я} - \frac{M_{наг}}{J}. \end{aligned}$$

Время переходных процессов по току в цепи якоря составляет примерно 0,03 с (рис. 4), что во много раз меньше времени переходных процессов для скорости вращения двигателя $\omega(t)$. Поэтому можно считать, что в контуре стабилизации тока имеет место равновесный режим по отношению к процессам для скорости вращения двигателя $\omega(t)$, т.е. $I_{я} = i_{я}^d$. Соответственно, предполагая равновесный режим для процессов по току в системе (16), получим вместо (16) расчетную модель для контура стабилизации скорости $\omega(t)$ пониженного порядка:

$$(17) \quad \frac{d\omega}{dt} = \frac{k_2}{J} i_{я}^d - \frac{M_{наг}}{J},$$

где задающее воздействие $i_{я}^d$ регулятора тока якоря рассматривается как управляющее воздействие для контура стабилизации скорости вращения двигателя (рис. 2).

Регулятор контура стабилизации скорости должен обеспечивать выполнение условия (2). Рассмотрим алгоритм управления для данного контура в виде дифференциального уравнения

$$(18) \quad \mu_{\omega} \frac{d}{dt} i_{я}^d = k_{\omega} \left[\frac{\omega^d - \omega}{T_{\omega}} - \frac{d}{dt} \omega \right],$$

где μ_{ω} – малый параметр, $\mu_{\omega} > 0$; $T_{\omega} > 0$.

Выполняя преобразование Лапласа для уравнения (18) при нулевых начальных условиях, получим

$$i_{я}^d(s) = \frac{k_{\omega}}{\mu_{\omega}} \left\{ \frac{1}{sT_{\omega}} [\omega^d(s) - \omega(s)] - \omega(s) \right\}.$$

Из данного выражения следует, что алгоритм управления (18) является пропорционально-интегральным (ПИ) регулятором, соответственно является реализуемым без применения операции дифференцирования.

Наличие малого параметра приводит к тому, что в замкнутой системе (17), (18) будут возникать быстрые и медленные

процессы, для анализа которых снова использован метод разделения движений. Применяя процедуру разделения движений [5, 6] для замкнутой системы (17), (18), получим уравнение ПБД контура стабилизации скорости, имеющее вид

$$(19) \quad \mu_{\omega} \frac{d}{dt} i_{\alpha}^d = k_{\omega} \left[\frac{\omega^d - \omega}{T_{\omega}} - \frac{k_2}{J} i_{\alpha}^d + \frac{M_{наз}}{J} \right],$$

где ω рассматривается как замороженная переменная на интервале времени переходных процессов в ПБД (19). Характеристический полином ПБД (19) имеет вид $\mu_{\omega} s + k_{\omega} k_2 / J$. Выбирая коэффициент усиления k_{ω} для регулятора (18) из условия $k_{\omega} = J/k_2$, получим $\mu_{\omega} s + 1$.

В силу указанного выбора параметров регулятора (18), обеспечивается устойчивость процессов в ПБД (19) контура стабилизации скорости. Тогда для квазиравновесного режима ПБД (19) после затухания переходных процессов получим $i_{\alpha}^d = (i_{\alpha}^d)^s$, где

$$(20) \quad (i_{\alpha}^d)^s = \frac{J}{k_2} \left[\frac{\omega^d - \omega}{T_{\omega}} + \frac{M_{наз}}{J} \right].$$

Подстановка $i_{\alpha}^d = (i_{\alpha}^d)^s$ вида (20) в выражение (17) приводит к уравнению подсистемы медленных движений (ПМД) для замкнутой системы (17), (18) в виде следующего дифференциального уравнения:

$$(21) \quad \frac{d\omega}{dt} = \frac{\omega^d - \omega}{T_{\omega}}.$$

Параметр T_{ω} регулятора (18) выбирается в зависимости от желаемого времени t_{ω} для переходных процессов по скорости ω в соответствии с условием $T_{\omega} \approx t_{\omega}/3$. Разделение темпов быстрых и медленных процессов в системе (17), (18) обеспечивается выбором параметра μ_{ω} таким образом, чтобы выполнялось условие $\mu_{\omega} \approx T_{\omega}/\eta_{\omega}$, где η_{ω} – степень разделения быстрых и медленных движений. Иерархия темпов формируемых процессов в контурах стабилизации тока якоря и скорости вращения двигателя обеспечивается соответствующим выбором парамет-

ров алгоритмов управления с учетом условий $\mu_{\text{я}} \ll T_{\text{я}} \ll \mu_{\omega} \ll T_{\omega}$.

Уравнение (21) для подсистемы медленных движений контура стабилизации скорости не зависит от параметров преобразователя, двигателя и контактной сети. В результате, в силу свойств решений уравнения (21), обеспечивается требование (2) для равновесного режима контура стабилизации скорости (в силу астатизма системы) при отсутствии переменных возмущающих воздействий в системе. Однако изменения величины момента нагрузки $M_{\text{нар}}$ и задающего воздействия по скорости ω^d порождают дополнительную ошибку в реализации условия (2). Аналогично соотношениям (13) и (14), величину относительной ошибки реализации условия (2) в режиме линейной заводки для момента нагрузки, $M_{\text{нар}} = c_M t$, можно оценить соотношением

$$\lim_{t \rightarrow \infty} [\omega^d(t) - \omega(t)] / c_M = T_{\omega} \mu_{\omega} / (k_{\omega} k_2).$$

Величину относительной ошибки реализации условия (2) в режиме линейной заводки для скорости, $\omega_{\text{я}}^d = c_{\omega} t$, можно оценить соотношением

$$\lim_{t \rightarrow \infty} [\omega_{\text{я}}^d(t) - \omega_{\text{я}}(t)] / c_{\omega} = T_{\omega}.$$

Результаты численного моделирования для двигателя постоянного тока с преобразователем (3) и широтно-импульсным модулятором (4)–(6) на его входе, с регулятором тока (8) и регулятором скорости (18), представлены на рис. 6–9 и были получены при следующих параметрах системы: $R_{\text{я}} = 0,34 \text{ Ом}$; $L_{\text{я}} = 0,003 \text{ Гн}$; $J = 150 \text{ кг} \cdot \text{м}^2$; $C_1 = C_2 = C_3 = C_4 = 0,002 \text{ Ф}$; $R_{\text{ин}} = 0,1 \text{ Ом}$; $E_1 = 12 \text{ кВ}$; $T_s = 0,001 \text{ с}$; $k_1 = k_2 = 27,56$. В соответствии с представленными выше соотношениями, параметры регуляторов имеют следующие значения: $k_{\text{я}} = -4L/E_1 = -1 \cdot 10^{-6}$; $d_{\text{я}} = 2$; $\mu_{\text{я}} = 0,0013 \text{ с}$; $T_{\text{я}} = 0,01 \text{ с}$; $\mu_{\omega} = 0,1 \text{ с}$; $T_{\omega} = 1 \text{ с}$; $k_{\omega} = J/k_2 = 5,44$.

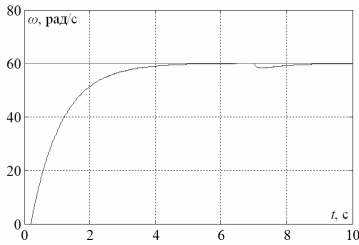


Рис. 6. Скорость вращения ω

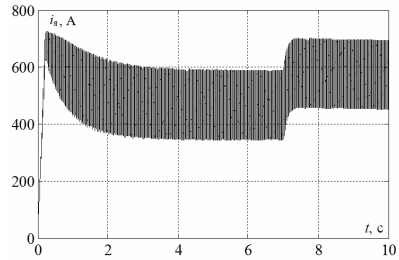


Рис. 7. Ток $i_{я}$ в цепи якоря

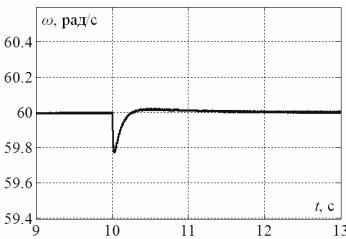


Рис. 8. Скорость вращения ω при изменении напряжения контактной сети

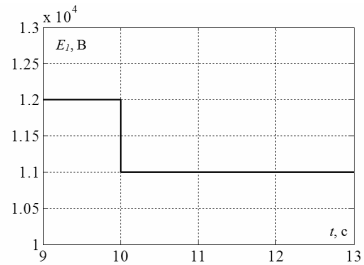


Рис. 9. Изменение напряжения контактной сети

Переходные процессы по скорости соответствуют решениям уравнения (21), где время переходного процесса примерно равно 3 с и задается параметром T_{ω} алгоритма управления (18). На рис. 6 в момент времени $t = 7$ с возникает отклонение скорости двигателя от заданной величины, что обусловлено ступенчатым изменением внешнего момента нагрузки $M_{наг}$ при $t = 7$ с от величины 9000 Н·м до 12000 Н·м. Возникающее возмущение скорости вращения двигателя компенсируется с помощью регулятора путем изменения тока цепи якоря (рис. 7). На рис. 8 показан график изменения скорости вращения двигателя при изменении напряжения контактной сети (рис. 9). Из результатов моделирования следует, что в обсуждаемой системе регулирования обеспечивается как стабилизация, так и формирование

заданного времени переходных процессов для скорости вращения двигателя при изменении внешнего момента нагрузки и напряжения контактной сети.

6. Заключение

Рассмотрена система управления для двигателя постоянного тока независимого возбуждения с многоуровневым преобразователем напряжения. Получена математическая модель многоуровневого преобразователя напряжения и предложена методика расчета двухконтурной системы управления скоростью вращения двигателя с предварительной стабилизацией величины тока якоря. В отличие от известных стандартных настроек систем автоматического управления электроприводами [1, 2], особенностью обсуждаемой методики синтеза является формирование разнотемповых процессов в замкнутой системе [7], что позволяет обеспечить заданные показатели качества переходных процессов при изменениях питающего напряжения контактной сети, переменной величине момента нагрузки, неполной информации о параметрах двигателя и преобразователя напряжения. Приведены расчетные соотношения для выбора параметров регуляторов тока якоря и скорости вращения двигателя и выполнено численное моделирование системы управления.

Литература

1. БАШАРИН А.В., НОВИКОВ В.А., СОКОЛОВСКИЙ Г.Г. *Управление электроприводами: Учебное пособие для вузов.* – Ленинград, Энергоиздат, Ленинградское отделение, 1982. – 392 с.
2. ВОРОНИН С.Г. *Математическое описание и стандартные настройки систем автоматического управления электроприводами.* [Электронный ресурс]. – Режим доступа: http://model.exponenta.ru/epivod/glv_070.htm (дата обращения: 12.04.2012).

3. ЗИНОВЬЕВ Г.С., ЛОПАТКИН Н.Н., ВАЙС Х. *Высоковольтный DC-DC конвертор для электровозов новой волны* // Электротехника. – 2009. – №12 – С. 46–51.
4. КИМ Д.П. *Теория автоматического управления. Т.1. Линейные системы.* – М.: ФИЗМАТЛИТ, 2003. – С. 78–79.
5. КРАСОВСКИЙ Н.Н. *Об устойчивости решений системы двух дифференциальных уравнений* // Прикладная математика и механика – 1953. – Т. 17. – Вып. 6. – С. 651–672.
6. ТИХОНОВ А.Н. *Системы дифференциальных уравнений, содержащие малые параметры при производных* // Математический сборник. – 1952. – Т. 31. – №3. – С. 575–586.
7. ЮРКЕВИЧ В.Д. *Синтез нелинейных нестационарных систем управления с разнотемповыми процессами.* – С.-Петербург: Наука, 2000. – 287 с.
8. SIRA-RAMIREZ H. *A geometric approach to pulse-width-modulated control in nonlinear dynamical systems* // *IEEE Trans. Automatic Control.* – 1989. – Vol. 34, №2 – P. 184–187.
9. YURKEVICH V.D., ZINOVIEV G.S., GORDEEV A.A. *PWM Current Controller Design for Multi-level DC-DC Converter via Singular Perturbation Technique* // *Proceedings of International Conference and Seminar on Micro/Nanotechnologies and Electron Devices EDM 2011, 12-th Annual. Erlagol, 2011.* – P. 390–398.

RESEARCH OF DC MOTOR CONTROL SYSTEM WITH MULTI-LEVEL VOLTAGE CONVERTER

Artem Gordeev, Novosibirsk State Technical University, Novosibirsk, student (artem.a.gordeev@gmail.com).

Valery Yurkevich, Novosibirsk State Technical University, Novosibirsk, Doctor of Science, professor ((383) 346-49-35, yurkev@ac.cs.nstu.ru).

Gennady Zinoviev, Novosibirsk State Technical University, Novosibirsk, Doctor of Science, professor ((383) 346-11-82, genstep@mail.ru).

Abstract: The problem of regulation for a DC motor with multi-level DC-DC converter is discussed. The control system is considered consisting of two feedback loops. In the first one the armature current control for a DC motor is provided by means of pulse-width modulated control for multi-level DC-DC converter. In the second one DC motor speed control is maintained. Proportional-integral (PI) controllers are designed for armature current and motor speed control based on singular perturbation technique such that multi-time-scale motions are artificially induced in the closed-loop system. Numerical simulations are performed in order to show efficacy of the proposed design technique.

Keywords: DC motor, DC-DC converter, pulse-width modulation, PI controller, singular perturbation method.

*Статья представлена к публикации
членом редакционной коллегии В. А. Уткиным*

УДК 519.71
ББК 32.817

УСТОЙЧИВОСТЬ СИСТЕМ С БЛОКАМИ НЕЧЕТКОГО ЛОГИЧЕСКОГО ВЫВОДА В ОБЪЕКТЕ УПРАВЛЕНИЯ

Усков А. А.¹

(Российский университет кооперации, Москва)

Предлагается подход к анализу абсолютной устойчивости одного класса систем с блоками нечеткого логического вывода в объекте управления. Разработанные методы исследования систем доведены до уровня простых и удобных в инженерной практике методик.

Ключевые слова: асимптотическая устойчивость, нечеткий логический вывод, система автоматического управления.

1. Введение

В настоящее время достаточно широкое распространение получили нечеткие системы управления. Применение нечеткой логики позволяет использовать субъективные знания (мнения) экспертов, что очень полезно при построении моделей сложных или плохо формализуемых объектов и процессов [3, 5].

В то же время блок нечеткого логического вывода с точки зрения теории автоматического управления представляет собой статическое звено с очень сложной нелинейной характеристикой, что значительно осложняет исследование систем с такими звеньями [5].

¹ Андрей Александрович Усков, доктор технических наук, профессор (andrey@uskov.net, www.uskov.net).

В статье предложен подход к анализу абсолютной устойчивости одного класса систем с блоками нечеткого логического вывода в объекте управления.

2. Постановка задачи

Рассмотрим замкнутую автономную нелинейную импульсную систему автоматического управления, приведенную на рис. 1.

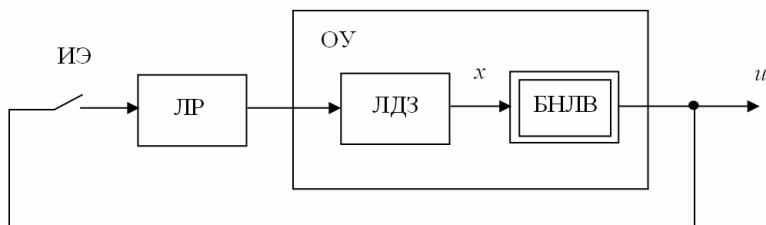


Рис. 1. Система управления с односвязным объектом

Система состоит из импульсного элемента с амплитудно-импульсной модуляцией ИЭ, линейного регулятора ЛР и объекта управления ОУ. Структурно объект управления состоит из последовательного соединения линейного динамического звена ЛДЗ и статического нелинейного элемента – блок нечеткого логического вывода (БНЛВ).

Зависимость между выходным u и входным x сигналами нелинейного элемента задается набором нечетких продукционных правил:

P_1 : если x есть A_1 , то $u = a_1$;

P_2 : если x есть A_2 , то $u = a_2$;

.....

P_N : если x есть A_N , то $u = a_N$,

где A_1, A_2, \dots, A_N – нечеткие множества, определенные на множестве действительных чисел R ; a_1, a_2, \dots, a_N – действительные числа.

Выходной сигнал НЭ u рассчитывается в соответствии с алгоритмом нечеткого вывода Сугэно (Sugeno) нулевого порядка [3, 5]:

$$(1) \quad u(x) = \frac{\sum_{i=1}^N a_i \mu_i(x)}{\sum_{i=1}^N \mu_i(x)},$$

где $\mu_i(x)$ – функции принадлежности нечетких множеств A_i .

Необходимо получить достаточное условие асимптотической устойчивости системы с БНЛВ со структурой, приведенной на рис. 1

3. Характеристика блока нечеткого логического вывода

Представим характеристику БНЛВ (1) в виде

$$(2) \quad u(x) = k(x) \cdot x,$$

где $k(x)$ – коэффициент передачи БНЛВ, зависящий от входного сигнала x .

В качестве характеристики нелинейной зависимости БНЛВ выберем коэффициент

$$(3) \quad K_H = \max_{\forall x \in R} |k(x)|,$$

Ввиду сложности определения численного значения K_H для произвольного БНЛВ, воспользуемся его оценкой.

Решая совместно (1)–(3), получаем

$$(4) \quad K_H = \max_{x \in R} \frac{\left| \sum_{i=1}^N a_i \mu_i(x) \right|}{|x| \cdot \left| \sum_{i=1}^N \mu_i(x) \right|}.$$

Воспользовавшись свойством функции «модуль» [1], из предыдущего выражения получим:

$$(5) \quad K_H \leq \max_{x \in R} \frac{\sum_{i=1}^N |a_i| \cdot \mu_i(x)}{|x| \cdot \sum_{i=1}^N \mu_i(x)}.$$

Отметим далее, что множитель

$$\frac{\sum_{i=1}^N |a_i| \cdot \mu_i(x)}{\sum_{i=1}^m \mu_i(x)} = X_c$$

определяет, по сути, координаты центра масс X_c невесомого стержня с расположенными на нем грузами с массами $\mu_i(x)$ в точках с координатами $|a_i|$, где $i = 1, 2, \dots, N$. Один из возможных вариантов расположения грузов на стержне показан на рис. 2.

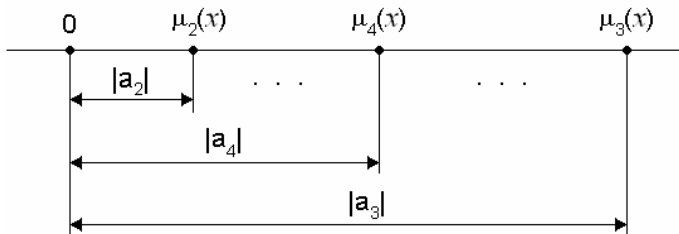


Рис. 2. Иллюстрация к выводу формулы

Очевидно, что координата точки центра масс X_c не может превышать координаты крайнего справа груза, имеющего массу, отличную от 0.

На основании свойства центра масс [2], можно записать:

$$(6) \quad K_H \leq \max_{x \in R} \frac{1}{|x|} \cdot \max(|a_1| \cdot 1_0(\mu_1(x)), |a_2| \cdot 1_0(\mu_2(x)), \dots, |a_n| \cdot 1_0(\mu_N(x))),$$

где $1_0(t) = \begin{cases} 1 & \text{при } t > 0, \\ 0 & \text{при } t = 0; \end{cases}$ – единичная функция.

Последнее выражение можно привести к виду

$$(7) \quad K_{H0} = \max \left(\max_{x \in B_1} \frac{|a_1|}{|x|}, \max_{x \in B_2} \frac{|a_2|}{|x|}, \dots, \max_{x \in B_N} \frac{|a_N|}{|x|} \right) \geq K_H,$$

где B_1, B_2, \dots, B_N – носители нечетких множеств A_1, A_2, \dots, A_N соответственно [3].

Методику использования соотношения (7) поясним на примере.

Пример 1

Рассмотрим БНЛВ, описываемый набором нечетких правил:

- П₁: если x есть Z , то $u = 0$;
- П₂: если x есть PS , то $u = 1$;
- П₃: если x есть PM , то $u = 2$;
- П₄: если x есть NS , то $u = -1$;
- П₅: если x есть NM , то $u = -2$.

На рис. 3 приведены функции принадлежности нечетких множеств Z, PS, PM, NS и NM .

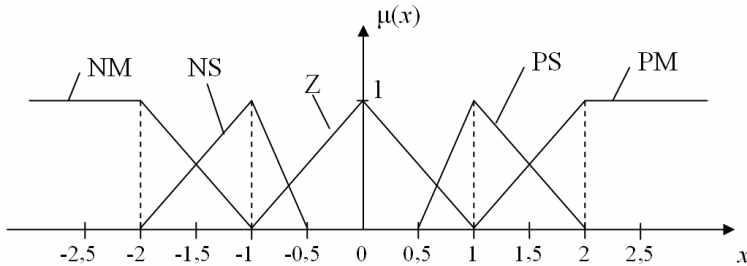


Рис. 3. Функции принадлежности нечетких множеств

Рассмотрим применение формулы (7) в данном случае.

Для правила П₁ носитель нечеткого множества Z – это $x \in [-1, 1]$, значение $a_1 = 0$, следовательно $\max_{x \in [-1, 1]} \frac{|0|}{|x|} = \lim_{|x| \rightarrow 0} \frac{0}{|x|}$.

Анализ приведенного предела показывает, что его значение не превышает 1.

Для правила Π_2 носитель нечеткого множества PS – $x \in [0,5, 2]$, значение $a_1 = 1$, следовательно $\max_{x \in [0,5, 2]} \frac{|1|}{|x|} = 2$.

Для правила Π_3 носитель нечеткого множества PM – $x \in [1, +\infty)$, значение $a_1 = 2$, следовательно $\max_{x \in [1, +\infty)} \frac{|2|}{|x|} = 2$.

Для правила Π_4 носитель нечеткого множества NS – $x \in [-2, -0,5]$, значение $a_1 = -1$, следовательно $\max_{x \in [-2, -0,5]} \frac{|-1|}{|x|} = 2$.

Для правила Π_5 носитель нечеткого множества NM – $x \in (-\infty, -1]$, значение $a_1 = -2$, следовательно $\max_{x \in (-\infty, -1]} \frac{|-2|}{|x|} = 2$.

Подставляя данные частные результаты в выражение (7), получим $K_{H0} = \max(1, 2, 2, 2, 2) = 2$.

4. Достаточное условие устойчивости

Используя описанную методику определения численного значения величины K_{H0} , можно получить достаточное условие асимптотической устойчивости системы на рис. 1.

Если характеристика БНЛВ находится в 1 и 3 квадрантах, а импульсная линейная часть системы устойчива, к рассматриваемой системе применим геометрический критерий абсолютной устойчивости Я.З. Цыпкина [2, 6]. Достаточное условие асимптотической устойчивости в целом положения равновесия в данном случае определяется неравенством:

$$(8) \quad \operatorname{Re} W^*(j\bar{w}) + \frac{1}{K_{H0}} > 0,$$

где $W^*(j\bar{w})$ – комплексный коэффициент передачи последовательно соединенных импульсного элемента ИЭ, линейного регулятора ЛР и линейного динамического звена ЛДЗ.

Отметим, что области устойчивости, полученные с помощью критерия Я.З. Цыпкина, не уже областей, полученных с

помощью второго метода А.М. Ляпунова с квадратичной функцией [2, 6].

Пример 2

Рассмотрим автономную систему, приведенную на рис. 1.

Последовательно соединенные ЛР и ЛДЗ описывается передаточной функцией $W(p) = \frac{k}{1 + p \cdot T}$. Импульсный элемент с

фиксатором нулевого порядка имеет период квантования T_0 . Используется БНЛВ из примера 1.

Для системы на рис. 1 справедливо разностное уравнение $x_{m+1} = e^{-\frac{T_0}{T}} \cdot x_m + u_m \cdot k \cdot (1 - e^{-\frac{T_0}{T}})$ [1]. Комплексный коэффициент передачи импульсной линейной части системы

$$W^*(j\omega) = \frac{(1 - e^{-\frac{T_0}{T}}) \cdot k}{e^{j\omega} - e^{-\frac{T_0}{T}}}$$

В соответствии с неравенством (8) не-

сложно получить достаточное условие асимптотической устойчивости в целом положения равновесия рассматриваемой системы:

$$k < \frac{1 + e^{-\frac{T_0}{T}}}{(1 - e^{-\frac{T_0}{T}}) \cdot K_{H0}}$$

Приняв $K_{H0} = 2$ (из примера 1) и $T_0 = 0,1$, определим область устойчивости рассматриваемой системы рис. 4 (область ниже линии 1, показана заштриховкой). Для сравнения на данном рисунке показана также действительная область устойчивости системы (область ниже линии 2).

Согласно рис. 4 предлагаемое достаточное условие позволяет определить около 50% истинной области асимптотической устойчивости в пространстве параметров k и T .

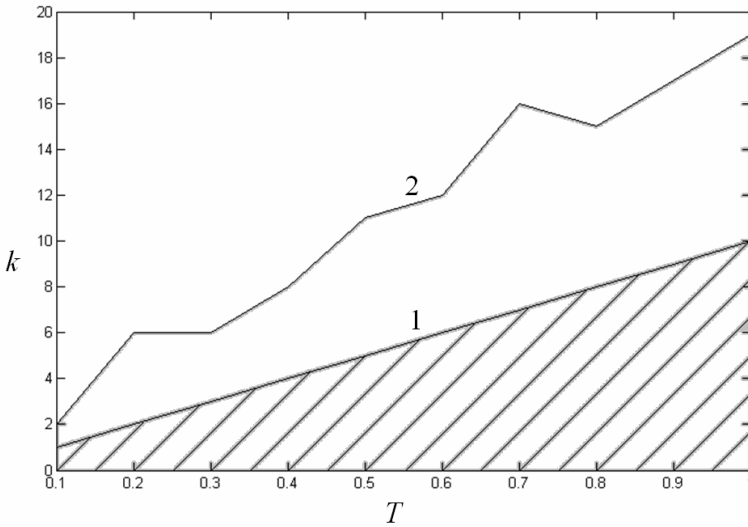


Рис. 4. Область устойчивости системы

5. Система с несколькими БНЛВ

Обобщим приведенные выше результаты на случай системы с несколькими БНЛВ.

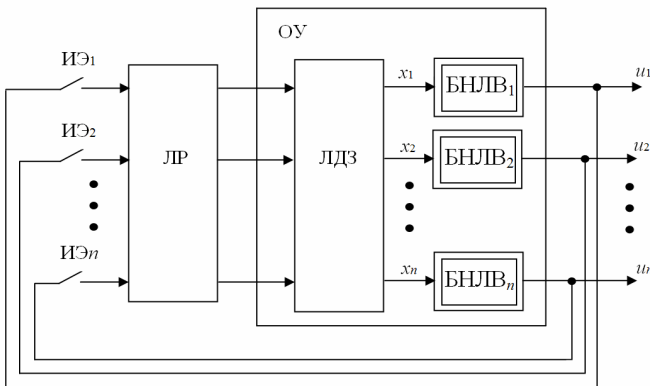


Рис. 5. Система управления с многосвязным объектом

Рассмотрим автономную импульсную систему с многосвязным объектом управления (рис. 5).

Система состоит из импульсных элементов с амплитудно-импульсной модуляцией ИЭ_{*i*}, *i* = 1, 2, ..., *n*, работающих синхронно, линейного регулятора ЛР и объекта управления ОУ. Структурно объект управления состоит из линейного динамического звена ЛДЗ и статических нелинейных элементов БНЛВ_{*i*}, *i* = 1, 2, ..., *n*.

Допустим, что характеристики БНЛВ описываются наборами нечетких продукционных правил и находятся в 1 и 3 квадрантах. Для каждого из БНЛВ по формуле (7) определены коэффициенты K_{H0i} , *i* = 1, 2, ..., *n*, которые представлены в виде матрицы

$$(9) \quad \mathbf{K}_{H0} = \begin{bmatrix} K_{H01} & 0 & \dots & 0 \\ 0 & K_{H02} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & K_{H0n} \end{bmatrix}.$$

Согласно критерию Я.З. Цыпкина для систем с несколькими нелинейностями [6], если импульсная линейная часть системы на рис. 4 устойчива и существует действительное число *p* такое, что выполняется неравенство

$$(10) \quad \frac{1}{2} \left[\mathbf{K}_{H0}^p \cdot \mathbf{W}^*(j\bar{\omega}) + \mathbf{W}^*(-j\bar{\omega}) \cdot \mathbf{K}_{H0}^p \right] + \mathbf{K}_{H0}^{p-1} > 0,$$

$$0 \leq \bar{\omega} \leq \pi,$$

где $\mathbf{W}^*(j\bar{\omega})$ – матрица комплексных коэффициентов передачи импульсной линейной части системы, то рассматриваемая система будет асимптотически устойчива в целом.

6. Выводы

В статье предложено достаточное условие асимптотической устойчивости в целом систем управления с одновязными блоками нечеткого логического вывода Сугэно нулевого порядка, позволяющее определять область устойчивости в пространстве параметров системы.

Разработанный метод исследования может быть полезен при анализе и синтезе систем управления с нечеткой логикой.

Литература

1. БРОНШТЕЙН И.Н., СЕМЕНДЯЕВ К.А. *Справочник по математике для инженеров и учащихся вузов.* – М.: Наука, 1981. – 723 с.
2. ВИДАЛЬ П. *Нелинейные импульсные системы.* – М.: Энергия, 1974. – 336 с.
3. КРУГЛОВ В.В., ДЛИ М.И. *Интеллектуальные информационные системы: компьютерная поддержка систем нечеткой логики и нечеткого вывода.* – М.: Издательство физико-математической литературы, 2002. – 256 с.
4. КРУГЛОВ В.В., УСКОВ А.А. *Достаточное условие устойчивости замкнутых систем управления с нечеткими логическими регуляторами // Известия РАН. Теория и системы управления.* – 2004. – №4. – С. 47–51.
5. УСКОВ А.А., КУЗЬМИН А.В. *Интеллектуальные технологии управления. Искусственные нейронные сети и нечеткая логика.* – М.: Горячая линия – Телеком, 2004. – 143 с.
6. ЦЫПКИН Я.З., ПОПКОВ Ю.С. *Теория нелинейных импульсных систем.* – М.: Наука, 1973. – 416 с.

STABILITY OF SYSTEMS WITH BLOCKS OF FUZZY INFERENCE IN PLANT

Andrey Uskov, Russian University of Cooperation, Moscow,
Doctor of Science, professor.

Abstract: The technique is suggested of analysis of terrain clearance stability of one class of systems with blocks of fuzzy inference in the plant. Simple and convenient engineering routines are developed on the basis of suggested techniques.

Keywords: asymptotic stability, fuzzy logic conclusion, system of automatic control

*Статья представлена к публикации
членом редакционной коллегии А. К. Погодаевым*

УДК 021.8 + 025.1

ББК 78.34

МОДЕЛЬ ЧИСЛЕННОЙ ОПТИМИЗАЦИИ СТРУКТУРЫ ОПЕРАЦИОННОГО ЯДРА ОРГАНИЗАЦИИ

Воронин А. А.¹, Харитонов М. А.²

(Волгоградский государственный университет, Волгоград)

Построена модель оптимизации операционного ядра организации, которое состоит из базового технологического модуля и сети модулей вспомогательных производств. Производственная функция операционного ядра представлена в виде суперпозиции производственных функций Леонтьева, отвечающих каждому из модулей. Задача оптимизации сведена к задаче линейного программирования с параметром, описывающим структуру операционного ядра. Разработан алгоритм автоматического построения исходных уравнений для каждого значения параметра. Представлены результаты численного решения оптимизационной задачи в широком диапазоне переменных и параметров модели. Данная модель может служить основой синтеза обобщенных механизмов управления организационной системой на большом интервале времени.

Ключевые слова: организационная система, операционное ядро, оптимизация структуры, производственная функция, линейное программирование.

1. Введение

Построение моделей управления организационной системой (ОС) на большом интервале времени на базе представлен-

¹ Александр Александрович Воронин, доктор физико-математических наук, профессор (a.voronin@volsu.ru).

² Михаил Алексеевич Харитонов, студент (kharitonov.mihail@gmail.com).

ной в [10] системы механизмов управления невозможно без учета ее структурно-функциональной изменчивости. Изменчивость внутренних и внешних связей, краткосрочных, среднесрочных и долгосрочных целей и критериев развития ОС мотивирует обобщение не только моделей базовых механизмов [10], но и обобщение термина «механизм управления» путем введения в его состав в дополнение к собственно методам принятия решений также и комплекса организационных структур, норм, параметров ОС, изменяющихся в течение долгосрочного периода управления согласно актуальной и проектируемой динамике ОС с учетом реализации временной последовательности простых и комплексных механизмов управления [10]. Такие обобщенные механизмы управления представляют собой сложно-структурированные системы, подсистемами и элементами которых являются обобщенные, комплексные или простые механизмы, оптимальность каждого из которых, таким образом, определяется уже в контексте задачи оптимизации всей системы – обобщенного механизма управления.

Эта концепция фактически совпадает с концепцией организационной структуры Г. Минцберга [8], поэтому термины «организационная структура» (в смысле Г. Минцберга) и «обобщенный механизм управления» являются синонимичными, а структурные конфигурации Г. Минцберга можно назвать оптимальными (в рамках формалистики менеджмента) обобщенными механизмами управления. Структурные и организационные изменения ОС можно таким образом трактовать как переходы между условно оптимальными структурными конфигурациями или обобщенными механизмами управления, выделяя в них (в соответствии с введенными признаками) траектории адаптации, развития и др. Развитие ОС при этом можно связывать с взаимодействием и развитием ее обобщенных механизмов управления.

Этапами синтеза обобщенных механизмов, т.е. этапами решения задач моделирования структурной динамики и динамической структурной оптимизации ОС являются построение семейств соответствующих уравнений динамики и критерия эффективности в структурнозависимой форме, описание пространства структурных состояний, корректное введение «прин-

ципов» динамики, разработка эффективных алгоритмов оптимизации. В качестве первых успешных шагов в этом направлении можно указать на работы [1–3, 5, 6, 9], в которых исследуются задачи статической и динамической оптимизации иерархий управления ОС с критерием эффективности в виде структурно зависимой функции затрат. Для синтеза обобщенных механизмов управления необходимо сопряжение этих моделей со структурно зависимой производственной функцией (ПФ) ОС с теоретико-игровыми моделями иерархического управления. Сказанное мотивирует актуальность синтеза системы многопараметрических динамических моделей оптимизации и управления, в совокупности образующих иерархическую структуру обобщенных механизмов управления ОС.

В настоящей работе представлена математическая модель оптимизации операционного ядра организации, под которым согласно терминологии Г. Минцберга [8] будем понимать совокупность структурных единиц ОС, непосредственно осуществляющих основное и вспомогательные производства (включая персонал и оборудование). Структура операционного ядра представлена в виде ориентированного графа специального вида с переменным числом слоев. Вершины графа обозначают технологические модули базового и вспомогательных производств, ребра – входы и выходы технологических модулей – «потoki» факторов производства. Целевая функция – производственная функция операционного ядра – представлена в структурнозависимой форме в виде суперпозиции производственных функций Леонтьева, отвечающих каждому из технологических модулей.

При заданных входах (факторах производства) и структуре операционного ядра значение производственной функции является решением задачи линейного программирования.

Увеличение числа вспомогательных производств, с одной стороны, обеспечивает устойчивость выпуска при нарушении оптимальных пропорций между факторами производства, с другой – увеличивает себестоимость. Альтернативой усложнения структуры операционного ядра является покупка недостающих факторов на рынке (при наличии такой возможности). В экономиках с высоко развитой институциональной средой

последний способ оказывается наиболее эффективным, что приводит к упрощению структуры ОС и появлению специализированных аутсорсинговых организаций, реализующих не только вспомогательные производства, но и ранее традиционные функции ОС («универсальная бухгалтерия», «универсальный офис-менеджер» и т.п.). А в экономиках с низким уровнем развития институциональной среды преобладает противоположная тенденция – вертикальная интеграция путем поглощения смежных и вспомогательных производств. Однако в предложенной в настоящей работе модели не исследуется задача поиска оптимального баланса между указанными тенденциями, и оптимизация структуры операционного ядра заключается в поиске минимального числа слоев ее структуры, обеспечивающего максимальное для заданных входов (факторов производства) значение производственной функции (то есть не увеличивающееся при дальнейшем наращивании числа слоев).

Для решения задачи разработан алгоритм автоматического построения исходных уравнений для каждого числа слоев. Представлены результаты численного решения задачи в широком диапазоне переменных и параметров модели.

2. Структура и производственная функция операционного ядра ОС

Построенная в работе модель структурно зависимой ПФ основана на моделях [4, 6]. В качестве элементарной ПФ каждого элемента структуры операционного ядра с неизменной технологией (простого преобразователя – ПП) используется ПФ Леонтьева $F = k \min(f_1/a_1, \dots, f_n/a_n)$ где f_i – величины аргументов – трансформационных факторов производства; a_1, \dots, a_n – технологические коэффициенты; k – нормирующий множитель. В настоящей работе будем использовать агрегированную макроэкономическую трехфакторную модель трансформационных факторов (природный – N , технический – T , человеческий – H). Учет векторной природы факторов производства влечет лишь увеличение числа переменных и уравнений модели.

В условиях устойчивых межфакторных диспропорций или флуктуаций трансформационных факторов среднее значение ПФ ПП значительно ниже ее максимума. В рамках ОС эластичность ПФ по аргументам обеспечивается построением вертикальной структуры вспомогательных производств, производящих недостающие для максимальной эффективности базового ПП части трансформационных факторов. Таким образом, структуру факторных производственных потоков операционного ядра ОС можно представить в виде сложного преобразователя (СП) – многоуровневой структуры ПП, в котором выходы одних ПП являются входами других. Структура СП имеет вид ориентированного графа, вершинам которого отвечают ПП, а ребрам – факторные потоки. Последняя вершина отвечает базовому, а остальные – вспомогательным производствам. Исходные факторы производства являются входами всех ПП, выходом последнего (базового) ПП является продукт ОС.

Задача оптимизации факторных пропорций ПФ ПП имеет вид

$$(1) \quad F = S \min \left(\frac{N}{A}, \frac{T}{B}, \frac{H}{C} \right) \rightarrow \max_{N, T, H}, \quad N + T + H = R, \quad S = A + B + C$$

и очевидное решение:

$$N = \frac{RA}{S}, \quad T = \frac{RB}{S}, \quad H = \frac{RC}{S}, \quad F = R.$$

На рис. 1 приведена структура ПФ, на рис. 2 – графики ПФ ПП для оптимальных и неоптимальных факторных пропорций.

На рис. 2 приведены некоторые графики функции (1).

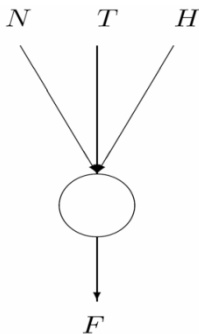


Рис. 1. Простой преобразователь

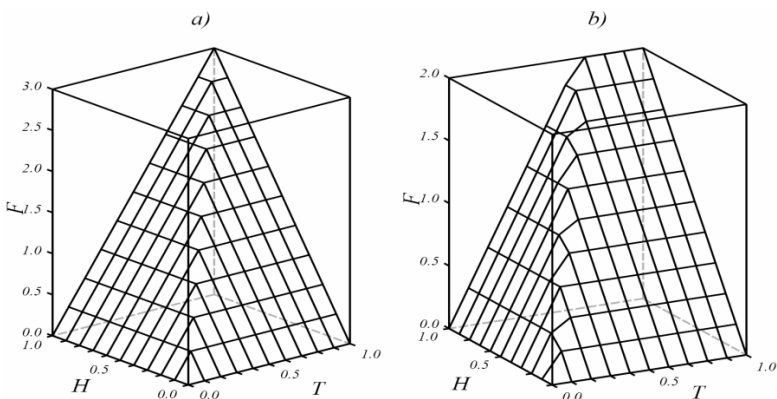


Рис. 2. График $F(1, T, H)$: а) $A = B = C = 1$;
б) $A = 1; B = 2; C = 3$

3. Задача оптимизации структуры операционного ядра ОС.

Рассмотрим СП с факторами производства N, TN, HN (TN и HN соответственно части факторов T и H , участвующие в производстве недостающей части фактора N) и возможностью производства части фактора N на n последовательных ПП (рис. 3).

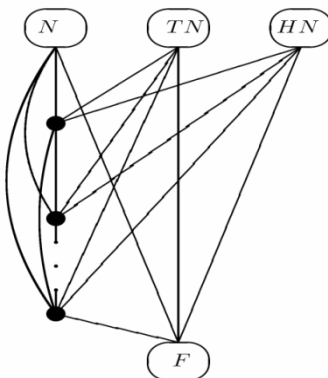


Рис. 3. Сложный преобразователь с производством части фактора N

ПФ этого СП задается следующей системой уравнений:

$$\sum_{l=1}^n N_l^0 + N_F^0 = N, \quad \sum_{l=1}^n T_l + T_F = TN, \quad \sum_{l=1}^n H_l + H_F = HN,$$

$$N^i = \sum_{k=i+1}^n N_k^i = (a_i + b_i + c_i) \min \left(\frac{1}{a_i} \sum_{l=0}^{i-1} N_l^i, \frac{T_i}{b_i}, \frac{H_i}{c_i} \right),$$

(2)

$$N_F^n = (a_n + b_n + c_n) \min \left(\frac{1}{a_n} \sum_{k=0}^{n-1} N_k^n, \frac{T_n}{b_n}, \frac{H_n}{c_n} \right),$$

$$F = S \min \left(\frac{N_F^n + N_F^0}{A}, \frac{T_F}{B}, \frac{H_F}{C} \right).$$

Здесь первые три уравнения описывают распределение исходных факторов производства по структуре СП. Четвертое уравнение дает балансовые соотношения для i -го вспомогательного ПП; N^i – общая величина фактора, произведенного на i -м ПП; N_k^i – величина фактора, произведенного на i -м и используемого на k -м ПП, $i = 1, \dots, n-1$. Пятое уравнение системы отвечает последнему вспомогательному ПП ($i = n$).

Из переменных системы (2) составим вектор факторных потоков СП

$$\Phi = \left(N_1^0, N_2^0, \dots, N_{n_1}^0, T_1, H_1, N_2^1, N_2^1, \dots, N_{n_1}^1, T_2, H_2, \dots \right. \\ \left. \dots, N_{n_1}^{n_1-1}, T_{n_1}, H_{n_1}, N_F^n, N_F^0, T_F, H_F, \varphi_P \right).$$

Задача оптимизации ПФ СП (2) с переменной структурой при фиксированных N, TN, HN имеет вид:

$$(3) \quad F = S \min \left(\frac{N_F^n + N_F^0}{A}, \frac{T_F}{B}, \frac{H_F}{C} \right) \rightarrow \max_{\Phi, n}.$$

Заметим, что задача (1) эквивалентна следующей задаче линейного программирования (ЛП):

$$-\frac{S}{A}N + F \leq 0, \quad -\frac{S}{B}T + F \leq 0, \quad -\frac{S}{C}H + F \leq 0, \quad F \rightarrow \max_{N, T, H}.$$

Таким же образом задача (2)–(3) сводится к следующей задаче параметрического ЛП:

$$(4) \quad \varphi_p(n) \rightarrow \max_{n, \Phi}, \quad \sum_{j=1}^p M_{ij} \varphi_j \leq \beta_i, \quad i = 1, \dots, u(n),$$

где $p = p(n)$ – число переменных; $u(n)$ – число неравенств, определяемых следующими формулами:

$$p(n) = \frac{n^2 + 5n + 8}{2}, \quad u(n) = 3n + 6.$$

В частности, для $n = 1, a_i = b_i = c_i = 1$ имеем:

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 & -3 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & 1 \\ -3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\Phi = (N_1^0, T_1, H_1, N_F^1, N_F^0, T_F, H_F, \varphi_P), \quad \beta = (N, TN, HN, 0, 0, 0, 0, 0).$$

Обозначим $\varphi_p^*(n)$ – решение задачи (4). Оптимальной структурой СП, описываемого системой (2), будем считать величину

$$n^* = \min_n \left(\text{Arg max}_n \varphi_p^* \right).$$

Обозначим $\lambda = (N + TN + HN)/F$. Внешними параметрами задачи (4) являются исходные межфакторные пропорции $N:TN:HN$ и пропорции технологических коэффициентов $a_i:b_i:c_i$. Для каждого СП величины a_i, b_i, c_i предполагаются заданными. В численных экспериментах эти величины задавались случайным образом, с сохранением их значений при варьировании числа слоев.

Результаты решения задачи (4) для некоторых значений переменных N, TN, HN и параметров a_i, b_i, c_i , принимающих случайные значения из отрезка $K \subset R^+$, представлены на рис. 4.

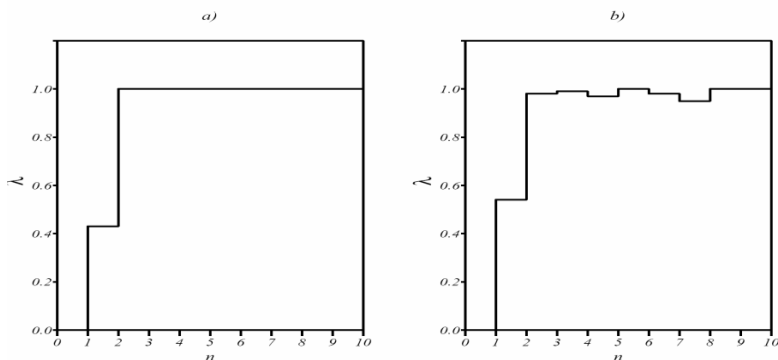


Рис. 4: График $\lambda(n)$:

- a) $N = 1, TN = HN = 10, K = \{1\}, n^* = 2,$
 b) $N = 1, TN = HN = 10, K = [1, 2], n^* = 3$

Структура полного СП с возможностью вспомогательного производства каждого из трансформационных факторов представлена на рис. 5.

ПФ СП на рис. 5 состоит из трех СП (см. рис. 3), каждый из которых соответственно производит недостающие части факторов N, T, H . На рис. 5 обозначено: TN, HN – соответственно части факторов T и H , участвующие в производстве недостающей части фактора N ; TT, HT – соответственно части факторов T и H ,

участвующие в производстве недостающей части фактора T ; TN, HN – соответственно части факторов T и H , участвующие в производстве недостающей части фактора H .

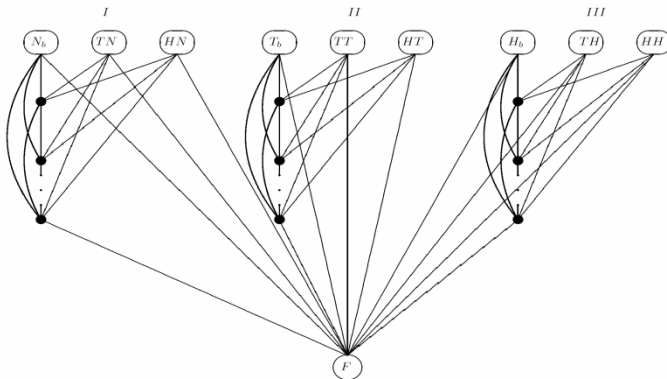


Рис. 5. Полный сложный преобразователь

ПФ этого СП имеет вид:

$$F = S \min \left(\frac{N}{A}, \frac{T}{B}, \frac{H}{C} \right),$$

$$(5) \quad N = N_F^{n_1} + N_F^0,$$

$$T = T_F^{n_2} + T_F^0 + TN_F + TT_F + TH_F,$$

$$H = H_F^{n_3} + H_F^0 + HN_F + HT_F + HH_F.$$

I:

$$\sum_{l=1}^{n_1} N_l^0 + N_F^0 = N, \quad \sum_{l=1}^{n_1} TN_l + TN_F = TN, \quad \sum_{l=1}^{n_1} HN_l + HN_F = HN,$$

$$N^i = \sum_{k=i+1}^{n_1} N_k^i = (a_i^1 + b_i^1 + b_i^1) \min \left(\frac{1}{a_i^1} \sum_{l=0}^{i-1} N_l^i, \frac{TN_i}{b_i^2}, \frac{HN_i}{c_i^2} \right),$$

$$N_F^{n_2} = (a_{n_1}^1 + b_{n_1}^1 + c_{n_1}^1) \min \left(\frac{1}{a_{n_1}^1} \sum_{k=0}^{n_1-1} N_k^1, \frac{TN_{n_2}}{b_{n_1}^2}, \frac{HN_{n_2}}{c_{n_1}^2} \right);$$

II:

$$\sum_{l=1}^{n_2} T_l^0 + T_F^0 = T_b, \quad \sum_{l=1}^{n_2} TT_l + TT_F = TT, \quad \sum_{l=1}^{n_2} HT_l + HT_F = HT,$$

$$T^i = \sum_{k=i+1}^{n_2} T_k^i = (a_i^2 + b_i^2 + b_i^2) \min \left(\frac{1}{a_i^2} \sum_{l=0}^{i-1} T_l^i, \frac{TT_i^2}{b_i^2}, \frac{HT_i^2}{c_i^2} \right),$$

$$T_F^{n_2} = (a_{n_2}^2 + b_{n_2}^2 + c_{n_2}^2) \min \left(\frac{1}{a_{n_2}^2} \sum_{k=0}^{n_2-1} N_{n_2}^k, \frac{TT_{n_2}^2}{b_{n_2}^2}, \frac{HT_{n_2}^2}{c_{n_2}^2} \right);$$

III:

$$\sum_{l=1}^{n_3} H_l^0 + H_F^0 = H_b, \quad \sum_{l=1}^{n_3} TH_l + TH_F = TH, \quad \sum_{l=1}^{n_3} HH_l^3 + HH_F^3 = HH,$$

$$H^i = \sum_{k=i+1}^{n_3} H_k^i = (a_i^3 + b_i^3 + b_i^3) \min \left(\frac{1}{a_i^3} \sum_{l=0}^{i-1} H_l^i, \frac{TH_i}{b_i^3}, \frac{HH_i}{c_i^3} \right),$$

$$H_F^{n_3} = (a_{n_3}^3 + b_{n_3}^3 + c_{n_3}^3) \min \left(\frac{1}{a_{n_3}^3} \sum_{k=0}^{n_3-1} H_{n_3}^k, \frac{TH_{n_3}}{b_{n_3}^3}, \frac{HH_{n_3}}{c_{n_3}^3} \right).$$

В обозначении технологических коэффициентов нижний индекс n_j , $j = 1, 2, 3$, обозначает номер слоя в j -м СП, что соответствует нумерации I, II, III на рис. 5. Верхний индекс (так же как и j) обозначает номер СП.

При заданных входах задача оптимизации структуры ОЯ с ПФ (5) имеет вид

$$(6) \quad F = S \min \left(\frac{N}{A}, \frac{T}{B}, \frac{H}{C} \right) \rightarrow \max_{\Phi, n},$$

$$\Phi = (\Phi^1, \Phi^2, \Phi^3, \Phi_p), \text{ где}$$

$$\Phi^1 = \left(N_1^0, N_2^0, \dots, N_{n_1}^0, TN_1, HN_1, N_2^1, N_2^1, \dots, N_{n_1}^1, TN_2, HN_2, \dots \right),$$

$$\left(\dots, N_{n_1}^{n_1-1}, TN_{n_1}, HN_{n_1}, N_F^{n_1}, N_F^0, TN_F, HN_F \right),$$

$$\Phi^2 = \left(T_1^0, T_2^0, \dots, T_{n_2}^0, TT_1, HT_1, T_2^1, T_2^1, \dots, T_{n_2}^1, TT_2, HT_2, \dots \right),$$

$$\left(\dots, T_{n_2}^{n_2-1}, TT_{n_2}, HT_{n_2}, T_F^{n_2}, T_F^0, TT_F, HT_F \right),$$

$$\Phi^3 = \left(\begin{array}{l} H_1^0, H_2^0, \dots, H_{n_3}^0, TH_1, HH_1, H_2^1, H_2^1, \dots \\ \dots, H_{n_3}^1, TH_2, HH_2, \dots, H_{n_3}^{n_3-1}, TH_{n_3}, HH_{n_3}, H_F^0, H_F^0, TH_F, HH_F \end{array} \right),$$

$\mathbf{n} = (n_1, n_2, n_3)$ где n_1, n_2, n_3 – число слоев каждого СП (I, II, III).

Сводя задачу построения полного СП (6) к задаче линейного параметрического программирования с векторным параметром \mathbf{n} , получим:

$$(7) \quad \Phi_P(\mathbf{n}) \rightarrow \max_{\mathbf{n}, \Phi} \sum_{j=1}^P \Xi_{ij} \Phi_j \leq \gamma_i, \quad i = 1, \dots, U(\mathbf{n}),$$

где $P = P(\mathbf{n})$ – число переменных; $U(\mathbf{n})$ – число неравенств, определяемых следующими формулами:

$$P(\mathbf{n}) = p(n_1) + p(n_2) + p(n_3) + 1, \quad U(\mathbf{n}) = u(n_1) + u(n_2) + u(n_3);$$

$$\gamma = \left(\underbrace{N, TN, HN, 0, \dots}_{u(n_1)}, \underbrace{T, TT, HT, 0, \dots}_{u(n_2)}, \underbrace{H, TH, HH, 0, \dots}_{u(n_3)} \right)^T,$$

$$\Xi = \begin{pmatrix} M_1 & 0 & 0 & F_1 \\ 0 & M_2 & 0 & F_2 \\ 0 & 0 & M_3 & F_3 \end{pmatrix},$$

(матрицы M_1, M_2, M_3 строятся аналогично задаче (4)),

$$F_1 = \left(0, 0, 0, \frac{A}{S}, \frac{B}{S}, \frac{C}{S}, 0, \dots, 0 \right), \quad F_2 = \left(0, 0, 0, \frac{B}{S}, \frac{B}{S}, \frac{C}{S}, 0, \dots, 0 \right),$$

$$F_3 = \left(0, 0, 0, \frac{C}{S}, \frac{B}{S}, \frac{C}{S}, 0, \dots, 0 \right), \quad \dim F_i = u(n_i), \quad i = 1, 2, 3, \quad S = A + B + C.$$

Для исследования зависимости решения задачи (7) от векторного параметра был разработан алгоритм автоматического вывода неравенств (7) для различных \mathbf{n} .

Обозначим Φ_P^* – решение задачи (7), $Q = \underset{\mathbf{n}}{\text{Arg max}} \Phi_P^*(\mathbf{n})$.

Множеством \mathbf{n}^* оптимальных структур полного СП для задачи (7) будем считать множество Парето \mathbf{n}_P многокритериальной задачи оптимизации:

$$n_i \rightarrow \min_{\mathbf{n} \in Q^*}, \quad i = 1, 2, 3.$$

4. Результаты численных экспериментов

Внешними параметрами задачи (7) являются пропорции входных переменных $N : T : H$ и технологических коэффициентов вспомогательных ПП. Далее представлены результаты решения двух частных оптимизационных задач (7):

7(а): задача (7) при фиксированных входах.

7(б): задача (7) при $n = n_1 = n_2 = n_3$.

Некоторые результаты решения параметрической задачи 7(а) симплекс-методом [11] представлены на рис. 6.

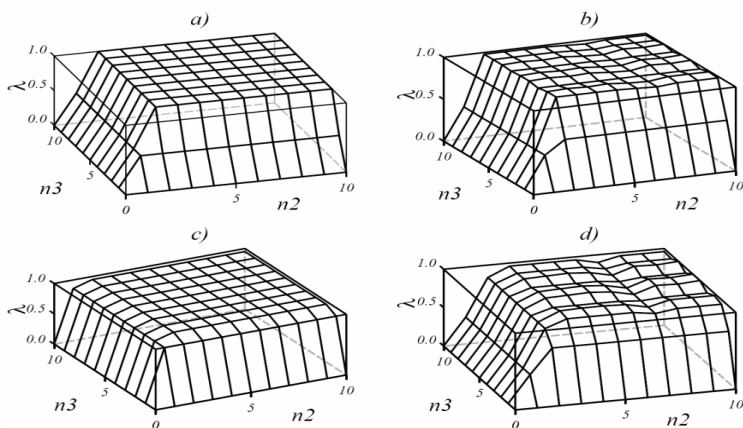


Рис. 6. График зависимости $\lambda(n_1, n_2, n_3)$ при $n_1 = 10$, $N_b = T_b = H_b = 1, TN = HN = TT = HT = TH = HH = 10$
 а) $K = \{1\}$, $\mathbf{n}_n = (10, 1, 1)$; б) $K = [1, 2]$, $\mathbf{n}_n = (10, 8, 4)$;
 в) $a_i = 1, b_i = 2, c_i = 3$, $\mathbf{n}_n = (10, 2, 1)$, д) $K = [1, 5]$, $\mathbf{n}_n = (10, 10, 4)$

На рис. 7, 8 представлены результаты решения задач 7(а, б) для некоторых значений параметров.

На рис. 10–12 приведены некоторые решения задачи 7(б): графики изменения λ при $T_b \in [0, 1]$, $H_b \in [0, 1]$ с шагом $\Delta = 0,1$ и значениях параметра $n = n_1 = n_2 = n_3$ от 1 до 8.

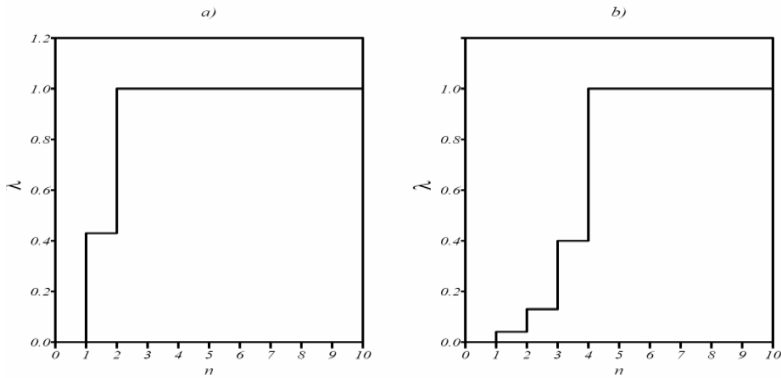


Рис. 7. График зависимости $\lambda(n)$

a) $N_b = T_b = H_b = 1, TN = HN = TT = HT = TH = HH = 10,$
 $K = \{1\}, n^* = 2;$ b) $N_b = T_b = H_b = 1,$
 $TN = HN = TT = HT = TH = HH = 100, K = \{1\}; n^* = 4$

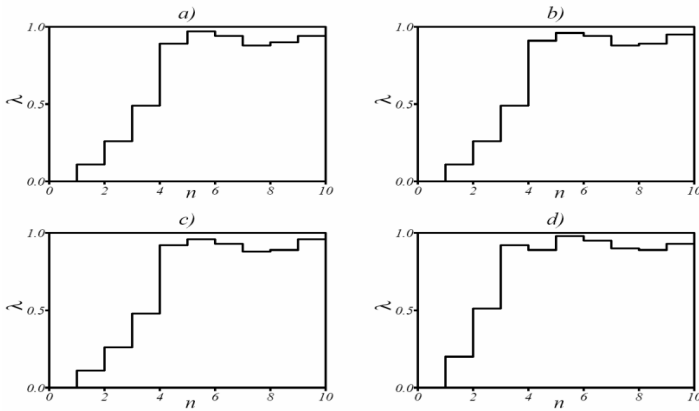


Рис. 8. График зависимости $\lambda(n), n^* = 5;$

при $N_b = 1, 0$ и $TN = HN = TT = HT = TH = HH = 10, K = [1, 5].$

a) $T_b = 0,3; H_b = 0,3;$ b) $T_b = 0,3; H_b = 0,6;$
 c) $T_b = 0,3; H_b = 0,9;$ d) $T_b = 0,6; H_b = 0,3$

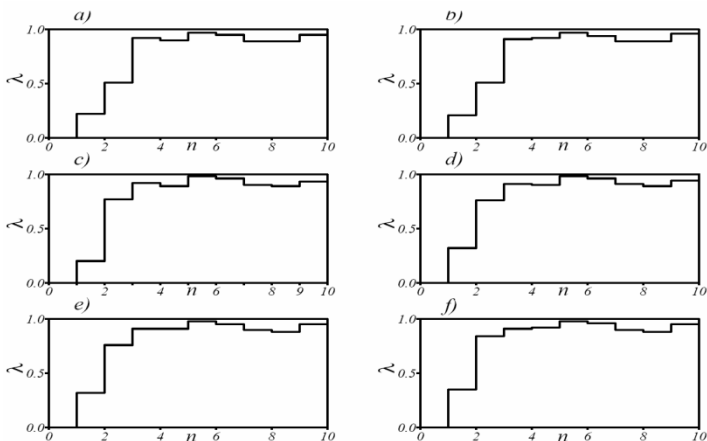


Рис. 9: График зависимости $\lambda(n)$. $n^* = 5$; при $N_b = 1, 0$
 $uTN = HN = TT = HT = TH = HH = 10, K = [1, 5]$.

- a) $T_b = 0,6; H_b = 0,6$; b) $T_b = 0,6; H_b = 0,9$; c) $T_b = 0,9; H_b = 0,3$;
 d) $T_b = 0,9; H_b = 0,6$; e) $T_b = 0,9; H_b = 0,9$; f) $T_b = 1,0; H_b = 1,0$

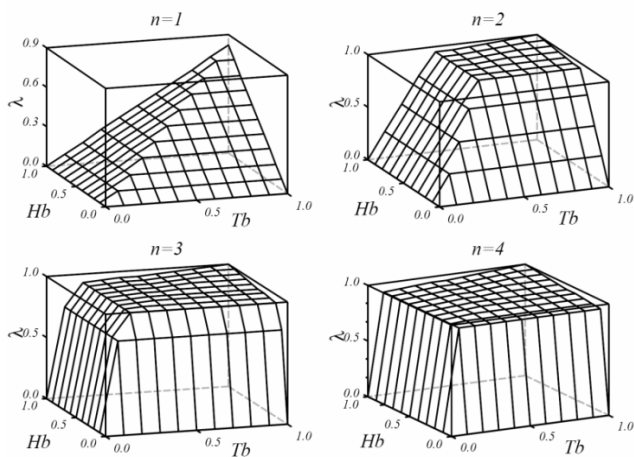


Рис. 10. График $F(T, H)$ при
 $TN = HN = TT = HT = TH = HH = 5, 0$ при $K = \{1\}$

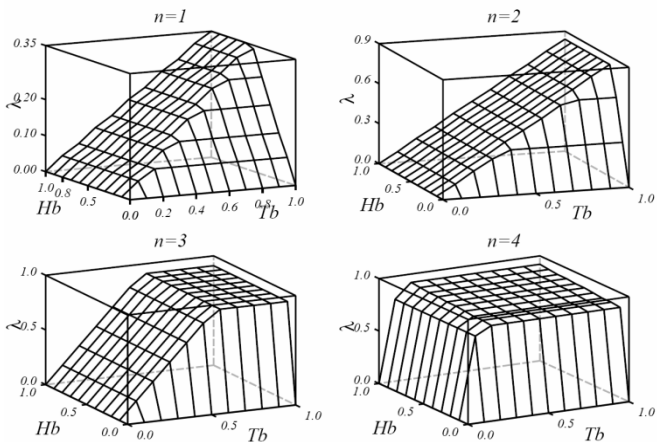


Рис. 11. График $F(1, T, H)$ при $TN = HN = TT = HT = TH = HH = 10,0$ и $\text{при } K = [1, 5]$.

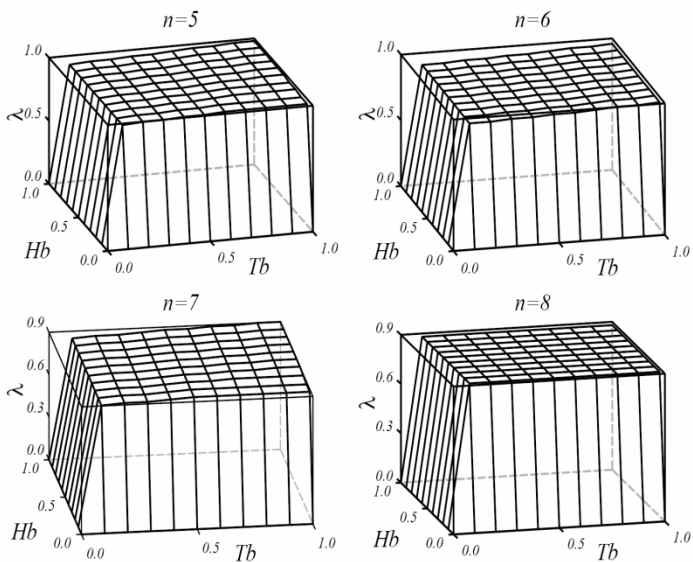


Рис. 12. График $F(1, T, H)$ при $TN = HN = TT = HT = TH = HH = 10,0$ и $\text{при } K = [1, 5]$

5. Заключение

Результаты численного исследования модели оптимизации структуры операционного ядра ОС показали, что

– значительные запасы факторов во вспомогательных производствах и достаточно большое число ПП во всех линиях полного СП позволяет достичь устойчивого максимума ПФ при значительных отклонениях пропорций факторов базового производства от оптимальных и разбросах технологических коэффициентах вспомогательных производств.

– увеличение разброса технологических коэффициентов вспомогательных производств при больших запасах факторов компенсируется ростом числа слоев ПП.

– уменьшение запасов факторов и числа ПП вспомогательных производств резко снижает устойчивость выпуска при отклонении пропорций факторов базового производства от идеальных.

Введение в данную модель параметров, характеризующих внутреннюю и внешнюю среду ОС, позволяет строить на ее основе модели оптимизации комплексной структуры ОС и механизмов управления. Второе направление непосредственного развития модели – модели структурной динамики и оптимального управления. Синтезом двух указанных направлений являются модели управления адаптацией и развитием ОС.

Литература

1. ВОРОНИН А.А., МИШИН С.П. *Алгоритмы поиска оптимальной структуры организационной системы* // Автоматика и телемеханика. – 2002. – №5. – С. 120–132.
2. ВОРОНИН А.А., МИШИН С.П. *Оптимальные иерархические структуры*. – М. ИПУ РАН, 2003. – 214 с.
3. ВОРОНИН А.А., МИШИН С.П. *Модель оптимального управления структурными изменениями организационной системы* // Автоматика и телемеханика. – 2002. – №8. – С. 136–150.
4. ВОРОНИНА И.Д. *Задача управления организационной структурой в условиях глобального инновационного процес-*

- са // Управление большими системами. –2006. – №12–13. – С. 51–59.
5. ГУБКО М.В. *Математические модели оптимизации иерархических структур*. – М.: ЛЕНАНД, 2006. – 264 с.
 6. ГУБКО М.В. *Структура оптимальной организации континуума исполнителей* // Автоматика и телемеханика. – 2002. – №12. –С. 116–130.
 7. КЛЕЙНЕР Г.Б. *Производственные функции. Теория, методы, применение*. – М.: Финансы и статистика, 1986. – 239 с.
 8. МИНЦБЕРГ Г. *Структура в кулаке. Создание эффективной организации* / Пер. с англ. под ред. Ю.Н. Каптуревско-го. – СПб.: Питер, 2002. – 512 с.
 9. МИШИН С.П. *Оптимальные иерархии управления в экономических системах*. – М.: ПМСОФТ, 2004. – 205 с.
 10. НОВИКОВ Д.А. *Теория управления организационными системами*. – М.: Московский психолого-социальный институт, 2005. – 584 с.
 11. ПАПАДИМИТРИУ Х., СТАЙГЛИЦ К. *Комбинаторная оптимизация. Алгоритмы и сложность*. Пер. сангл. – М.: Мир, 1985. – 512 с.

CONSTRAINED OPTIMIZATION MODEL OF ORGANIZATION'S OPERATING CORE

Alexander Voronin, Volgograd State University, Volgograd, Doctor of Science, professor (a.voronin@volsu.ru).

Mikhail Kharitonov, Volgograd State University, Volgograd, student (kharitonov.mihail@gmail.com).

Abstract: We suggest a constrained optimization model of an organization's operating core, whose structure consists of a basic technological module and modules of support facilities. The production function of the operating core is represented as a superposition of Leontief production functions corresponding to each of the modules. We reduce the optimization problem to a linear program with a parameter describing operational structure of the core. We develop an algorithm of automatic construction of basic equations for each value of the parameter. We solve the problem numerically for a wide range of model variables and parameters. The proposed model may serve as the basis for generalized mechanisms synthesis of organizational system control on the long time interval.

Keywords: organizational system, operating core, structure optimization, production function, linear program.

Статья представлена к публикации членом редакционной коллегии В. В. Клочковым

УДК 338.2:504.03

ББК 20.1+65.050

НАЛОГОВЫЕ МЕХАНИЗМЫ СТИМУЛИРОВАНИЯ ПОВЫШЕНИЯ ЭКОЛОГИЧНОСТИ ОБОРУДОВАНИЯ ДЛИТЕЛЬНОГО ПОЛЬЗОВАНИЯ¹

Клочков В. В.²

*(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)*

Гривский С. А.³, Игнатъева А. И.⁴

*(Московский физико-технический институт
(Национальный исследовательский университет), Москва)*

С помощью экономико-математических моделей проведен анализ экономической заинтересованности владельцев долговечного оборудования в его ускоренной замене на более экологически чистое. Выявлены условия, при которых государству придется стимулировать приобретение более экологически чистого оборудования, даже если оно более экономично. Проведен сравнительный анализ эффективности дискреционных инструментов такого стимулирования (например, налогообложения изделий старого поколения) и встроенных регуляторов (например, налогообложения вредных выбросов).

Ключевые слова: изделия длительного пользования, экология, экономическая эффективность, гипотеза Портера, стимулирование, коррупционные риски

¹ Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 11-08-00986).

² Владислав Валерьевич Клочков, доктор экономических наук, кандидат технических наук, ведущий научный сотрудник (vlad_klochkov@mail.ru).

³ Сергей Александрович Гривский, аспирант (serg.mipt@gmail.com).

⁴ Анастасия Ивановна Игнатъева, аспирант.

1. Введение

Повышение экологической чистоты производственных технологий, как правило, требует замены или модернизации долговечного и дорогостоящего оборудования. Причем нередко такая замена не является добровольной. Она может быть вызвана ужесточением экологических норм и запретом на дальнейшую эксплуатацию изделий, не удовлетворяющих новому уровню требований. Однако более экологически чистая техника, как правило, и экономически эффективнее, а ущерб окружающей среде сопряжен и с экономическими потерями вследствие перерасхода дефицитных ресурсов. Так, например, высокий расход топлива может быть связан с его неполным, неэффективным сгоранием, что, в свою очередь, порождает значительные выбросы сажи и др. вредных веществ. Таким образом, устраняя источники экологического ущерба, можно одновременно устранить и источники экономических потерь. Нередко улучшение экологических параметров тепловых двигателей (по крайней мере в части выбросов CO_2) происходит одновременно с повышением топливной экономичности, поскольку выбросы CO_2 пропорциональны расходу топлива¹. В связи с этим возникает следующая гипотеза: частные агенты сами, без государственного принуждения, заинтересованы в замене старой техники на более экологически чистую. Следовательно, государственное регулирование в области экологии нецелесообразно: свободный рынок автоматически обеспечит повышение экологической чистоты техники. Наиболее известный сторонник такой точки зрения – известный

¹ При неизменном индексе эмиссии, т.е. отношении объема выбросов к объему потребляемых энергоресурсов. Если в ряде отраслей изделия новых поколений могут обеспечивать более низкие значения индекса эмиссии (например, в стационарной энергетике – за счет улавливания CO_2 , подробнее см. [13]), то на транспорте подобные решения вряд ли реализуемы. Например, CO_2 содержится в реактивной струе авиадвигателя, которая и обеспечивает движение самолета.

экономист неоклассического направления М. Портер, поэтому данная гипотеза часто называется *гипотезой Портера*. Примечательно, что, будучи приверженцем либеральных подходов в экономической политике, он не отрицает полностью необходимости принятия законов об охране окружающей среды, однако полагает, что они необходимы лишь в силу ограниченной рациональности предпринимателей. В отсутствие подобных законов они могут не осознавать, что замена техники на более экологически чистую экономически выгодна им самим. Уточненный вариант гипотезы Портера таков: государственное принуждение может явиться начальным импульсом, побуждающим предпринимателей к инновациям. И даже если в краткосрочной перспективе экологическое регулирование приведет к снижению эффективности работы фирм, налагая на них дополнительную нагрузку, то в долгосрочной перспективе инновации, индуцированные этим принуждением, приведут к росту конкурентоспособности. Сами эти тезисы были высказаны в работах [18] и [19], вышедших соответственно в 1991 и 1995 гг. С тех пор в зарубежной литературе появился обширный массив исследований, посвященных как теоретической, так и эмпирической проверке данной гипотезы (см., например, [20] и обзорную статью [16]). Отличие предлагаемого здесь подхода от большинства предшествующих заключается в управленческой постановке проблемы и в непосредственном учете технико-экономических факторов. Из числа работ, известных авторам, относительно близкой представляется лишь работа [21], в которой также рассматривается проблема замены долговечной техники на более новую и экологически чистую. Однако во всех зарубежных работах, посвященных анализу гипотезы Портера, внимание уделяется лишь тому, действительно ли повышение экологической чистоты технологий (под влиянием экологической политики) приведет – хотя бы в долгосрочной перспективе – к повышению их экономической эффективности. На наш взгляд, проверка такой гипотезы «в узком смысле» -

задача не столько экономического, сколько инженерного анализа¹. Здесь же вопрос ставится иначе: допустим, что экологически чистая техника действительно экономичнее. Но достаточны ли рыночные стимулы для того, чтобы побудить владельцев к замене техники на более экологически чистую (т.е. справедлива ли гипотеза Портера «в широком смысле»)? Причем, поскольку в данной работе, в отличие от практически всех предшествующих, рассматривается долговечное оборудование, которое в момент появления более экологически чистых технологий может обладать значительным остатком ресурса, основное внимание будет уделено именно ускоренной замене техники, до полной выработки ее ресурса. Для того чтобы такая замена была экономически выгодной, новые изделия, как показано в работах [5, 8], должны обладать существенно большей экономической эффективностью, чем современные. То есть даже в тех случаях, когда гипотеза Портера выполняется «в узком смысле», она необязательно будет справедлива «в широком смысле». Но именно последнее определяет необходимость или необязательность государственного вмешательства в процессы обновления долговечной техники. И если государственное вмешательство все-таки потребуется, в каких формах его предпочтительнее осуществлять?

В этой работе основным объектом приложения разработанного инструментария и полученных с его помощью рекомендаций является гражданская авиация. И хотя данная отрасль ответственна всего лишь за 13% суммарного объема выбросов CO₂ всеми видами транспорта и лишь около 2% общего объема антропогенных выбросов CO₂, в ней уже несколько десятилетий уделяется значительное внимание

¹ Разумеется, изменения могут заключаться не только во внедрении новой техники, но и в оптимизации организации бизнеса, и экологические ограничения могут стимулировать фирмы к поиску таких внутренних резервов.

повышению экологической чистоты, а экологические характеристики являются важнейшим фактором в конкурентной борьбе на рынках авиатехники (см. [3]). Ужесточение экологических норм становится инструментом устранения конкурентов с важнейших рынков и стимулирования продаж авиатехники в периоды стагнации на рынках авиаперевозок¹. В то же время предлагаемые в данной работе методические подходы и полученные в итоге качественные выводы применимы без ограничения общности во всех отраслях, в которых актуально повышение экологической чистоты долговечной техники.

2. Анализ экономической заинтересованности владельцев долговечной техники в повышении ее экологической чистоты

Вначале необходимо получить условия, определяющие экономическую заинтересованность владельца в досрочном списании старой техники и ее замене изделиями нового поколения. Для этого проведем сравнение двух альтернатив: продолжение эксплуатации старого изделия до полной выработки его ресурса либо его немедленная замена изделием нового типа. Целевой функцией будем считать затраты владельца. Это оправданно, если старое и новое изделия выполняют сопоставимую работу или можно привести затраты к единице продукции: например, к летному часу использования воздушного судна или к пассажиро-километру. Как показано в работах [5, 8], досрочное списание авиатехники становится целесообразным с точки зрения снижения эксплуатационных затрат, когда выполняется следующее неравенство:

$$(1) \quad a^{\text{нов}} = \frac{P^{\text{нов}}}{T^{\text{нов}}} < c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}}$$

¹ Эти аспекты подробнее исследованы в предшествующих работах авторов, см. [5, 6, 9].

где $c_{\text{опер}}^{\text{стар}}$, $c_{\text{опер}}^{\text{нов}}$ - текущие *операционные* затраты в расчете на летный час самолетов соответственно старого и нового типов; $a^{\text{нов}}$ - средняя ставка амортизации нового типа самолетов в расчете на летный час, определяемая как отношение цены самолета нового типа $P^{\text{нов}}$ к его назначенному ресурсу $T^{\text{нов}}$, выраженному в летных часах. Полученное условие интуитивно очевидно: стоимость приобретения нового изделия, приведенная к летному часу, должна быть ниже экономии текущих, операционных затрат. Тогда досрочная замена еще исправного изделия старого поколения будет выгодной. В свою очередь, операционные затраты складываются прежде всего из затрат на горюче-смазочные материалы (ГСМ), а также прочих составляющих: расходов на техническое обслуживание и ремонт (ТОиР), на оплату труда экипажей, платежей за услуги аэропортов и аэронавигационных служб и др., подробнее см. [10]:

$$(2) \quad c_{\text{опер}} = c_{\text{пр}} + g * p_{\text{ГСМ}},$$

где $c_{\text{пр}}$ - средние «нетопливные» (т.е. прочие, в свете данной работы) затраты в расчете на летный час; g - удельный расход топлива, тонн на летный час; $p_{\text{ГСМ}}$ - цена тонны авиатоплива. В то же время досрочная замена долговечных изделий, которые могли бы еще безопасно эксплуатироваться, сопряжена не только с дополнительными финансовыми затратами их владельцев, но и с дополнительным расходом различных природных ресурсов, энергии, с экологическим воздействием. По аналогии с условиями экономической эффективности досрочной замены техники можно получить условия эффективности с точки зрения экономии энергоресурсов и сокращения вредных выбросов.

Несмотря на существенное различие видов энергоносителей, используемых при производстве и эксплуатации авиатехники, энергозатраты на этих этапах ЖЦИ можно привести к сопоставимому виду, выразив их, например, в т.н.э. – тоннах нефтяного эквивалента. Если приближенно принять энергетическую ценность 1 т авиатоплива равной 1 т.н.э., тогда текущее энергопотребление в расчете на летный

час эксплуатации самолета, выраженное в т.н.э., равно среднему часовому расходу авиатоплива самолетом данного типа g . Суммарные затраты энергоносителей на производство изделия обозначим G . Они включают в себя энергозатраты на получение необходимых конструкционных материалов (выплавку металлов и сплавов, синтез пластмасс и т.п.), на их обработку, а также на содержание зданий и сооружений (отопление, освещение и т.п.).

Количественные показатели эмиссии вредных веществ тепловыми двигателями также весьма многообразны. Большое значение имеют выбросы окислов азота, несгоревших частиц сажи, а также одного из основных газов, вызывающих парниковый эффект, – углекислого газа (CO_2). Если учитывать только последний показатель, тогда эмиссия в процессе эксплуатации изделий (в расчете на летный час) связана с потреблением энергоносителей следующим образом:

$$(3) \quad x_{\text{экспл}} = g * e_{\text{экспл}}$$

где $e_{\text{экспл}}$ – индекс эмиссии CO_2 на этапе эксплуатации, определяющий объем выбросов CO_2 при сжигании тонны авиатоплива. Его можно принять равным 3,125 т CO_2 /т.н.э., см. [1]. Что касается объема эмиссии CO_2 в процессе производства изделий, обозначенного X , он имеет специфический характер. Непосредственно производство может почти не создавать вредных выбросов, однако с выбросами сопряжена выработка потребляемой машиностроительными предприятиями энергии. Как правило, это электроэнергия, вырабатываемая на электростанциях. Они, в свою очередь, могут работать на различных видах топлива (природный газ, уголь, нефтепродукты, и т.д.), каждому из которых соответствует свое значение индекса эмиссии либо могут вообще не сжигать углеводородного топлива (ГЭС, АЭС). Если обозначить средний индекс эмиссии на стадии производства $\bar{e}_{\text{произв}}$, можно выразить суммарную эмиссию в процессе производства одного изделия следующим образом:

$$(4) \quad X = G * \bar{e}_{\text{произв}}$$

Показатели расхода энергоносителей G и g , а также показатели эмиссии X и $x_{\text{экспл}}$ можно трактовать соответственно

как «энергетические цены» и «экологические цены» производства и эксплуатации изделий. Строго говоря, экологические и энергетические (ресурсные) цены можно трактовать единообразно. С одной стороны, изъятие природных ресурсов, как и производство выбросов, оказывает нагрузку на природу, т.е. и энергетические цены можно рассматривать как экологические. С другой стороны, в современной экономике природопользования производство антропогенных выбросов также трактуется как эксплуатация природных ресурсов (см. [12]), но в расширенном смысле. Под ресурсами подразумеваются и способности окружающей среды поглощать выбросы. Таким образом, и экологические цены можно назвать ресурсными. В связи с этим можно ввести обобщающие показатели экологических цен производства и эксплуатации техники, в которые войдут (например, с некоторыми весовыми коэффициентами) потребление ресурсов и производство выбросов. Такие интегральные показатели в любом случае придется вводить даже для измерения экологических цен, поскольку, например, при сжигании топлива образуются вредные выбросы многих видов, подробнее см. [1, 11]. Аналогично производство изделий сопряжено с расходом не только энергетических, но и других природных ресурсов.

Целесообразно ли с точки зрения энергосбережения или охраны окружающей среды досрочное списание изделий старого поколения и их замена на изделия нового типа? Как и при выводе условия (1), сравним две альтернативы: продолжение эксплуатации старого изделия до полной выработки остатка ресурса либо его немедленная замена на новое изделие. В итоге получим условия, подобные условию (1) (с точностью до вида цен: экономические, экологические или энергетические). Так, досрочное списание изделий старого поколения и приобретение изделий нового поколения будет эффективным с точки зрения энергосбережения только при выполнении следующего неравенства:

$$(5) \quad \frac{G^{\text{нов}}}{T^{\text{нов}}} < g^{\text{стар}} - g^{\text{нов}}.$$

То есть энергозатраты на производство нового изделия, приходящиеся на 1 летный час его ресурса, должны быть ниже, чем разница эксплуатационных энергозатрат старого и нового изделий, приходящихся на летный час. Иначе говоря, энергетическая цена производства нового изделия должна быть ниже разницы энергетических цен эксплуатации старого и нового изделий. Аналогичное соотношение можно получить и для выбросов CO_2 . Сокращение суммарного ущерба окружающей среде при немедленной замене старых изделий на новые достигается лишь в том случае, если выполняется следующее неравенство:

$$(6) \quad \frac{X^{\text{нов}}}{T^{\text{нов}}} < x_{\text{опер}}^{\text{стар}} - x_{\text{опер}}^{\text{нов}} .$$

Итак, досрочное списание изделий старого поколения по соображениям сокращения выбросов целесообразно при выполнении условия (6), а по соображениям экономии энергоресурсов – при выполнении условия (5). Всегда ли при этом владельцы будут экономически заинтересованы в досрочном списании старых изделий, т.е. будет ли выполняться условие (1)? По объективным причинам, соотношения экономических, энергетических и экологических цен производства и эксплуатации авиатехники существенно различаются:

$$\frac{P}{c_{\text{опер}}} \neq \frac{G}{g}, \quad \frac{G}{g} \neq \frac{X}{x_{\text{экспл}}}, \quad \frac{X}{x_{\text{экспл}}} \neq \frac{P}{c_{\text{опер}}} .$$

Необходимо учитывать, что цена изделия, помимо стоимости потребленных в процессе производства энергоресурсов, включает в себя прибыль производителя, а также целый ряд иных статей затрат: на разработку (*научно-исследовательские и опытно-конструкторские работы*, НИОКР), на оплату труда производственных рабочих, на закупку сырья и комплектующих изделий и т.д., а значит справедливо следующее:

$$(7) \quad P = p_{\text{энерг}} * G + P_{\text{проч}} ,$$

где $P_{\text{энерг}}$ – средняя цена энергоресурсов, потребляемых в процессе производства изделий; $P_{\text{проч}}$ – прочие составляющие цены изделия, не связанные непосредственно с потреблением энергоресурсов.

Пусть вместо неравенства (1) выполняется строгое равенство:

$$(8) \quad \frac{P^{\text{нов}}}{T^{\text{нов}}} = c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}},$$

т.е. с экономической точки зрения одинаково выгодно как продолжить эксплуатацию старого изделия, так и немедленно заменить его новым. Это граничная ситуация с точки зрения экономической заинтересованности владельцев техники в ее досрочной замене. Перепишем левую и правую части последнего равенства в следующей форме:

$$\frac{P^{\text{нов}}}{T^{\text{нов}}} = \frac{P_{\text{энерг}} * G^{\text{нов}} + P_{\text{проч}}^{\text{нов}}}{T^{\text{нов}}},$$

$$c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}} = (c_{\text{пр}}^{\text{стар}} - c_{\text{пр}}^{\text{нов}}) + (g^{\text{стар}} - g^{\text{нов}}) * p_{\text{гсм}},$$

где $c_{\text{пр}}^{\text{стар}}$, $c_{\text{пр}}^{\text{нов}}$ – средние «нетопливные» затраты на ТОиР в расчете на летный час самолетов соответственно старого и нового типов.

Доля стоимости потребленных энергоресурсов в цене сложной наукоемкой техники, как правило, невелика. Так, по данным источника [15], доля затрат на электроэнергию и топливо в общих издержках всех предприятий авиационного двигателестроения США составляет лишь 1-2%, т.е. $P_{\text{энерг}} * G < P_{\text{проч}}$. В то же время на стадии эксплуатации именно затраты на ГСМ становятся при нынешних ценах на авиатопливо одной из главных, если не преобладающей статьей издержек авиакомпаний. И даже относительно небольшая экономия топлива существенно сокращает операционные издержки. То есть выполняются следующие соотношения:

$$g * p_{\text{гсм}} < c_{\text{пр}},$$

$$(g^{\text{стар}} - g^{\text{нов}}) * p_{\text{гсм}} < c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}}.$$

Следовательно, при выполнении равенства (8), как правило, выполняется следующее неравенство:

$$P_{\text{энерг}} * \frac{G^{\text{нов}}}{T^{\text{нов}}} < P_{\text{гсм}} * (g^{\text{стар}} - g^{\text{нов}}),$$

или (считая, что $P_{\text{энерг}} < P_{\text{гсм}}$),

$$\frac{G^{\text{нов}}}{T^{\text{нов}}} < (g^{\text{стар}} - g^{\text{нов}}).$$

Это неравенство совпадает с условием (5) и означает, что досрочная замена техники эффективна с точки зрения энергосбережения. Заметим, что такая ситуация складывается при выполнении равенства (8), граничного, с точки зрения экономической эффективности такой замены. Таким образом, при выполнении неравенства (1) условие (5) тем более выполняется, но обратное справедливо далеко не всегда. То есть экономическая заинтересованность владельцев в досрочной замене старой техники не гарантирована, даже когда такая замена целесообразна с точки зрения энергосбережения.

В свою очередь, досрочное списание техники становится целесообразным с точки зрения снижения потребления энергоресурсов при условии (5), а при выполнении следующего равенства станет одинаково выгодно как продолжить эксплуатацию старого изделия, так и немедленно заменить его новым:

$$(9) \quad \frac{G^{\text{нов}}}{T^{\text{нов}}} = (g^{\text{стар}} - g^{\text{нов}}).$$

Эта ситуация является граничной с точки зрения энергетической эффективности досрочной замены техники. Как правило, среднее значение коэффициента удельной эмиссии на стадии производства ниже (по крайней мере не выше), чем на стадии эксплуатации ($\bar{e}_{\text{произв}} < e_{\text{экспл}}$). Тогда при выполнении условия (9) справедливо следующее неравенство:

$$\bar{e}_{\text{произв}} * \frac{G^{\text{нов}}}{T^{\text{нов}}} < e_{\text{экспл}} * (g^{\text{стар}} - g^{\text{нов}}).$$

Но, согласно ранее введенным обозначениям, $\bar{e}_{произв} * G = X$, а $e_{экспл} * g = x_{экспл}$. Поэтому последнее неравенство можно переписать в следующей форме:

$$\frac{X_{нов}}{T_{нов}} < x_{экспл}^{стар} - x_{экспл}^{нов},$$

т.е. оно совпадает с условием (6). Это означает, что досрочная замена техники оправданна с экологической точки зрения при выполнении равенства (9), граничного с точки зрения энергетической эффективности такой замены. Таким образом, при выполнении условия (5) условие (6), как правило, выполняется, но обратное справедливо далеко не всегда. То есть целесообразность ускоренной замены старых изделий с экологической точки зрения достигается в более широком диапазоне условий, чем с точки зрения энергосбережения.

Сопоставляя полученные результаты, можно сделать вывод о том, что экономическая заинтересованность эксплуатирующих организаций в ускоренной замене изделий на технику нового поколения будет достигаться (по мере сокращения расхода топлива изделиями новых типов) позже, чем такая замена станет оправданной как с экологической, так и с энергетической точки зрения. Поэтому, вопреки выводу, сделанному на основе гипотезы Портера противниками государственного вмешательства в экономику, иногда оно необходимо для обеспечения экономии энергоресурсов и сокращения техногенной нагрузки на окружающую среду.

Этот вывод усиливается, если принять во внимание следующие факторы. Всегда ли повышение экологической чистоты техники сопровождается повышением ее экономической эффективности? В качестве единственного вида воздействия тепловых двигателей на окружающую среду здесь рассматривались выбросы CO_2 , приблизительно пропорциональные расходу топлива, поэтому повышение экологической чистоты изделий сопровождалось и повышением их экономической эффективности. Однако наряду с выбросами парниковых газов имеет значение и эмиссия иных видов вредных веществ: угарного газа (CO), сажевых частиц (C),

окислов азота (NO_x) и т.п., а также шум, производимый летательными аппаратами и другими транспортными средствами, тепловое загрязнение и др.. Связь между уменьшением этих вредных воздействий и сокращением расхода топлива уже не является прямой и даже монотонной. Ряд специалистов в области проектирования авиадвигателей отмечает [4, 11], что до определенного момента соображения улучшения экологических характеристик и прочих показателей совершенства авиадвигателей не противоречили друг другу. Например, повышение полноты сгорания топлива, как говорилось выше, сокращает не только удельный расход топлива, но и выбросы несгоревших остатков топлива, дымность выхлопа. Как отмечено в работе [4], сам по себе переход к *двухконтурным турбореактивным авиадвигателям* и повышение *степени их двухконтурности*, целесообразные с точки зрения повышения удельной тяги и топливной экономичности, попутно привели и к радикальному (приблизительно на 20 дБ) сокращению уровня шума. Однако для дальнейшего снижения уровня шума авиадвигателей и эмиссии вредных веществ производителям приходится внедрять новые, все более сложные и дорогостоящие конструктивно-технологические решения, неоднозначные с экономической точки зрения. Нередко даже приходится оптимизировать конструкцию двигателя не по критериям повышения тяги или сокращения расхода топлива, а именно из соображений снижения уровня шума. Дальнейшее снижение уровня вредных выбросов (за исключением CO_2 , выбросы которого пропорциональны расходу топлива) также противоречит соображениям экономии топлива. Например, при повышении температуры перед турбиной повышается полнота сгорания топлива, но растет и образование окислов азота (NO_x) и т.п. С аналогичными противоречиями сталкиваются и разработчики автомобильных двигателей (весьма показательна динамика экономичности и экологических показателей двигателей, удовлетворяющих нормам *Euro 1 – Euro 6*). На данном этапе развития технологий практически исчерпаны резервы улучшения по Парето характеристик тепловых двигателей

различных типов, т.е. одновременного улучшения экономических и экологических параметров. Как видно из приведенных примеров, даже требования снижения разных видов экологического воздействия тепловых двигателей на природу могут вступать в противоречие друг с другом. Дальнейшее повышение экологической чистоты техники практически по любому критерию приводит, как правило, к повышению затрат – как на производство изделий, так и на их эксплуатацию. Поэтому, если с общественной точки зрения будет признано целесообразным дальнейшее улучшение экологических параметров тепловой энергетики, автомобильного и авиационного транспорта, вероятнее всего, для достижения этой цели потребуются государственное вмешательство, поскольку даже «в узком смысле» гипотеза Портера уже не выполняется, т.е. повышение экологичности техники ухудшает ее экономические показатели. В связи с этим попытки ее эмпирической проверки (и тем более формирование на ее основе рекомендаций в отношении экологической политики) без учета стадии инновационного развития технологий в той или иной отрасли принципиально некорректны.

Как показано выше, в настоящее время, как правило, вначале выполняются «экологическое» и «энергетическое» условия эффективности досрочной замены изделий длительного пользования, т.е. условия (5) и (6), и лишь затем – «экономическое», т.е. условие (1). В итоге приходится стимулировать досрочную замену техники, поскольку она целесообразна с экологической точки зрения¹. Но представим

¹ Здесь рассматривается именно сокращение удельного потребления ресурсов и удельной эмиссии вредных веществ, т.е. приходящихся на единицу продукции (например, на летный час или пассажиро-километр). Однако суммарное потребление ресурсов и суммарная эмиссия при этом (и благодаря этому) могут даже возрасти. В этом состоит так называемый эффект рикошета (см., например, [17]), который также необходимо учитывать при формировании экологической политики. Однако в данной работе он не рассматрива-

себе, что последовательность выполнения соответствующих условий по мере совершенствования технологий была бы обратной. Это означало бы, что экономические субъекты уже экономически заинтересованы в досрочной замене техники, но с экологической и ресурсной точек зрения она неэффективна. Такая гипотетическая ситуация отнюдь не является невероятной – напротив, она не менее распространена, чем та, что обсуждалась в этой работе до сих пор. В разных отраслях и на разных этапах технологического развития могут иметь место различные соотношения экономических, энергетических и экологических цен производства и эксплуатации изделий, а также их изменения по мере совершенствования технологий. И если, например, в гражданском авиастроении сокращение энергетических и экологических цен эксплуатации (т.е. расхода топлива, выбросов CO_2) даже на верхнем участке S-образной кривой существенно выше соответствующих цен производства новой техники, то в некоторых отраслях – таких как производство бытовой техники и электроники – напротив, энергетические и экологические цены производства сами по себе существенно выше соответствующих цен эксплуатации (которые, в принципе, могут и не снижаться существенно при появлении новых поколений продукции). То есть досрочная замена таких изделий будет заведомо неэффективной с энергетической или экологической точки зрения, но при этом вполне может быть выгодной с экономической точки зрения¹. И в соответствующих отраслях, напротив, необходимо принимать

ется, а рассматривается лишь процесс улучшения удельных параметров технологий.

¹ Кроме того, ускоренное обновление потребительских благ длительного пользования может быть вызвано не столько объективными экономическими факторами, сколько субъективными мотивами – изменчивостью моды, рекламой и пропагандой. Для любых производителей товаров длительного пользования их ускоренная замена – один из главных источников спроса, в том числе в периоды неблагоприятной экономической конъюнктуры, см., например, [2].

меры государственного регулирования, которые тормозили бы экологически неэффективные или даже опасные процессы ускоренного обновления продукции. Провести формальный анализ эффективности таких мер и выбрать наиболее предпочтительные можно по аналогии с анализом инструментов стимулирования ускоренной замены техники, который и проводится в данной работе.

Таким образом, идеальной была бы ситуация, когда пропорции различных видов цен (экономических, энергетических и экологических), а также их изменения в ходе научно-технического прогресса совпадали бы (хотя бы приблизительно). Однако, в силу объективных технико-экономических причин, такая пропорциональность вряд ли будет соблюдаться естественным образом во всех отраслях и на всех стадиях технологического развития. И корректировать возможные дисбалансы придется с помощью государственного вмешательства. Обобщая результаты анализа различных сочетаний экологических и экономических цен, можно сформулировать следующую основную задачу государственной экологической политики на рынках изделий длительного пользования. Необходимо обеспечить сбалансированность системы экологических и экономических цен таким образом, чтобы досрочная замена техники становилась эффективной приблизительно одновременно как по экономическим, так и по экологическим критериям – без значительного опережения или отставания.

3. Сравнительный анализ прямого налогообложения эксплуатации старой техники и налогообложения выбросов

В ряде отраслей – например, в гражданской авиации, на автомобильном транспорте – ускоренная замена техники на более экологически чистую обеспечивается ужесточением соответствующих стандартов, ограничивающих эксплуатацию изделий старого поколения. Иногда такое ужесточение формально не затрагивает уже эксплуатируемые изделия, а

касается только продажи новых, однако реальная практика применения экологических норм свидетельствует о том, что они нацелены именно на ускоренное обновление парка. Впрочем, и в этих случаях владельцы старых изделий могут избежать финансовых потерь, продавая старую технику в страны третьего мира, где приняты менее жесткие экологические нормы, подробнее см. [9]. Однако, помимо прямого административного принуждения, стимулирование государством ускоренного обновления парка долговечной техники может принимать более мягкие формы. Например, эксплуатация изделий старых поколений может облагаться экологическими налогами. Методический аппарат, предложенный авторами в работе [7], предполагает следующий подход к экономическому обоснованию ставок таких налогов¹. Они должны обеспечить экономическую заинтересованность эксплуатирующих организаций в досрочном списании старой техники, если такое списание целесообразно с экологической точки зрения. Используя этот факт, получим оценки минимально необходимого уровня ставок экологических налогов для различных форм налогообложения. Итак, необходимо при выполнении условия (6) добиться выполнения следующего неравенства:

$$(10) \left(c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}} \right)_{\text{нал}} > a^{\text{нов}},$$

где $\left(c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}} \right)_{\text{нал}}$ – разность операционных затрат старого и нового изделий с учетом экологических налогов².

¹ Здесь рассматривается лишь непосредственное влияние экологического налогообложения на поведение эксплуатирующих организаций. Проблема надлежащего использования собираемых налогов остается за рамками данной работы.

² В принципе, добиться выполнения условия (1) можно и воздействуя на цену нового изделия, но лишь в сторону ее уменьшения, т.е. дотируя покупку более экологически чистой техники. Такой механизм также используется в ряде отраслей в наиболее экономически развитых странах мира.

Для достижения желаемого соотношения затрат владельцев старой и новой техники, государством могут применяться различные формы налогообложения:

а) экологические налоги могут взиматься непосредственно с каждой единицы использования (например, с летного часа) изделий старого поколения¹ по ставке t , исчисляемой в денежных единицах за летный час:

$$(11) \left(c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}} \right)_{\text{нал}} = e + c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}} ;$$

б) экологическими налогами может облагаться потребляемое [как старыми, так и новыми изделиями] авиатопливо по ставке s , исчисляемой в денежных единицах за тонну топлива:

$$(12) \left(c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}} \right)_{\text{нал}} = \left(c_{\text{пр}}^{\text{стар}} - c_{\text{пр}}^{\text{нов}} \right) + \left(g^{\text{стар}} - g^{\text{нов}} \right) * (p_{\text{ГСМ}} + s) ;$$

в) налогами могут облагаться сами выбросы CO₂ по ставке λ денежных единиц за тонну выбросов:

(13)

$$\left(c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}} \right)_{\text{нал}} = \left(c_{\text{пр}}^{\text{стар}} - c_{\text{пр}}^{\text{нов}} \right) + \left(g^{\text{стар}} - g^{\text{нов}} \right) * (p_{\text{ГСМ}} + e_{\text{экспл}} * \lambda) .$$

Последнюю форму экологического налогообложения предполагается применять во многих отраслях. Фактически этот принцип и положен в основу так называемого *Киотского протокола* [14] – одного из самых масштабных проектов экологического регулирования в мировом масштабе. Поэтому особенно актуален экономический анализ эффективности

¹ Либо наоборот, может дотироваться эксплуатация изделий нового поколения. Однако, несмотря на эквивалентность такого механизма дотированию их приобретения (т.е. уменьшению цены новых изделий) с арифметической точки зрения, именно второй вариант более реалистичен. Принимая решение о выборе изделий длительного пользования, владелец предпочтет единовременную дотацию в виде скидки к цене, чем периодические дотации за их эксплуатацию, которые могут и прекратиться при изменении государственной политики.

данного механизма, стимулирующего повышение экологической чистоты технологий. Сравнение формул (12) и (13) показывает, что принцип действия двух последних форм налогообложения одинаков (при фиксированном индексе эмиссии, делающем выбросы пропорциональными потреблению топлива).

Пользуясь формулами (11)–(13), можно получить оценки минимального уровня ставок экологических налогов, соответствующих описанным формам налогообложения, при которых они будут действенными (т.е. будет выполняться условие (10)):

(14)

$$t_{min} = a^{нов} - (c_{опер}^{стар} - c_{опер}^{нов}) = a^{нов} - (c_{пр}^{стар} - c_{пр}^{нов}) - (g^{стар} - g^{нов}) * P_{гсм},$$

(15)
$$s_{min} = \frac{1}{g^{стар} - g^{нов}} * \left[a^{нов} - (c_{пр}^{стар} - c_{пр}^{нов}) \right] - P_{гсм},$$

(16)

$$\lambda_{min} = \frac{1}{e_{экспл}} * \left\{ \frac{1}{g^{стар} - g^{нов}} * \left[a^{нов} - (c_{пр}^{стар} - c_{пр}^{нов}) \right] - P_{гсм} \right\} = \frac{s_{min}}{e_{экспл}}.$$

Рассмотрим следующий реалистичный пример. Пусть самолеты нынешнего поколения, имеющие значительный остаток ресурса, потребляют в среднем 2,5 т авиатоплива на летный час. На рынке появляются воздушные суда того же класса, но нового поколения, со следующими характеристиками: $T^{нов} = 75\ 000$ л.ч., $P^{нов} = 60$ млн. долл.,

$$g^{нов} = 0,5 \frac{\text{Т}}{\text{л.ч.}}$$

(заметим, что такое сокращение удельного

расхода топлива на данном этапе считается очень значительным). Для упрощения расчетов предположим, что прочие, «нетопливные» составляющие эксплуатационных затрат изделий старого и нового поколений различаются незначительно. Тогда различие операционных затрат изделий нынешнего и нового поколения вызвано лишь различными значениями расхода топлива. Условие (8) выполняется при цене авиатоплива, равной 1600 долл./т:

$$\frac{P^{\text{нов}}}{T^{\text{нов}}} = c_{\text{опер}}^{\text{стар}} - c_{\text{опер}}^{\text{нов}} = (g^{\text{стар}} - g^{\text{нов}}) * \tilde{p}_{\text{гсм}},$$

$$\tilde{p}_{\text{гсм}} = \frac{P^{\text{нов}}}{(g^{\text{стар}} - g^{\text{нов}}) * T^{\text{нов}}} = \frac{60 \text{ млн. долл.}}{0,5 \frac{\text{т}}{\text{л.ч.}} * 75000 \text{ л.ч.}} = 1600 \frac{\text{долл.}}{\text{т}}.$$

Если в настоящее время цена авиатоплива ниже этого уровня, авиакомпании будут заинтересованы продолжать эксплуатацию имеющихся воздушных судов до полной выработки ресурса. В то же время по экологическим соображениям может быть целесообразным их немедленное списание и замена на изделия нового поколения. Даже если принять пессимистическую оценку энергозатрат на производство самолета нового поколения: $G^{\text{нов}} = 12\,000 \text{ т.н.э.}$ ¹, и считать, что индексы эмиссии CO_2 на этапах производства и эксплуатации совпадают: $\bar{e}_{\text{произв}} = e_{\text{экспл}} = 3,125 \frac{\text{т. CO}_2}{\text{т.н.э.}}$, условие (б) выполняется более чем с трехкратным запасом:

$$\bar{e}_{\text{произв}} * \frac{G^{\text{нов}}}{T^{\text{нов}}} = 3,125 \frac{\text{т. CO}_2}{\text{т.н.э.}} * \frac{12000 \text{ т.н.э.}}{75000 \text{ л.ч.}} = 0,5 \frac{\text{т. CO}_2}{\text{л.ч.}} <$$

$$< e_{\text{экспл}} * (g^{\text{стар}} - g^{\text{нов}}) = 3,125 \frac{\text{т. CO}_2}{\text{т.н.э.}} * 0,5 \frac{\text{т}}{\text{л.ч.}} = 1,5625 \frac{\text{т. CO}_2}{\text{л.ч.}}$$

В данном примере ускоренная замена старого поколения авиатехники на новое позволит сократить выбросы CO_2 более чем на 1 т в расчете на летный час, даже с учетом дополнительных выбросов в процессе производства новых изделий. Таким образом, в данном случае государство, руководствуясь соображениями защиты окружающей среды, может прибегнуть к прямому запрету эксплуатации изделий старого поколения либо ввести экологические налоги.

¹ Некоторые простые оценки энергоемкости производства авиатехники, как из металлов и сплавов, так и из полимерно-композитных материалов, выполнены авторами в работе [7].

Какая форма экологического налогообложения, из числа описанных выше, наиболее предпочтительна? Экономисты либерального направления рекомендуют использовать в государственной экономической политике *встроенные регуляторы*, которые, задавая общие для всех экономических агентов правила игры, автоматически обеспечивают желаемые изменения. Этому принципу лучше соответствует налогообложение потребляемого авиатоплива или выбросов CO₂. В то же время непосредственное налогообложение эксплуатации изделий старого поколения, как и прямой запрет на их эксплуатацию, относятся к *дискреционным* мерам. Если при использовании встроенных регуляторов всех экономических агентов помещают в однородную экономическую среду, стимулирующую желательные изменения, то дискреционные меры избирательны и направлены на конкретных агентов. Высокая селективность, присущая таким мерам, порождает коррупционные риски.

Однако прежде чем учитывать коррупционные аспекты, следует проанализировать принципиальную реализуемость и действенность той или иной политики. Пользуясь формулами (14)-(16), оценим минимальный уровень ставок экологических налогов, при котором авиакомпания в приведенном примере примут решение о досрочном обновлении парка авиатехники, если текущая цена авиатоплива равна 1000 долл./т.

Если налогом непосредственно облагается эксплуатация старых изделий, минимально необходимая ставка за летный час составит (см. формулу (14))

$$t_{min} = a^{нов} - (g^{стар} - g^{нов}) * p_{гсм} = \frac{60 \text{ млн.долл.}}{75000 \text{ л.ч.}} - 0,5 \frac{т}{\text{л.ч.}} * 1000 \frac{\text{долл}}{т} = 300 \frac{\text{долл}}{\text{л.ч.}}$$

Если налогом облагается потребляемое авиатопливо, ставка налога фактически должна повысить его цену до уровня, при котором авиакомпания добровольно откажутся от эксплуатации авиатехники старого поколения, т.е. до $\tilde{p}_{гсм}$ (см. формулу (15)):

$$s_{min} = \tilde{p}_{гсм} - p_{гсм} = 1600 \frac{\text{долл}}{т} - 1000 \frac{\text{долл}}{т} = 600 \frac{\text{долл}}{т}.$$

Если налогами облагаются сами выбросы CO₂, минимально необходимая ставка составит (согласно формуле (16))

$$\lambda_{min} = \frac{s_{min}}{e_{экспл}} = \frac{600 \frac{\text{долл}}{т}}{3,125 \frac{т \cdot CO_2}{т.н.э.}} = 192 \frac{\text{долл}}{т \cdot CO_2}.$$

Заметим, что в случае применения двух последних форм государственного регулирования возрастут затраты на эксплуатацию не только старых, но и новых изделий. Величину этого прироста можно оценить, подставив формулы минимально необходимых ставок экологических налогов (15), (16) в выражения для эксплуатационных затрат. Если налог взимается с тонны авиатоплива, стоимость летного часа самолетов нового поколения возрастет по меньшей мере на

$$g^{\text{нов}} * s_{min} = 2 \frac{т}{л.ч.} * 600 \frac{\text{долл}}{т} = 1200 \frac{\text{долл}}{л.ч.}$$

При налогообложении выбросов CO₂ минимальный (т.е. достигаемый при минимально действенных ставках налогов¹) прирост эксплуатационных затрат для самолетов нового поколения будет таким же в силу эквивалентности этих механизмов:

$$g^{\text{нов}} * e_{экспл} * \lambda_{min} = 2 \frac{т}{л.ч.} * 3,125 \frac{т \cdot CO_2}{т.н.э.} * 192 \frac{\text{долл}}{т \cdot CO_2} = 1200 \frac{\text{долл}}{л.ч.}.$$

То есть последствия введения таких налогов эквивалентны последствиям удорожания авиатоплива на 60%. Такое повышение эксплуатационных затрат (полученное в рамках вполне реалистичного примера) является чрезвычайно значимым (если

¹ Меньшие ставки налогов не приведут к желаемому результату, т.е. не обеспечат экономической заинтересованности авиакомпаний в ускоренной замене авиатехники на более экологически чистую.

не катастрофичным) для авиакомпаний. Возможно, это лишь частный случай, характеризующийся чрезвычайно неблагоприятным набором параметров? Оценим в общем виде прирост издержек для изделий нового поколения в том случае, если налог взимается с тонны выбросов. Прежде всего, сравнивая формулы (14) и (16), заметим, что минимально необходимая ставка штрафа за выбросы может быть аналитически выражена через минимально необходимую ставку прямого налога на эксплуатацию старых изделий:

$$(17) \quad \lambda_{min} = \frac{a^{нов} - (c_{пр}^{стар} - c_{пр}^{нов}) - p_{гсм} * (g^{стар} - g^{нов})}{e_{экспл} * (g^{стар} - g^{нов})} = \frac{a^{нов} - (c_{опер}^{стар} - c_{опер}^{нов})}{e_{экспл} * (g^{стар} - g^{нов})} = \frac{t_{min}}{e_{экспл} * (g^{стар} - g^{нов})}$$

Теперь остается оценить прирост издержек для изделий нового поколения в том случае, если взимается налог с выбросов (по минимально необходимой ставке, определяемой формулой (16)):

(18)

$$\Delta c_{\lambda}^{нов} = g^{нов} * e_{экспл} * \lambda_{min} = \frac{g^{нов} * e_{экспл} * t_{min}}{e_{экспл} * (g^{стар} - g^{нов})} = \frac{g^{нов}}{(g^{стар} - g^{нов})} * t_{min}$$

Как следует из полученного соотношения, прирост затрат владельцев новой техники был бы ниже, чем прямой налог на владельцев старых изделий, только в том случае, если бы показатели эмиссии при переходе на новое поколение авиатехники сокращались по меньшей мере вдвое:

$$\Delta c_{\lambda}^{нов} < t_{min}, \text{ при } g^{нов} < g^{стар} - g^{нов}, \text{ т.е. } g^{нов} < \frac{g^{стар}}{2}.$$

Тогда и доля экологических налогов в структуре эксплуатационных затрат становилась бы при покупке изделий

нового поколения пренебрежимо малой. Однако это возможно лишь при замене чрезвычайно «грязных» в экологическом отношении изделий на существенно более экологически чистые. Кроме того, при многократном сокращении выбросов CO_2 , вероятнее всего, и расход топлива сокращался бы в несколько раз, что автоматически обеспечивало бы экономическую заинтересованность авиакомпаний в досрочном списании старых изделий. Тогда и необходимость в государственном вмешательстве отпала бы. Однако к настоящему моменту соответствующий этап развития технологий в большинстве отраслей транспорта и в тепловой энергетике уже пройден. Как в отношении повышения экономичности, так и в отношении улучшения экологических характеристик тепловых двигателей возможности совершенствования традиционных конструкций и технологий близки к исчерпанию. Поэтому, напротив, имеют место следующие соотношения:

$$g^{\text{нов}} \approx g^{\text{стар}}, \text{ т.е. } g^{\text{нов}} > g^{\text{стар}} - g^{\text{нов}} \text{ и } \Delta c_{\lambda}^{\text{нов}} > t_{\text{min}}.$$

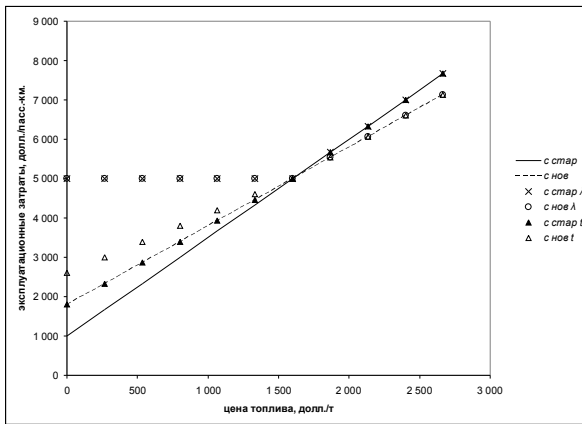


Рис. 1. Сравнение эксплуатационных затрат для старых и новых изделий при различных способах стимулирования обновления парка (пример 1)

Подчеркнем, что в случае налогообложения потребляемого топлива или выбросов значительное повышение затрат затронет

не только авиакомпании, эксплуатирующие старые воздушные суда, но даже авиакомпании, заменившие их на изделия нового поколения. В результате ослабевают стимулы к ускоренному обновлению техники, поскольку достигаемое благодаря ему сокращение эксплуатационных затрат на фоне существенно возросшего общего уровня затрат в относительном выражении невелико. Кроме того, столь существенное повышение эксплуатационных затрат подрывает финансовые возможности авиакомпаний, блокируя процессы обновления парка.

Различие принципов действия описанных механизмов стимулирования и их эффективности можно наглядно показать на рис. 1.

На нем изображены графики изменения по мере увеличения цены топлива следующих величин (расчеты проведены с использованием исходных данных вышеприведенного примера; прочие составляющие эксплуатационных расходов для старого и нового типов изделий приняты равными 1000 долл./л.ч.):

- стоимости летного часа самолетов старого и нового типов¹ в отсутствие государственного регулирования (немаркированные линии – соответственно сплошная и штриховая). Видно, что досрочная замена авиатехники становится выгодной лишь при цене авиатоплива свыше 1600 долл./т;

- стоимости летного часа самолетов старого и нового типов (соответственно крестообразные и круглые маркеры) при налогообложении выбросов по минимально необходимым, при данной цене топлива, ставкам. Видно, что при ценах авиатоплива, меньших 1600 долл./т, этот механизм повышает эксплуатационные затраты как для старого, так и для нового типов самолетов до 5000 долл./л.ч., т.е. до уровня, достигаемого при цене авиатоплива 1600 долл./т;

- стоимости летного часа самолетов старого типа (черные треугольные маркеры) при непосредственном налогообложении

¹ Для изделий нового типа – с учетом амортизации, поскольку на их приобретение придется нести дополнительные издержки, тогда как изделия старого типа уже имеются.

их эксплуатации по минимально необходимому, при данной цене топлива, ставкам. Видно, что при ценах авиатоплива, меньших 1600 долл./т, этот механизм повышает эксплуатационные затраты для старого типа самолетов до уровня эксплуатационных затрат нового типа самолетов, с учетом амортизации;

- стоимости летного часа самолетов нового типа (белые треугольные маркеры) при непосредственном налогообложении их эксплуатации по тем же ставкам, которыми облагается эксплуатация самолетов старого типа. Видно, что даже такое, очевидно несправедливое налогообложение повысит эксплуатационные затраты для нового типа самолетов существенно меньше, чем «справедливое» налогообложение выбросов (так, например, при цене авиатоплива 1000 долл./т – до 4100 долл./л.ч. против 5000 долл./л.ч.).

Качественно иная ситуация наблюдалась бы лишь в том случае, если бы удельный расход топлива при переходе к новому поколению авиатехники сокращался бы более, чем вдвое – например, до 1 т/л.ч. (что, очевидно, нереально в обозримом будущем). Соответствующие графики приведены на рис. 2.

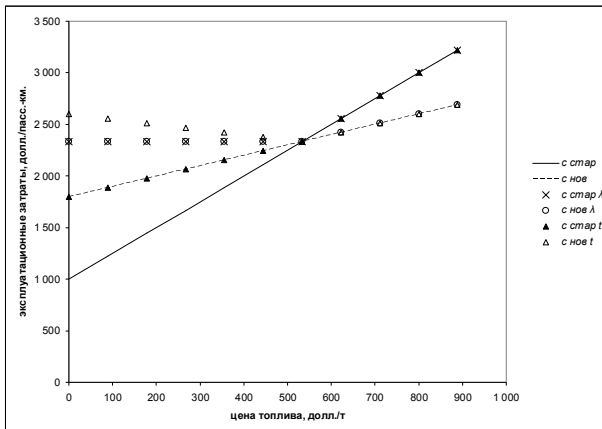


Рис. 2. Сравнение эксплуатационных затрат для старых и новых изделий при различных способах стимулирования обновления парка (пример 2)

Впрочем, в этом случае ускоренную замену морально устаревших изделий не пришлось бы стимулировать при современных ценах на авиатопливо, поскольку условие (1) выполнялось бы естественным образом.

4. Определение ставок экологического налогообложения с учетом коррупционных рисков

Итак, даже минимально необходимые ставки налогов на выбросы могут привести к существенному повышению затрат и для владельцев новой, более экологически чистой техники, в то время как налогообложение старых изделий на них официально вообще не распространяется. Однако наибольшие – и отнюдь не безосновательные – опасения у экономистов либерального направления вызывает реальная практика применения дискреционных мер. То, что управляющие воздействия реализуются не в «автоматическом», а «в ручном режиме», открывает значительные возможности коррупции, оппортунистического поведения чиновников, принимающих решения, поиска ренты с их стороны. Вполне возможно, что при непосредственном налогообложении эксплуатации старой техники владельцам новой, более экологически чистой техники, придется доказывать чиновникам, что их не следует облагать соответствующими налогами. И нельзя исключать возможность несправедливого применения к ним этих налогов – несмотря на то, что они уже понесли дополнительные издержки на повышение экологической чистоты своего имущества. Однако, как это ни парадоксально, даже такое, очевидно несправедливое применение дискреционных мер (в данном случае – непосредственного налогообложения старой техники) может оказаться для владельцев новой техники более выгодным, чем справедливое (по определению) применение встроенных регуляторов (т.е. налогообложения выбросов). Рассмотрим затраты владельца новой техники, которого несправедливо облагают налогом на эксплуатацию старой техники по ставке t_{min} :

$$c_t^{\text{нов}} = c^{\text{нов}} + t_{\text{min}},$$

и сравним их с затратами владельца новой техники в случае налогообложения выбросов по ставке λ_{min} :

$$c_{\lambda}^{\text{нов}} = c^{\text{нов}} + \Delta c_{\lambda}^{\text{нов}} = c^{\text{нов}} + \frac{g^{\text{нов}}}{(g^{\text{стар}} - g^{\text{нов}})} * t_{\text{min}}.$$

Найдем пороговое соотношение удельного расхода топлива старых и новых изделий, при котором первый вариант будет выгоднее для владельца изделий:

$$c_t^{\text{нов}} < c_{\lambda}^{\text{нов}}, \text{ при } t_{\text{min}} < \frac{g^{\text{нов}}}{(g^{\text{стар}} - g^{\text{нов}})} * t_{\text{min}}, \text{ или}$$

$$g^{\text{нов}} > g^{\text{стар}} - g^{\text{нов}}, \text{ т.е. } g^{\text{нов}} > \frac{g^{\text{стар}}}{2}.$$

Итак, если для изделий нового поколения не удается снизить удельный расход топлива хотя бы вдвое, минимально действенные ставки налогов на выбросы будут столь высоки, что даже владельцы новой техники скорее предпочтут стать жертвой несправедливого применения прямых налогов на эксплуатацию старых изделий. На данном этапе развития технологий не приходится рассчитывать на столь радикальное повышение топливной экономичности тепловых двигателей – напротив, как было отмечено выше, вероятнее всего, $g_{\text{нов}} \approx g_{\text{стар}}$. Следовательно, на данном этапе технологического развития многих отраслей экономики дискреционные меры стимулирования ускоренной замены техники на более экологически чистую более предпочтительны для объектов регулирования, чем встроенные регуляторы, даже с учетом коррупционных рисков. Этот вывод, неочевидный без проведенного выше технико-экономического анализа, может иметь большое значение для обоснования государственной экономической политики.

При анализе эффективности дискреционных инструментов и определении их количественных параметров необходимо учитывать коррупционные риски. Если владелец оборудования

руководствуется критерием ожидаемой полезности, можно найти пороговую вероятность несправедливого обложения владельцев новой техники налогом на эксплуатацию старых изделий, по достижении которой всякие стимулы к обновлению оборудования пропадают (если же агенты не расположены к риску, этот порог сокращается). Пусть вероятность стать жертвой несправедливого применения налога равна $p < 1$. Тогда владелец старой техники, продолжая ее эксплуатацию, гарантированно несет издержки в размере $c_i^{\text{стар}} = c^{\text{стар}} + t$. В то же время, заменяя технику на новую и более экологически чистую, он приобретает лотерею:

- с вероятностью $(1 - p)$ он будет нести издержки в размере $c^{\text{нов}}$,
- с вероятностью p он подвергнется несправедливому обложению налогом t , и его издержки возрастут до уровня $c_i^{\text{нов}} = c^{\text{нов}} + t$.

Таким образом, ожидаемый уровень затрат владельца новой техники с учетом коррупционных рисков составит

$$(19) \bar{c}^{\text{нов}} = p * (c^{\text{нов}} + t) + (1 - p) * c^{\text{нов}} = c^{\text{нов}} + p * t.$$

До сих пор считалось (см. формулу (14)), что минимально действенная ставка налога на эксплуатацию старого изделия t_{\min} должна уравнивать эксплуатационные затраты владельцев новой техники (с учетом амортизации) и старой, т.е. $c_i^{\text{стар}} = c^{\text{нов}}$. Однако, если учесть коррупционные риски, при такой ставке налога замена техники на новую станет заведомо невыгодной. Следовательно, минимально действенная ставка налога на эксплуатацию старой техники должна удовлетворять следующему условию:

$$\bar{c}^{\text{нов}} < c_i^{\text{стар}}, \quad \text{или} \quad c^{\text{нов}} + p * t < c^{\text{стар}} + t, \text{ т.е.}$$

$$(20) t_{\min} = \frac{c^{\text{нов}} - c^{\text{стар}}}{1 - p}.$$

Таким образом, чем выше коррупционные риски, тем выше должна быть ставка дискреционного налога, чтобы он оказывал стимулирующее воздействие. В то же время выше было

обосновано, что дискреционное налогообложение предпочтительнее встроенного регулятора (в виде платежей за выбросы) именно потому, что ставки могут быть сравнительно невысокими. Сохранится ли такое соотношение с учетом коррупционных рисков? Фактически ставку налога на эксплуатацию старых изделий следует умножить на $1/(1+p)$. Но и с учетом коррупционного риска такой налог останется априори предпочтительнее для владельца новой техники, чем налог на выбросы, при выполнении следующего условия:

$$\bar{c}^{\text{нов}} = c^{\text{нов}} + p^*t < c_{\lambda}^{\text{нов}},$$

где

$$c_{\lambda}^{\text{нов}} = c^{\text{нов}} + \Delta c_{\lambda}^{\text{нов}} = c^{\text{нов}} + \frac{g^{\text{нов}}}{(g^{\text{стар}} - g^{\text{нов}})} * (c^{\text{нов}} - c^{\text{стар}}).$$

Таким образом, ставка налога на летный час эксплуатации изделий старого типа должна быть не выше следующего уровня (условие предпочтительности перед налогообложением выбросов):

$$(21) \quad t_{\max} = \frac{g^{\text{нов}}}{(g^{\text{стар}} - g^{\text{нов}})} * \frac{c^{\text{нов}} - c^{\text{стар}}}{p}$$

и одновременно, как показано выше – не менее $t_{\min} = (c^{\text{нов}} - c^{\text{стар}})/(1-p)$ (условие действенности с точки зрения стимулирования обновления техники, см. формулу (20)).

Диапазон допустимых ставок налога $(t_{\min}; t_{\max})$ будет непустым, если $\frac{g^{\text{нов}}}{(g^{\text{стар}} - g^{\text{нов}})} > \frac{p}{1-p}$. Вероятность того, что владелец новой, более экологически чистой техники подвергнется штрафу за использование старой техники, должна удовлетворять следующему неравенству (полученному преобразованием вышеуказанного условия непустоты диапазона допустимых ставок налога):

$$p < p_{\text{кр}} = \frac{g^{\text{нов}}}{g^{\text{стар}}}.$$

Заметим, что чем радикальнее повышение топливной экономичности изделий нового поколения, тем ниже допустимый уровень коррупционного риска. И наоборот, если улучшение технико-экономических параметров изделий замедляется ($g_{\text{нов}} \approx g_{\text{стар}}$), экономические агенты будут более «терпимо» относиться даже к высокому риску несправедливого применения дискреционных мер, предпочитая его «справедливым» и «беспристрастным», но слишком жестким встроеным регуляторам.

Разумеется, если коррупция станет тотальной, т.е. если все владельцы новой техники будут гарантированно¹ несправедливо облагаться налогом на эксплуатацию старых изделий ($p \equiv 1$), стимулирующий эффект таких мер пропадет, поскольку априори не будет выгодно инвестировать в приобретение новой техники. В то же время в рамках данной работы не представляется возможным описать все многообразие проявлений «отказов государства» (начиная хотя бы с уклонения экономических агентов от уплаты любых экологических налогов и штрафов), не говоря уже о выработке мер противодействия этим явлениям.

5. Заключение

1. Досрочная замена оборудования, имеющего остаток ресурса, эффективна, если цена производства нового изделия ниже разности цен эксплуатации старого и нового изделий (т.е. если экономия в эксплуатации превышает дополнительные затраты, связанные с производством нового изделия). Такое правило справедливо для экономических цен (денежных затрат), энергетических цен (энергозатрат) и экологических цен (вредных выбросов).

2. На данном этапе технологического развития ускоренная замена тепловых двигателей вначале становится эффективной с

¹ Впрочем, в этом случае уже не приходится говорить о «коррупционных рисках», поскольку понятие «риск», при всей широте его интерпретации, все-таки обязательно связано с неопределенностью.

экологической и энергетической точек зрения, и лишь затем – с экономической. Поэтому даже если повышение экологической чистоты транспортного и энергетического оборудования сопровождается повышением его экономичности, естественных экономических стимулов может быть недостаточно для ускоренной замены изделий на более экологически чистые. Для обеспечения обновления парка государство может использовать административное принуждение либо экономическое стимулирование. Для некоторых видов изделий длительного пользования (например, бытовой техники и электроники), наоборот, вначале досрочная замена становится эффективной с экономической точки зрения, и лишь затем – с экологической. В соответствующих отраслях, во избежание излишнего потребления ресурсов и избыточной нагрузки на окружающую среду, целесообразно блокировать излишне частую замену изделий длительного пользования (также экономическими инструментами либо прямым административным принуждением).

3. При замедлении темпов совершенствования технологий, встроенные регуляторы (например, налогообложение вредных выбросов) становятся менее эффективными, чем дискреционные меры (например, прямое налогообложение эксплуатации старых изделий). Если уровни выбросов при эксплуатации новых изделий лишь незначительно ниже, чем старых, для стимулирования ускоренной замены изделий ставки налогов на выбросы должны быть настолько высокими, что даже владельцы новой и более экологически чистой техники испытают многократное повышение эксплуатационных расходов. Поэтому дискреционные меры становятся более предпочтительными для владельцев оборудования, чем встроенные регуляторы, даже с учетом коррупционных рисков.

Литература

1. БУРИЧЕНКО Л.А., ЕНЕНКОВ В.Г., НАУМЕНКО И.М., ПРОТОЕРЕЙСКИЙ А.С. *Охрана окружающей среды в гражданской авиации*. - М.: Машиностроение, 1992. - 320 с.

2. ГУСМАНОВ Т.М., КЛОЧКОВ В.В. *Экономические проблемы развития авиационной промышленности в условиях нестабильного спроса на авиаперевозки* // Экономическая наука современной России. – 2008. – №3. – С. 98–109.
3. ДМИТРИЕВ В.Г., МУНИН А.Г. *Экологические проблемы гражданской авиации* // Аэрокосмический курьер. – 2003. – №2. – С. 15–17.
4. КЛИНСКИЙ Б., НАЗАРЕНКО Ю. *К вопросу об антропогенном изменении климата, и о проблемах с Монреальским и Киотским протоколами* // Двигатель. – 2005. – №6.
5. КЛОЧКОВ В.В. *Управление инновационным развитием гражданского авиастроения*. – М.: ГОУ ВПО МГУЛ, 2009. – 280 с.
6. КЛОЧКОВ В.В., ГУСМАНОВ Т.М. *Экологические стандарты как инструмент стимулирования спроса на продукцию авиационной промышленности* // Маркетинг в России и за рубежом. – 2007. – №3. – С. 39–45.
7. КЛОЧКОВ В.В., ИГНАТЬЕВА А.И. *Эколого-экономические проблемы обновления мирового парка авиатехники* // Экономика природопользования. – 2009. – №2. – С. 23–40.
8. КЛОЧКОВ В.В., ШКАДОВА А.А., ЖДАНОВСКИЙ А.В. *Экономические аспекты морального устаревания техники* // Технология машиностроения. – 2008. – №11. – С. 65–70.
9. КЛОЧКОВ В.В., ШУСТОВ А.В., ГУСМАНОВ Т.М. *Экологические нормы как фактор конкурентной борьбы на рынках авиаперевозок и авиатехники* // Авиакосмическая техника и технология. – 2007. – №3. – С. 61–70.
10. КОСТРОМИНА Е.В. *Экономика авиакомпании в условиях рынка* - М.: НОУ ВКШ «Авиабизнес», 2002. – 304 с.
11. СКИБИН В., ВОЛКОВ С. *Выбросы вредных веществ от авиационных двигателей* // Аэрокосмический курьер. – 2003. – №2. – С. 18–19.
12. УШАКОВ Е.П. *Экологически чистые технологии в преодолении экономического кризиса: модельный анализ* // Экономическая наука современной России. – 2009. – №4. – С. 54–64.

13. ФЕДОРОВ Б.Г. *Экономико-экологические аспекты выбросов углекислого газа в атмосферу* // Проблемы прогнозирования. – 2004. – №5. – С. 86–101.
14. ФЕДОРОВ Б.Г. *Посткиотская экономика России* // Проблемы прогнозирования. – 2007. – №4. – С. 74–83.
15. *Aircraft Engine and Engine Parts Manufacturing: 2002* // In: 2002 Economic Census. Manufacturing. Industry series. U.S. Census Bureau, 2004. – P.47.
16. BRÄNNLUND R. *Environmental policy without costs? A review of the Porter hypothesis* // S-WoPEc, Swedish Working Papers in Economics, Umeå Economic Studies. – 2009. – №766. – P. 49.
17. HERRING H. *Rebound effect* // In: Encyclopedia of Earth. Eds. Cutler J. Cleveland (Washington, D.C.: Environmental Information Coalition, National Council for Science and the Environment). – 2008. URL: http://www.eoearth.org/article/Rebound_effect.
18. PORTER M.E. *America's Green Strategy* // Scientific American. – 1991. – Vol. 264, №4. – P. 96.
19. PORTER M.E., VAN DEN LINDEN C. *Toward a New Conception of the Environment-Competitiveness Relationship* // Journal of Economic Perspectives. – 1995. – Vol. 9, № 4.
20. WAGNER M. *The Porter Hypothesis Revisited. A Literature Review of Theoretical Model and Empirical Test* // Lüneburg: Centre for Sustainability Management, 2003. - P. 46. (http://www2.leuphana.de/umanagement/csm/content/nama/downloads/download_publicationen/38-2downloadversion.pdf)
21. XEPAPADEAS A. DE ZEEUW A. *Environmental Policy and Competitiveness: The Porter Hypothesis and the Composition of Capital* // Journal of Environmental Economics and Management. – 1999. – Vol.37, №2. – P. 165–182. – URL: <http://ideas.repec.org/a/eee/jee/v37y1999i2p165-182.html>

THE ROLE OF TAXES IN ECOLOGICAL COMPATIBILITY IMPROVEMENT OF DURABLE EQUIPMENT

Vladislav Klochkov, Institute of Control Sciences of RAS, Moscow, Doctor of Science (Moscow, Profsoyuznaya st., 65, vlad_klochkov@mail.ru).

Sergey Grivskiy, Moscow Institute of Physics and Technology, Moscow, postgraduate (serg.mipt@gmail.com).

Anastasia Ignatieva, Moscow Institute of Physics and Technology, Moscow, postgraduate.

Abstract: Mathematical and economical models are used to analyze incentives of durable equipment owners to accelerate its renewal to ecologically pure one. Conditions are revealed when the government has to stimulate acquisition of ecologically safer equipment (even when this equipment is money-saving). Comparative analysis is performed of stimulation discretionary instruments (such as taxation of old goods) and built-in regulators (such as taxation of toxic emission).

Keywords: durable equipment, ecology, economic effectiveness Porter hypothesis, stimulation, corruption risks.

*Статья представлена к публикации
членом редакционной коллегии Р. М. Нижегородцевым*

УДК 519.234.3
ББК 2.2.22.172

ОБ ОДНОЙ НЕЙРОСЕТЕВОЙ МОДЕЛИ ДИАГНОСТИКИ ВЕНОЗНЫХ ЗАБОЛЕВАНИЙ

Веденяпин Д. А.¹, Лосев А. Г.²

(Волгоградский государственный университет, Волгоград)

Изучается задача диагностики заболеваний вен нижних конечностей, в частности, предлагается использовать нейронные сети для классификации наблюдений, полученных методом комбинированной термометрии. Установлено, что нейронные сети являются достаточно эффективным методом диагностики заболеваний по данным комбинированной термометрии, а получаемая точность допускает их использование в экспертных комплексах.

Ключевые слова: нейронные сети, диагностика венозных заболеваний, комбинированная термометрия.

1. Введение

Несмотря на высокий уровень развития математического моделирования, и вычислительной техники, задачи классификации объектов различной природы остаются весьма нетривиальными. Это обусловлено рядом обстоятельств, среди которых и сложность формализации объектов, и проблематичность их описания в количественной шкале, и множество других факторов. Подавляющее большинство исследователей сходится во мнении, что создание универсальной распознающей системы, без приложения к какой-либо реальной задаче, оказывается

¹ Дмитрий Александрович Веденяпин, аспирант
(stmatont@gmail.com).

² Александр Георгиевич Лосев, доктор физико-математических наук, профессор (alexander.losev@volsu.ru).

бесплодной затеей. Значительно более плодотворным является подход, при котором вначале решается реальная практическая задача, а затем применяемые подходы обобщаются на группу проблем, близких первоначальной.

Одним из важнейших видов задач классификации являются проблемы медицинской диагностики. Разработка диагностических технологий на основе интеграции медицинских знаний и различных математических методов искусственного интеллекта для интерпретации и анализа полученных результатов, является крайне актуальной задачей. При этом, зачастую трудности диагностики возникают не из-за дефицита необходимой информации, а из-за отсутствия соответствующих объективных методов ее структуризации, выделения симптом-комплексов, выявления наиболее значимых и существенных признаков того или иного заболевания вен, и т.д. Многие из указанных выше проблем вполне можно решить созданием интеллектуальных аппаратно-программных комплексов. Кроме того отметим, что в настоящее время приоритетными в медицине являются неинвазивные, безопасные как для пациентов, так и для медицинского персонала, способы диагностики. Одним из таких методов является комбинированная термометрия [2, 4, 5, 7], применяемая с диагностическими целями в различных областях медицины (маммология, гинекология, урология, неврология, хирургия и др.) и доказавшая свою высокую информативность и абсолютную безвредность. Однако, скорее всего в силу «молодости», диагностика по данным комбинированной термометрии остается искусством врача, а разработка математических моделей и методов диагностики находится в самом начале пути. Данная работа посвящена изучению проблем диагностики заболеваний вен нижних конечностей по данным комбинированной термометрии.

2. Методика комбинированной термометрии в диагностике венозных заболеваний

В течение последних нескольких лет разрабатывается методика обследования нижних конечностей с помощью комбинированной термометрии (см., [2, 7, 8]). Обследование пациентов проводится посредством последовательного измерения кожной (ИК) температуры и глубокой микроволновой термометрии (РТМ) в 12 симметричных точках, расположенных по задней поверхности обеих голени пациента, в положении пациента «лежа на животе» и «стоя». Точки измерения были выбраны, исходя из анатомического строения голени и внутреннего кровотока ноги, следующим образом (см. рис. 1).

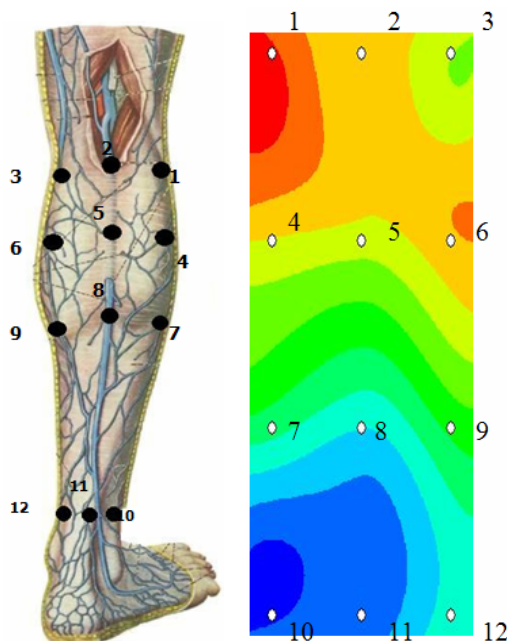


Рис. 1.

Первые три точки расположены: на вершине латеральной головки икроножной мышцы (1), в подколенной ямке (2), на

вершине медиальной головки икроножной мышцы (3). Второй ряд точек расположен: по центру латеральной головки икроножной мышцы (4), между головками икроножной мышцы (5), и по центру медиальной головки икроножной мышцы (6). Третий ряд точек расположен в нижней части икроножной мышцы – латерально (7), в центре (8) и медиально (9). Последние точки измерения температуры находятся на наружной стороне ахиллова сухожилия в области латеральной лодыжки (10), по центру ахиллова сухожилия (11) и по его внутренней поверхности в области медиальной лодыжки (12).

Характер данных групп и выбор точек измерения были определены исходя из ранее накопленного материала и биологических особенностей строения нижних конечностей. После анализа полученной информации, врач ставит диагноз пациенту и, при необходимости, направляет пациента на дальнейшее обследование более дорогостоящими и/или более опасными методами.

При исследовании была создана экспертная база (обучающая выборка) пациентов, которая была разделена на 2 контрольные группы – группу здоровых людей (Норма) и группу больных варикозной болезнью. Группа больных, в свою очередь, была разбита на 3 подгруппы в зависимости от степени заболевания: варикозная болезнь (ВБ), посттромботическая болезнь (ПТБ), острый венозный тромбоз (ОВТ). На каждого исследуемого заполнялась индивидуальная карта, в которую, помимо температурных данных, также была внесена информация о пациенте, температура окружающей среды во время обследования, информация, основанная на осмотре пациента, и пр. После заполнения карт собранная информация вносилась в общую базу данных.

Обучающая выборка состоит из данных на 214 голеней пациентов. Среди них 38 голеней полностью здоровых пациентов, 56 голеней больных ВБ, 12 голеней больных ОВТ, 58 голеней больных ПТБ, 50 здоровых голеней у пациентов со второй больной голенью.

Одним из самых важных этапов при решении задачи нейросетевого прогнозирования является формирование обучающей выборки [1]. Традиционно, для получения входных параметров нейронной сети применяют следующие методы: предварительная обработка входных параметров таким образом, чтобы среднее значение по всему обучающему множеству было близко к нулю, а параметры некоррелированы [12]; предобработка качественных признаков [10, 14]; линейный сдвиг данных в отрезок $[-1, 1]$; проведение нелинейной предобработки; метод пошагового отбора входных параметров. В данной работе предлагается метод использования в качестве дополнительных входных параметров функций от температуры, предоставляющие высокоинформативные качественные признаки заболеваний. Ряд этих функций был предложен врачами-флебологами и был получен ими на основании исследований физиологических особенностей голени нижних конечностей. Например, наличие экстремумов в некоторых точках. Большинство других было получено в результате статистической обработки данных (смотри, например, [2, 8]).

В результате ряда численных экспериментов в качестве входных параметров нейронной сети для диагностики «Здоров/Болен» были выбраны 67 входных параметра:

- температуры в 12 точках голени, полученные датчиком РТМ в положении лежа;
- температуры в 12 точках голени, полученные датчиком ИК в положении стоя;
- 4 условия на наличие боли, отека, кожных изменений;
- 39 функций от температур, описывающих наличие аномально горячих (холодных) зон, асимметрий и других качественных признаков, определяющих наличие патологий в голени.

Для диагностики «Здоров/ВБ/ПТБ/ОВТ» для каждой из четырех нейронных сетей были выбраны свои параметры:

- 26 функций от температур, предоставляющих качественные признаки наличия патологий в голени;

- 10 функций от температур, предоставляющих признаки, характерные для пациентов с диагнозом «ОВТ»;
- 14 функций от температур, предоставляющих признаки, характерные для пациентов с диагнозом «ПТБ»;
- 7 функций от температур, предоставляющих признаки, характерные для пациентов с диагнозом «ВБ».

Отдельно отметим, что выбор величин, таких как температурные экстремумы, наличие боли, «степень» асимметрии и т.д., объясняется их высокой информативностью при диагностике.

3. Нейронная сеть

Для построения нейронных сетей, способных решать задачи диагностирования варикозной болезни, необходимо сформировать их топологии, определить механизм обучения и алгоритм тестирования. Помимо этого, необходимо сформировать обучающую базу данных для проведения процесса обучения нейронной сети.

На основе анализа задачи и была выбрана наиболее популярная в подобных задачах сигмоидальная функция активации [3]. В работе использовалась многослойная полносвязная нейронная сеть, подтвердившую свою применимость в задачах диагностики в медицине (см., например, [3, 11, 12]). Были опробованы различные архитектуры нейронных сетей, так как обычно количество нейронов скрытого слоя подбирается экспериментально [9]. На основании проведенных тестов было принято решение остановиться на модели двухслойной нейронной сети прямого распространения. Обучение нейронной сети производилось одним из наиболее популярных алгоритмов обучения многослойных нейронных сетей – алгоритмом обратного распространения ошибки [6, 13].

Для нахождения оптимальных топологий нейронных сетей был выбран метод перекрестной проверки данных, который был

апробирован в [12]. Суть метода перекрестной проверки данных заключается в том, что из обучающей выборки последовательно исключаются пациенты, затем нейронная сеть последовательно для каждого исключенного пациента обучается на получившейся обучающей выборке и, затем, считается процент верно диагностированных пациентов.

В большинстве случаев в задачах диагностики используется одна нейронная сеть. В данной работе мы предлагаем использовать при дифференциальной диагностике («ПТБ/ОВТ/ВБ/ Здоров») последовательно три нейронные сети. Первая нейронная сеть отделяет голени с диагнозом ПТБ от остальных. Вторая нейронная сеть отделяет голени с диагнозом ОВТ от оставшихся голеней. Третья нейронная сеть отделяет голени с диагнозом ВБ от оставшихся голеней. Оставшимся голеням автоматически ставится диагноз «Здоров». Не очень давно в работе [6] была предложена подобная идея. Отметим, что в нашем случае применение данного метода дало более чем приличные результаты.

Ниже представлены (см. таблицу 1) топологии нейронных сетей показавшие наилучшие средние результаты при описанном выше методе проверки результатов обучения:

Таблица 1. Результаты обучения нейронных сетей

	Диагностика «Здоров / Болен»	Диагностика «ПТБ / ОВТ / ВБ / Здоров»
Количество входных параметров	67	57
Количество выходных параметров	2	2
Число нейронов в первом скрытом слое	4	4
Число нейронов во втором скрытом слое	2	2
Алгоритм обучения	Обратное распространение	Обратное распространение

	Диагностика «Здоров / Болен»	Диагностика «ПТБ / ОВТ / ВБ / Здоров»
Функция активации	Сигмоидаль- ная	Сигмоидаль- ная
Процент верно диагностиро- ванных пациентов	83,3	59

Традиционно, на выходе нейронной сети используется один выходной параметр, на основании анализа которого осуществляется диагностика. В нашем случае для диагностики «Здоров/Болен» предлагается использовать на выходе нейросети двумерный вектор (x_1, x_2) , определяющий диагноз голени пациента: в случае, если $x_1 > x_2$, то предположительный диагноз голени пациента «Здоров», в случае, если $x_2 > x_1$, то предположительный диагноз голени пациента «Болен».

В случае диагностики «Здоров/ВБ/ПТБ/ОВТ» на выходе каждой нейросети также предлагается двумерный вектор (x_1, x_2) , определяющий диагноз пациента. В случае, если набор текущей нейронной сети состоит из двух значений таких, что $x_1 > x_2$, то ставится диагноз, соответствующей текущей нейронной сети, в противном случае управление передается следующей нейронной сети. Если текущая нейронная сеть – последняя, то автоматически ставится диагноз «ОВТ».

4. Результаты

Перечислим основные выводы, которые можно сделать по результатам работы нейронной сети:

1. В случае построения нейронных сетей для диагностики заболевания по данным термометрии различных органов:
 - а. В качестве входных параметров разумно применять функции, учитывающие физиологию органа, и, предоставляющие высокоинформативные диагностические признаки.

Они могут быть получены как в результате экспертного анализа, так и в результате статистической обработки данных.

2. При дифференциальной диагностике, более высокую точность дает комбинация отдельных сетей для каждого заболевания с двумя выходами в каждой сети.
3. Архитектура нейронной сети выбирается экспериментально, что является наиболее широко применяемым способом для задач классификации с помощью нейронных сетей.

В результате обучения описанных выше нейронных сетей методом обратного распространения были получены следующие результаты:

При дифференциальной диагностике («Здоров/ВБ/ПТБ/ОВТ») точность составила 59 %. При диагностике «Здоров/Болен» точность составила 83,3 %.

Работа выполнена при финансовой поддержке Министерства образования и науки РФ по ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2012 годы» (госконтракт № 16.513.11.3067)

Литература

1. АВДЕЕВА В.М., КРЮЧКОВА И.Н. *Обработка статистических данных и определение состава входов нейросети в процессе формирования информационной базы для прогнозирования.* // Территория науки. – 2007. – №2(3). – С. 196–204.
2. АНИСИМОВА Е.В., ЗАМЕЧНИК Т.В., ЛОСЕВ А.Г., МАЗЕПА Е.А. *О некоторых характерных признаках в диагностике венозных заболеваний нижних конечностей методом комбинированной термометрии.* // Вестник новых

- медицинских технологий. – 2011. – Т. XVIII., №2. – С. 329–330.
3. АРАВИН О.И. *Применение искусственных нейронных сетей для анализа патологий в кровеносных сосудах.* // Российский журнал биомеханики. – 2011. – Т.5, №3 (53).
 4. ВАЙСБЛАТ А.В. *Радиотермография как метод диагностики в медицине.* – М.: НЦЗД РАМН, 2003. – 80 с.
 5. ВЕСНИН С.Г. *Микроволновая радиотермометрия – национальное достояние России.* // Здоровоохранение. – 2007. – №9. – С. 159–164.
 6. ДУМЛЕР А.А., ПОЛЕЩУК А.Н., БОГДАНОВ К.В., ЧЕРЕПАНОВ Ф.М., ЯСНИЦКИЙ Л.Н. *Опыт создания нейросетевой системы для диагностики сердечно-сосудистых заболеваний.* // Вестник Пермского университета. – 2011. – Вып. 1(5). – С. 95–101.
 7. ЗАМЕЧНИК Т.В., ЛАРИН С.И., ЛОСЕВ А.Г., ОВЧАРЕНКО Н.С. *Способ комбинированной термометрии и математические модели вероятностной диагностики заболеваний вен нижних конечностей.* // Вестник новых медицинских технологий. – 2009. – Т. XVI., №4. – С. 14–16.
 8. ЛОСЕВ А.Г., СТАВРОВ Т.А. *Об одном алгоритме классификации в методе комбинированной термометрии диагностики венозных заболеваний.* // Естественные и технические науки. – 2011. – №5. – С. 268–270.
 9. МЕДВЕДЕВ В.С., ПОТЕМКИН В.Г. *Нейронные сети. MATLAB 6* – М: Диалог-МИФИ, 2002. – 496 с.
 10. МИРКЕС Е.М. *Нейрокомпьютер: проект стандарта.* – Новосибирск: Наука, Сибирская издательская фирма РАН., 1998. – 337 с.
 11. ПЯТАКОВИЧ Ф.А., ЯКУНЧЕНКО Т.И., ХЛИВНЕНКО Л.В., ВАСИЛЬЕВ В.В., МАККОНЕН К.Ф., МАСЛОВА О.В. *Разработка моделей и алгоритмов нейросетевой классификации степени активности автономной нервной системы и оценка их адекватности на обучающей и экзаменационных выборках.* // Фундаментальные исследования. – 2011. – №2 – С. 136–141.

12. СОЛОВОВ В.А., ФРОЛОВА И.Г. *Использование логистических регрессий и нейронных сетей в выявлении рака предстательной железы.* // Сибирский онкологический журнал. – 2006. – №1. – С. 14–17.
13. ХАЙКИН С. *Нейронные сети. Полный курс.* – М.: Вильямс, 2006. –995 с.
14. ЦАРЕГОРОДЦЕВ В.Г. *Оптимизация предобработки данных: константа Липшица обучающей выборки и свойства обученных нейронных сетей.*// Нейрокомпьютеры: разработка, применение. – 2003. – №7. – С. 3–8.

NEURAL NETWORKS IN VASCULAR DISEASES DIAGNOSIS

Dmitriy Vedenyapin, Volgograd State University, Volgograd, P.G. (dima.vedenyapin@gmail.com).

Alexander Losev, Volgograd State University, Volgograd, Doctor of Science, professor (alexander.losev@volsu.ru).

Abstract: We consider the problem of vascular diseases diagnosis and suggest using neural networks for classification of combined thermometry observation results. We show that neural networks are an efficient method of vascular disease diagnosis. The accuracy of this method allows using neural networks in expert systems.

Keywords: Neural networks, vascular diseases diagnosis, combined thermometry.

*Статья представлена к публикации
членом редакционной коллегии Д. А. Новиковым*

УДК 629.764
ББК 39.62

АНАЛИЗ ВОЗМОЖНОСТЕЙ ПОВЫШЕНИЯ БЕЗОПАСНОСТИ ЭКСПЛУАТАЦИИ ПЕРСПЕКТИВНЫХ РАКЕТНЫХ СРЕДСТВ ВЫВЕДЕНИЯ НА ОРБИТЫ

Андриенко А. Я.¹, Тропова Е. И.², Чадаев А. И.³
*(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)*

Приводятся результаты исследований по перестройке базовой стратегии выведения ракет-носителей (РН) на околоземные орбиты – перестройке, направленной на повышение безопасности эксплуатации РН.

Ключевые слова: ракета-носитель, безопасность эксплуатации, стратегия выведения на орбиты.

1. Введение

В самом конце XX века в Институте проблем управления им. В.А. Трапезникова РАН была предложена концепция [1 и др.] перестройки базовой стратегии управления выводением ракеты-носителя (РН) на околоземные орбиты – перестройки, направленной на смену приоритетов в критериях качества выведения: в качестве основного принимается критерий, характеризующий безопасность эксплуатации средств выведения (СрВ). В результате этой перестройки оказывается возможным, в частно-

¹ Анатолий Яковлевич Андриенко, заведующий лабораторией, доктор технических наук, профессор (vladgic@ipu.ru).

² Елена Ивановна Тропова, научный сотрудник (тел. (495) 334-88-71).

³ Александр Иванович Чадаев, старший научный сотрудник, кандидат технических наук (тел. (495) 334-88-71).

сти, выведение в космос по любому азимуту, поэтому стратегию [1] будем в дальнейшем называть всеазимутальной.

Посредством полноразмерного моделирования на ЭВМ процессов выведения применительно к *существующему* парку РН были выявлены последствия гипотетического перехода на всеазимутальную стратегию управления выведения. Отметим наиболее примечательные из этих последствий.

1. Повышается безопасность выведения полезного груза (ПГ) на орбиту из-за снижения аэродинамических нагрузок на РН и облегчения условий спасания ПГ в нештатных ситуациях на атмосферном участке полёта РН.

2. Снижается до предельно минимального значения ущерб землепользования при штатных пусках РН вследствие ликвидации территориальных районов штатного падения отработавших частей РН вне космодрома.

3. Сокращается (в десятки раз) вероятность нанесения экологического ущерба вне территории космодрома при аварийных ситуациях в процессе выведения.

4. Снижается, как следует из дальнейшего изложения, загрязнение космического пространства при эксплуатации СрВ.

Однако обескураживающе глубоким оказалось выявленное падение (почти двукратное) грузоподъёмности существующего парка ракет-носителей при перестройке стратегии управления их выведением. И основной причиной тому явилось несоответствие облика традиционных РН всеазимутальной стратегии управления выведением. Этот облик – с примерно равномерным распределением по ступеням общего запаса характеристической скорости РН и уменьшающейся от ступени к ступени тяговооруженностью – сложился, главным образом, из условия максимизации массы ПГ, выводимого на орбиту.

2. Облик перспективных ракетных средств выведения на орбиты

Перспективные СрВ, адекватные задаче безопасного выведения, при всём своём потенциальном разнообразии заметно отлича-

ются от традиционных РН. Приведём краткое, сопоставительное с традиционными РН, описание предельного облика таких СрВ.

1. Первая, лифт-бустерная, ракетная ступень предназначена для доставки РН на безатмосферный участок полёта (на высоту свыше 50 км) по практически вертикальной траектории (независимо от азимута пуска), выбираемой из условия приведения всех потенциальных точек падения¹ ступени на заданное безопасное удаление от стартовых сооружений (в данных исследованиях оно принималось равным 5 км). По тяговооружённости эта ступень не отличается от I ступеней традиционных РН, но в энергетическом отношении (по запасу топлива, характеристической скорости) она значительно слабее; масса сухой конструкции заметно больше, чем в традиционных РН, так как данная ступень СрВ должна оснащаться а) бортовой системой предохранения старта от разрушения (при аварийном выключении двигателей в районе стартовых сооружений), а также б) системами спасания (мягкой посадки) отработавшей I ступени и обеспечения многоразового её использования.

2. Вторая, разгонная, ступень предназначена для скорейшего увода РН с территории страны по заданному азимуту при относительно малых значениях угла тангажа, обеспечивающих приведение точки падения отработавшей второй ступени в акваторию Мирового океана². Эта ступень должна быть энергетически более совершенной, чем вторые ступени традиционных РН (в части отношения конечной массы к начальной и по удельному импульсу тяги). Тяговооружённость её должна быть больше обычной, а двигательная установка допускать глубокое дросселирование тяги (с возможным отключением основного двигателя ступени и с доразгоном РН на работающем рулевом

¹ Потенциальная точка падения для текущего момента времени определяется в предположении, что двигатель в этот момент выключается.

² Предполагается, что при любом азимуте запуска на дальности свыше 6000 км найдётся акватория Мирового океана, пригодная для захоронения отработавшей ступени РН.

двигателе) для ограничения перегрузок в условиях, порождаемых сильно меняющейся текущей массы ступени.

3. Третья, экстракторная¹, ступень предназначена для довыведения ПГ на орбиту при штатном режиме выведения и быстрого извлечения и увода ПГ от РН в аварийных ситуациях. В двигательную установку этой ступени входит блок твёрдотопливных двигателей (ТТД), при одновременном включении которых обеспечивается аварийное спасание ПГ (при ускорении до 8g); при последовательных включениях ТТД обеспечивает совместно с ЖРД малой тяги довыведение в штатном режиме полёта с небольшими перегрузками. Запас топлива (полного импульса тяги) – на порядок меньше, чем в традиционных ступенях.

3. Пример возможной реализации перспективных СрВ

Представленное в предыдущем разделе описание облика СрВ следует проиллюстрировать примером численного оценивания этого облика. В качестве такого примера принималась РН среднего класса, прототипом которой является РН «Зенит». Этот пример интересен в том отношении, что он связан с переводом данной РН с украинской на российскую конструкторско-производственную базу: к сожалению, этот перевод не сопровождался сколько-нибудь заметным совершенствованием российской РН по сравнению с прототипом.

Ввиду большого объёма работ, связанных с расчётом характеристик РН и траекторий их полёта, ограничимся здесь представлением двух содержательных версий перспективных РН, доставляющих некоторые приближения к предельному облику перспективных СрВ среднего класса. При построении этих версий, во-первых, будем ориентироваться на реальные возможности существующих ЖРД; во-вторых, будем воспроизводить

¹ Термин заимствован из зарубежных публикаций (см., например, [2]).

здесь только часть свойств предельного облика перспективных СрВ.

В первом цифровом столбце таблицы 1 приведены типизированные данные прототипа – РН «Зенит». Одна из особенностей прототипа, учитываемая и при баллистических расчётах версий перспективных СрВ, состоит в программировании тяги двигателей из условия обеспечения заданных ограничений по перегрузке.

Основной результат построения версий РН представлен во втором столбце таблицы 1 в виде значений параметров перспективной РН с общим запасом топлива, равным запасу топлива прототипа, но перераспределённым по ступеням в соответствии с всеазимутальной стратегией управления выведением. Баллистические расчёты РН *версии 1* проводились в предположении, что в составе версии используются

– на I ступени – тот же двигатель 11Д520, что и в прототипе;

– на II ступени для обеспечения потребной тяговооружённости (см. п.2, раздел 2) вместо рулевого двигателя 11Д513 прототипа – четырёхкамерный двигатель 11Д451 разработки КБХА, вместо основного двигателя 11Д123 – так называемая «четвертушка» двигателя 11Д520 (с высотным соплом);

– на III ступени – двигатель 11Д58М разработки НПО «Энергия».

Как следует из данных таблицы 1, перекомпоновка РН из двухступенчатого прототипа в трёхступенчатую РН версии 1 при сохранении общей массы топлива приводит к возрастанию массы сухой конструкции на 20%, к малозаметному изменению грузоподъёмности ракеты при традиционной стратегии управления выведением и к четырёхкратному снижению потерь в грузоподъёмности, вызванных перестройкой стратегии управления выведением. Выявленное здесь 10%-е снижение грузоподъёмности (с 12,0 до 10,7 т массы ПГ – см. последние две строки таблицы для РН прототипа и РН версии 1) из-за смены приоритетов в критериях качества управления выведением не столь уж существенно.

Таблица 1. Основные характеристики РН среднего класса¹

Наименование параметра РН	РН прото- тип	РН версии 1	РН версии 2
Стартовая масса, т	458	466	536
Масса рабочих запасов топлива			
I ступени, т	324	230	272
II ступени, т	81,2	173,4	196
III ступени, т	–	2,9 – 4,4	5,0
Масса сухой конструкции			
I ступени, т	32	32	34
II ступени, т	7,5	15,1	16,3
III ступени, т	–	0,4	0,5
Номинальный расход топлива			
I ступени, т/с	2,4	2,4	2,492
II ступени, т/с	0,266	0,6893	0,6893
III ступени, т/с	–	0,0242	0,0242
Удельный импульс тяги в пустоте на			
I ступени, с	336,2	336,2	340
II ступени, с	349	356	356
III ступени, с	–	360	360
Масса ПГ, выводимого на орбиту			
– при традиционной стра- тегии управления выве- дением, т	12,0	12,2	–
– при всеазимутальной стратегии управления выведением, т	5,5	10,7	12,2

¹ Примечание к таблице 1: данные по массе ПГ приведены применительно к выведению на круговую орбиту высотой 383 км и наклоном 56°.

В РН *версии 2* (последний столбец таблицы 1) повышение грузоподъёмности РН до первоначальной достигается с использованием тех же двигателей, что и в РН *версии 1* (с форсированием двигателя 11Д520 до предельного значения – на 3,8 %), так что стартовая масса РН возрастает на 15 %.

Наконец, отметим, что масса части конструкции СрВ, выводимой на орбиту вместе с ПГ во всеазимутальных РН *версий 1 и 2*, на порядок меньше, чем в РН прототипе; соответственно снижается и интенсивность загрязнения космического пространства при эксплуатации СрВ.

4. Частные проблемы создания перспективных СрВ

К ещё нерешённым проблемам, связанным с созданием перспективных СрВ, относится разработка методик обоснования

1) размеров и координат района посадки блоков первой ступени с учётом динамики активного и пассивного участков полёта в реальных условиях атмосферы над космодромом;

2) ограничений на зоны возможных районов падения отделяющихся элементов и блоков верхних ступеней СрВ в Мировом океане;

3) распределения допустимых перегрузок по ступеням СрВ;

4) управляемости и устойчивости движения СрВ на всех участках полёта.

5. Заключение

Представленные в статье результаты исследований возможностей экологизации эксплуатации РН свидетельствует, в частности, о том, что при переводе РН среднего класса на российскую конструкторско-промышленную базу имеется вполне реальная возможность создания ракеты-носителя нового поколения – всеазимутальной с повышенной безопасностью и экологичностью эксплуатации. Дальнейшее совершенствование такой РН связано с решением проблем предохранения стартовых

сооружений и спасания ПГ в аварийных ситуациях, обеспечения многоразового использования отдельных частей СрВ и др.

Литература

1. АНДРИЕНКО А.Я., ВОЛКОВ В.Я., ИВАНОВ В.П., ЧАДАЕВ А.И. *Исследование возможностей повышения безопасности средств выведения на основе выбора базовых стратегий управления выведением* // Отчёт Института проблем управления РАН по теме 059-94/08. – 1994. – С. 23–30.
2. ГРОДЗОВСКИЙ Г.Л., ИВАНОВ Ю.Н., ТОКАРЕВ В.В. *Механика космического полёта (проблемы оптимизации)*. – М.: Наука, 1975. – 704 с.

ANALYSING POSSIBILITIES OF OPERATION SAFETY INCREASE OF PROSPECTIVE ROCKET LAUNCHERS

Anatolii Andrienko, Institute of Control Sciences of RAS, Moscow, Laboratory Head, Doctor of Science, professor (Moscow, Prof-soyuznaya st., 65, (495) 334-88-71, vladguc@ipu.rssi.ru).

Elena Tropova, Institute of Control Sciences of RAS, Moscow, researcher, (495) 334-88-71, vladguc@ipu.rssi.ru

Alexander Chadaev, Institute of Control Sciences of RAS, Moscow, cand. Sc., senior scientific worker, (495) 334-88-71, vladguc@ipu.rssi.ru

Abstract: We suggest reorganization of the basic strategy of launching vehicles into near-earth orbits,, which is directed to increase safety of operation during the launch.

Keywords: launch vehicle, operation safety, strategy of launching.

*Статья представлена к публикации
членом редакционной коллегии М. В. Губко*

УДК 517.977.58
ББК 73.6

РАСПРЕДЕЛЕНИЕ РЕСУРСОВ ПРОИЗВОДСТВЕННОЙ СИСТЕМЫ С ИСПОЛЬЗОВАНИЕМ СЕТЕЙ ПЕТРИ И ГЕНЕТИЧЕСКОГО АЛГОРИТМА

Сочнев А. Н.¹

(ФГАОУ ВПО «Сибирский федеральный университет»,
Красноярск)

Описывается метод определения оптимального количества ресурсов с использованием имитационной модели на основе сетей Петри, параметры которой определяются генетическим алгоритмом. Приводится пример использования предлагаемого метода для предварительного планирования производства. Формулируются рекомендации по использованию генетического алгоритма для задач распределения ресурсов.

Ключевые слова: сеть Петри, генетический алгоритм, распределение ресурсов.

1. Описание объекта исследования

Объектом исследования является инструментальное производство одного из предприятий. По своему назначению производство относится к вспомогательным (обслуживающим) подразделениям.

Инструментальному производству в области управления изготовлением, ремонтом, эксплуатацией, хранением и выдачей инструмента и оснастки функционально подчиняются инструментально-раздаточные кладовые и группы подготовки производства цехов основного и вспомогательного производства.

Структура инструментального цеха представлена на рис. 1.

¹ Алексей Николаевич Сочнев, кандидат технических наук, доцент (lesek@mail.ru).



Рис. 1. Структура инструментального производства

Основными задачами цеха являются: планирование потребности и производства технологической оснастки и инструмента на предприятии; изготовление специального технологического оборудования; своевременное обеспечение цехов основного и вспомогательного производства технологической оснасткой и инструментом нужного ассортимента и специальным технологическим оборудованием; контроль за правильной эксплуатацией технологической оснастки и инструмента в подразделениях и др.

Функциями инструментального производства в области производственной деятельности являются изготовление технологической оснастки по планам и графикам, утвержденным в установленном порядке; выполнение всех работ в соответствии с чертежами и технологическими процессами; доставка заготовок в цеха; изготовление средств механизации и автоматизации; учет и анализ простоев оборудования на участках инструментального производства; обеспечение экономичного расхода сы-

рья, материалов; снижение производственных затрат на единицу продукции.

Номенклатура изделий, производимых инструментальным цехом в течение планового периода (одной смены), представлена в таблице 1. Требуемое количество изделий каждого типа – 20 единиц.

Таблица 1. Номенклатура инструментального цеха

№	Наименование изделия	Технологические операции							
		заготовительная, мин	слесарная 1, мин	разметочная, мин	фрезерная 1, мин	слесарная 2, мин	шлифовальная, мин	слесарная 3, мин	фрезерная 2, мин
1	Плита нижняя	1	1,3 5	1	14	–	15	1	4
2	Стенка матрицы	1	–	1,2	8	5	16	2	–
3	Плита верхняя	1,7	–	1,2 5	4	2	–	–	–
4	Гайка	1	–	1,5	3	5	7	–	–
5	Пресс-форма 1	2	2,4	3	12	–	15	–	2
6	Пресс-форма 2	1	–	1	10	–	12	3	–
7	Плита	1,2	2	2	–	7	–	1	–
8	Шайба	1	1	1	3	2	–	–	–
9	Пресс-форма 3	2	2,2	2,7	11	–	8	5	–
10	Корпус	1,5	–	2	3	2	3	–	–

2. Имитационная модель производственной системы

Для моделирования дискретных производственных систем широко применяется математический аппарат сетей Петри [3]. Среди основных преимуществ сетей Петри можно отметить то, что они позволяют с требуемой точностью представлять ветвящиеся, параллельные, циклические процессы, обладают средст-

вами анализа, а также моделирования в реальном масштабе времени.

Процесс моделирования можно условно разделить на две стадии.

1. Формирование структуры и параметров модели на основе свойств исходной системы (объекта управления).

2. Имитационные эксперименты с целью достижения требуемого значения показателя качества.

В литературе описано множество различных классов сетей Петри. Имитационная модель инструментального цеха (см. рис. 2) должна отображать временные свойства процессов, поэтому для моделирования используются временные сети Петри [4, 7].

Позиции p_1, p_2, \dots, p_{10} моделируют заготовки десяти типов изделий; позиции $p_{55}, p_{56}, \dots, p_{64}$ – готовые изделия; переходы t_1, t_2, \dots, t_{10} – заготовительные операции; $t_{11}, t_{12}, \dots, t_{14}$ – операции «слесарная 1», $t_{16}, t_{17}, \dots, t_{25}$ – «разметочная» и т.д. Маркеры, моделирующие изделия различных типов, перемещаются в модели от входных позиций к выходным в соответствии с описанием технологических процессов, приведенным в таблице 1.

Моделирование выполняется с использованием такта, равного 0,1 минуты.

Обозначения позиций сети буквами p и m с порядковым номером – это особенность используемой программы *VisObjNet*. Все дуги представленной сети имеют единичный вес. Нули рядом с переходами представляют начальные значения временных задержек, далее они заменяются значениями из таблицы 1.

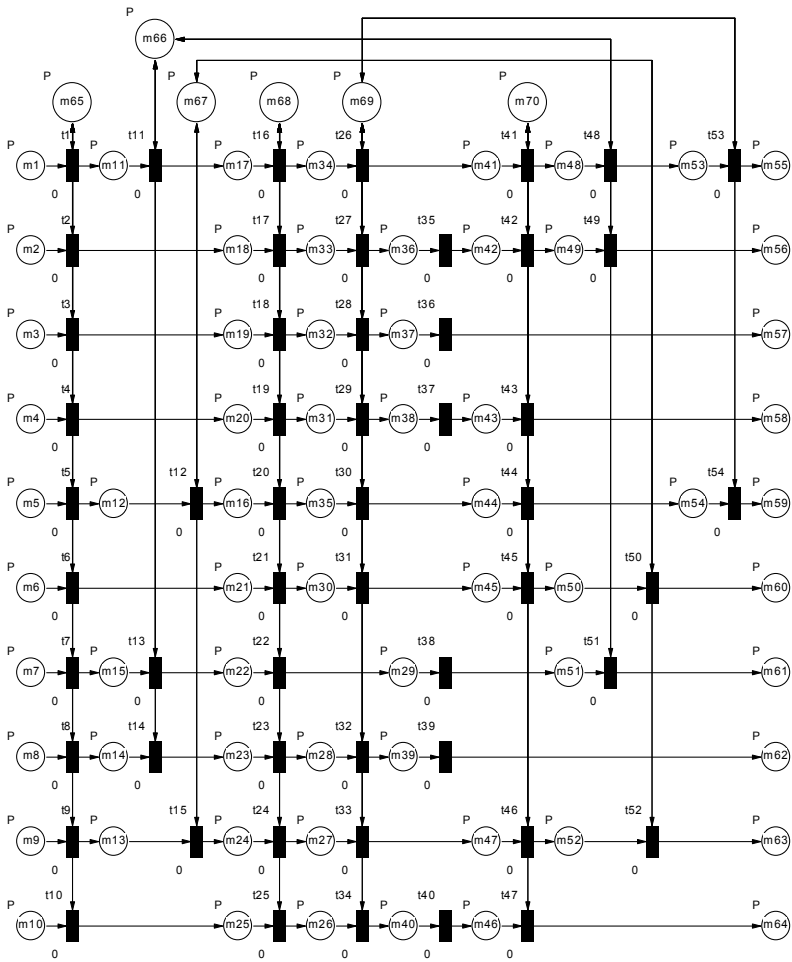


Рис. 2. Сетевая модель производственного процесса

3. Постановка задачи распределения ресурсов

Сформулируем типичную задачу планирования производственного процесса, состоящую в определении необходимого количества исполнителей на каждом участке [9]. Для ее решения требуется в модель ввести элементы, представляющие ресурсы системы (позиции сети $p_{65}, p_{66}, \dots, p_{70}$), см. рис.2.

Из условий задачи, приведенных выше, определяем длительность производственного процесса для одной смены: $T_{max} = 8 \times 60 \times 10 = 4800$ тактов.

Соответственно, задача состоит в том, чтобы определить минимальное количество исполнителей на каждом участке, при котором время процесса не превысит приведенной величины.

В модельном представлении это означает найти начальную маркировку позиций $\mu_0(p_{65}), \mu_0(p_{66}), \dots, \mu_0(p_{70})$. Формально требуется одновременно минимизировать две целевые функции

$$Q_1(\mu_0) = (\mu_0(p_{65}) + \mu_0(p_{66}) + \dots + \mu_0(p_{70})) \rightarrow \min_{\mu_0 \in M_0},$$
$$(1) \quad Q_2(\mu_0) = T_{max} - T(\mu_0) \rightarrow \min_{\mu_0 \in M_0},$$

$$M_0 = \{\mu_0 : T \leq T_{max}; \mu_0 \geq 1; \mu_0 \leq \mu_{0max}; \mu_0 \in N\}.$$

Диапазон изменения маркировок позиций определяется максимально возможным количеством каждого ресурса, см. таблицу 2.

Определение маркировки указанных позиций – процесс итерационный, многократно повторяемый и сопровождаемый имитационными экспериментами. На каждом шаге решения выполняется проверка условия $T \leq T_{max}$.

Практически наиболее сложно определить алгоритм изменения начальной маркировки позиций $p_{65}, p_{66}, \dots, p_{70}$. Этот процесс может быть выполнен следующими поисковыми методами: полного перебора, половинного деления (дихотомии), золотого сечения, случайного поиска и др. В данной статье будет исследован вопрос решения описанной задачи генетическим алгоритмом.

Таблица 2. Диапазон изменения маркировок

Позиция	Минимальное количество маркеров	Максимальное количество маркеров	Максимальное количество (двоичный код)
p_{65}	1	10	1010
p_{66}	1	12	1100
p_{67}	1	4	0100
p_{68}	1	10	1010
p_{69}	1	11	1011
p_{70}	1	7	0111

4. Метод полного перебора маркировок

Полный перебор – метод решения задачи путем перебора всех возможных вариантов. Сложность полного перебора зависит от количества всех возможных решений задачи. Если пространство решений очень велико, то полный перебор может не дать результатов в течение нескольких лет или даже столетий.

Любая задача оптимизации может быть решена полным перебором. При этом даже если вычисление целевой функции от каждого конкретного возможного решения задачи может быть осуществлено за полиномиальное время, в зависимости от количества всех возможных решений полный перебор может потребовать экспоненциального времени работы.

Для рассматриваемого примера использование метода полного перебора потребует проверить 369600 вариантов начальных маркировок сети ($10 \times 12 \times 4 \times 10 \times 11 \times 7$). Если учесть, что один имитационный эксперимент с сетью Петри занимает по времени примерно одну минуту, то общее время решения 369600 минут или 256 суток является совершенно недопустимой величиной.

В действительности это время можно уменьшить, исключая из рассмотрения заведомо неудачные варианты решения.

Дальнейшее уменьшение начальной маркировки приводит к нарушению условия $T \leq T_{max}$. Очевидный недостаток рассмотренного метода – большое количество шагов решения. В таблице 3

приведены только некоторые наиболее значимые из них, влияющие на величину T .

Таблица 3. Результаты поиска начальной маркировки методом полного перебора

Вектор маркировки $\mu_0 = (\mu_0(p_{65}) \ \mu_0(p_{66}) \ \dots \ \mu_0(p_{70}))$	Продолжительность производственного процесса T
$\mu_0 = (10 \ 12 \ 4 \ 10 \ 11 \ 7)$	3372
$\mu_0 = (5 \ 6 \ 1 \ 5 \ 5 \ 5)$	3760
$\mu_0 = (5 \ 6 \ 1 \ 5 \ 4 \ 5)$	4552
$\mu_0 = (2 \ 4 \ 1 \ 2 \ 4 \ 5)$	4552
$\mu_0 = (1 \ 4 \ 1 \ 2 \ 4 \ 5)$	4752
...	...
$\mu_0 = (1 \ 2 \ 1 \ 2 \ 4 \ 5)$	4792

5. Генетический алгоритм

Вопросы совместного использования генетического алгоритма и сетей Петри рассмотрены в работах [2, 6]. Они посвящены методам решения задачи структурного синтеза модели объекта с заданными свойствами. Наличие подобных публикаций подтверждает актуальность исследования вопросов совместного применения сетей Петри и генетических алгоритмов.

Основные преимущества генетических алгоритмов [1].

1. Они не требуют никакой информации о поверхности ответа.
2. Разрывы, существующие на поверхности ответа, имеют незначительный эффект на полную эффективность оптимизации.
3. Они стойки к попаданию в локальные оптимумы.
4. Они хорошо работают при решении крупномасштабных проблем оптимизации.
5. Могут быть использованы для широкого класса задач.
6. Могут быть использованы в задачах с изменяющейся средой.

Необходимость поиска всех решений не является помехой. Существует по крайней мере три класса задач, которые могут быть решены генетическим алгоритмом [8]:

- задача быстрой локализации одного оптимального значения;
- задача определения нескольких (или всех) глобальных экстремумов;
- задача описания ландшафта исследуемой функции, которая может сопровождаться выделением не только глобальных, но и локальных максимумов.

Задача распределения ресурсов, сформулированная выше, может быть отнесена к первому из приведенных классов задач.

Решение многокритериальной задачи оптимизации, сформулированной выше, предполагается выполнить стандартным образом: сведением двух критериев к одному путем введения априорных (экспертных) весовых коэффициентов для каждого из них.

Полученная задача оптимизации формулируется следующим образом. Целевая функция (функция приспособленности)

$$(2) \quad G(\mu_0) = k_1 \cdot Q_1(\mu_0) + k_2 \cdot Q_2(\mu_0) \rightarrow \min_{\mu_0 \in M_0}$$

где $M_0 = \{\mu_0 : T \leq T_{max}; \mu_0 \geq 1; \mu_0 \leq \mu_{0max}; \mu_0 \in N\}$.

Экспертные оценки коэффициентов целевой функции приняты одинаковыми: $k_1 = 1$ и $k_2 = 1$.

Поскольку предполагаемые значения функций $Q_1(\mu_0)$ и $Q_2(\mu_0)$ различаются в сотни раз, то для правильной оценки значения функции $G(\mu_0)$ их необходимо привести в нормированный или процентный вид.

Генетический алгоритм, как правило, требует в классической реализации определения переменных в двоичной форме. В рассматриваемом примере диапазоны изменения всех переменных определяются четырьмя двоичными разрядами. Нижняя и верхняя границы диапазонов изменения каждой переменной представлены в таблице 2.

Особенность рассматриваемой задачи также в том, что экстремальное значение целевой функции требуется найти за минимальное количество шагов. Это связано с тем, что для каждой особи в популяции требуется выполнить имитационный эксперимент. В сети Петри, представленной выше, процесс имитации длится при-

мерно минуту (для выбранной постоянной времени). Для более сложных моделей имитация может длиться до нескольких часов. Сравнительно высокая скорость решения задачи является основной причиной выбора генетического алгоритма.

Первый шаг генетического алгоритма – это формирование начальной популяции. Для решаемой задачи количество особей в начальной популяции должно быть минимальным. При этом необходимо учитывать, что в популяциях с малым размером гены распространяются слишком быстро. Все особи становятся похожими (популяция сошлась) еще до того, как найдено решение задачи. Получается так, что новый генотип с лучшей оценкой быстро вытесняет менее хорошие комбинации ген, исключая таким образом возможность получения более хорошего решения на их базе.

Вторая проблема заключается в том, что потомки, получаемые в большей популяции, скорее всего будут разнообразнее, чем потомки из меньшей популяции. С одной стороны, при малых популяциях внимание алгоритма будет сконцентрировано на некоторых направлениях, т.е. эти направления будут изучены быстрее. Но если направление изначально было ошибочным, то алгоритм остановится вблизи ложного оптимума. Выходом может быть увеличение популяции, при условии что время работы алгоритма не критично.

Также может помочь самоадаптация алгоритма при выборе операторов (в тупиковых ситуациях – увеличить вероятность применения случайного генератора особей).

Еще можно попытаться складировать отдельно особи, генотип которых был утерян популяцией, и временами добавлять эти особи (случай удачного кровосмешения в племенных формациях).

Далее по алгоритму оценивается приспособленность каждой особи, производится отбор и скрещивание. В каждой новой популяции применяется механизм мутации. Применительно к решаемой задаче скрещивание и мутирование имеет следующие особенности.

1. В результате скрещивания могут появляться особи, у которых некоторые признаки равны нулю. Поскольку это противоречит условиям задачи и особенностям системы, то они заменяются на ненулевые значения (например на единицу).

2. На каждом шаге некоторые особи мутируют, но неслучайным образом, а с использованием экспертного опыта. Эксперт участвует в процессе эволюции, отслеживает важные и неважные признаки с точки зрения величины их влияния на целевую функцию и меняет их в сторону предполагаемых оптимальных значений. Опыт показывает, что это значительно ускоряет скорость работы алгоритма и, соответственно, минимизирует количество экспериментов с имитационной моделью, а также вносит необходимое разнообразие в небольшую по объему популяцию.

Далее представлен процесс решения задачи в программе *Microsoft Excel* в виде последовательности шагов.

1. Формируются шесть векторов начальной маркировки позиций $p_{65}, p_{66}, \dots, p_{70}$, см. таблицу 4. Минимальное возможное количество особей в популяции определено эмпирически. При большем количестве особей генетический алгоритм проигрывает в скорости поиска оптимума методу половинного деления или золотого сечения. При меньшем количестве процесс может сойтись к ложному оптимуму.

Таблица 4. Первая популяция

Начальная маркировка (<i>chn</i> – хромосома <i>n</i>)						
	<i>ch</i> 1	<i>ch</i> 2	<i>ch</i> 3	<i>ch</i> 4	<i>ch</i> 5	<i>ch</i> 6
$\mu_0(p_{65})$	4	8	3	3	3	8
$\mu_0(p_{66})$	7	11	2	5	3	5
$\mu_0(p_{67})$	4	3	3	2	2	2
$\mu_0(p_{68})$	4	3	7	9	4	5
$\mu_0(p_{69})$	6	6	3	5	7	6
$\mu_0(p_{70})$	5	3	4	7	4	4

2. Сформированные векторы представляются в двоичной форме, см. таблицу 5.

3. Проводятся шесть имитационных экспериментов на модели и определяются значения целевых функций $Q_1(\mu_0)$, $Q_2(\mu_0)$ и $G(\mu_0)$. Полученные значения $Q_1(\mu_0)$ и $Q_2(\mu_0)$ нормируются, и $G(\mu_0)$ определяется от нормированных значений, см. таблицу 6.

Таблица 5. Двоичное представление маркировки

0110	1101	0101	0101	0100	1100
1001	1111	0010	0110	0100	0110
1111	1011	1011	1000	0110	0111
0110	0100	1011	1110	0110	1000
1000	1001	0101	0111	1010	1001
1011	0110	1000	1111	1010	1000

Таблица 6. Результаты экспериментов (* – нормированное значение функции)

$Q_1(\mu_0)$	1385	675	964	1040	24	25
$Q_1^*(\mu_0)$	0,34	0,16	0,23	0,25	0,01	0,01
$Q_2(\mu_0)$	28	33	21	30	24	28
$Q_2^*(\mu_0)$	0,17	0,20	0,13	0,18	0,13	0,17
$G^*(\mu_0)$	0,51	0,37	0,36	0,44	0,14	0,18

4. На основе полученных значений функции приспособленности $G(\mu_0)$ выбираются особи для скрещивания по принципу рулетки и формируется новая популяция, см. таблицы 7, 8.

Таблица 7. Скрещиваемые особи

$ch\ 6 \leftrightarrow ch\ 4$	$ch\ 3 \leftrightarrow ch\ 5$	$ch\ 1 \leftrightarrow ch\ 2$
-------------------------------	-------------------------------	-------------------------------

Таблица 8. Второе поколение и мутировавшие особи

Начальная маркировка ($ch\ n$ – хромосома n)						
	ch1	ch2	ch3	ch4	ch5	ch6
$\mu_0(p_{65})$	0110	0100	1101	0100	1100	0101
$\mu_0(p_{66})$	0101	1000	0110	0010	0110	0111
$\mu_0(p_{67})$	0111	1110	0111	1011	1011	0111
$\mu_0(p_{68})$	0110	0110	1011	1000	0100	1000
$\mu_0(p_{69})$	1000	1010	1001	0101	1001	1001
$\mu_0(p_{70})$	1011	1010	1000	1000	1100	1010

В новой популяции выполняются мутации. Как указывалось выше, мутации производятся экспертом с теми генами, которые с точки зрения процесса оптимизации заведомо неудачные. Напри-

мер, до мутации у пятой особи второй ген образовывала комбинация 1110_2 или 14_{10} (11_{10} с учетом ограничений количества маркеров) и имелись две важные причины уменьшить значение этого гена. Во-первых, эксперимент с предыдущими особями, в которых этот ген гораздо меньше, а значение функции приспособленности получилось вполне удовлетворительным. Во-вторых, вторая ячейка производственной системы по условиям задачи является недогруженной, поэтому использование большого количества станков представляется неправильным.

Процесс решения задачи состоит в повторении пунктов 1–4 до достижения желаемых значений целевых функций $Q_1(\mu_0)$, $Q_2(\mu_0)$ и $G(\mu_0)$.

В рассматриваемом примере приемлемое решение задачи появляется уже в третьем поколении, см. таблицы 9, 10.

Таблица 9. Третье поколение популяции

Начальная маркировка (<i>ch n</i> – хромосома <i>n</i>)						
	<i>ch 1</i>	<i>ch 2</i>	<i>ch 3</i>	<i>ch 4</i>	<i>ch 5</i>	<i>ch 6</i>
$\mu_0(p_{65})$	3	3	3	3	3	3
$\mu_0(p_{66})$	2	3	5	1	5	5
$\mu_0(p_{67})$	4	2	4	3	2	2
$\mu_0(p_{68})$	3	6	3	5	7	5
$\mu_0(p_{69})$	6	7	6	4	6	6
$\mu_0(p_{70})$	4	5	4	4	4	4

Таблица 10. Функции приспособленности третьего поколения (* – нормированное значение функции)

24	1245	24	1500	24	24
0,01	0,32	0,01	0,39	0,01	0,01
22	25	23	19	26	26
0,14	0,16	0,15	0,12	0,17	0,17
0,15	0,48	0,15	0,51	0,17	0,17

Процесс оптимизации может быть продолжен с использованием генетического алгоритма или иных методов, например метода

половинного деления, покоординатного спуска или подобных [5]. В описываемом примере оптимальное значение уточняется половинным делением и координатным спуском, см. таблицу 11. Поскольку длительность производственного процесса по условию задачи не должна превышать 4800 тактов, то первый шаг делается для того, чтобы вернуть ее в область допустимых значений (4824 → 3761).

Таблица 11. Поиск оптимального решения

$\mu_0(p_{65})$	3	3	2	2	1	1	1
$\mu_0(p_{66})$	2	2	1	2	2	2	2
$\mu_0(p_{67})$	4	4	3	3	3	1	1
$\mu_0(p_{68})$	3	3	2	2	2	2	2
$\mu_0(p_{69})$	6	6	5	5	5	5	4
$\mu_0(p_{70})$	4	5	5	5	5	5	5
T	4824	3761	6300	3952	4384	4384	4792
$Q_1(\mu_0)$	24	1039	1500	848	416	416	8
$Q_2(\mu_0)$	22	23	18	19	18	16	15

Описанный процесс решения задачи может быть представлен в виде блок-схемы обобщенного алгоритма, см. рис. 3.

6. Заключение

Основные научные результаты, полученные при работе над темой статьи, можно определить следующими пунктами.

1. Сформулирована многокритериальная задача оптимального распределения ресурсов, имеющая практическое применение в системах управления производством.

2. Сформулированная задача формализована с использованием терминов теории сетей Петри.

3. Получена сетевая модель системы, в структуре которой отдельно выделены позиции, описывающие количество ресурсов.

4. Определены основные особенности генетического алгоритма в применении его к задаче поиска оптимального количества ресурсов на рассмотренном инструментальном производстве.



Рис.3. Обобщенный алгоритм решения задачи

Литература

1. ГЛАДКОВ Л.А., КУРЕЙЧИК В.В., КУРЕЙЧИК В.М. *Генетические алгоритмы: Учебное пособие* – 2-е изд. М: Физматлит, 2006. – С. 320.
2. ЕЛЬЧАНИНОВ Д.Б., САМИМЕХАНА, ПЕТРОСОВ Д.А. *Применение генетических алгоритмов при проектировании компьютерной техники* / Вестник Херсонского государственного университета. – Херсон: Изд-во ХГУ, 2003. – № 2(18). – С. 146–148.
3. ЕМЕЛЬЯНОВ В.В., ГОРНЕВВ.Ф., ЕМЕЛЬЯНОВВ.В., ОВСЯННИКОВ М.В. *Оперативное управление в ГПС*. – М.: Машиностроение, 1990. – 256 с.
4. КОТОВ В.Е. *Сети Петри*. – М.: Наука, 1984. – 160 с.

5. КУРЕЙЧИК В.М., ЛЕБЕДЕВЬ.К., ЛЕБЕДЕВ О.К. *Поисковая адаптация: теория и практика.* – М: Физматлит, 2006. – 272 с.
6. ПЕТРОСОВ Д.А., ЛОБОДАВ.Г., ЕЛЬЧАНИНОВ Д.Б. *Представление генетических алгоритмов Сетями Петри в задачах проектирования компьютерной техники /* Материалы научно-практической конференции «Информационные технологии в науку и образование». – Х. : ХНУРЭ, 2005. – С. 48–51.
7. ПИТЕРСОН ДЖ. *Теория сетей Петри и моделирование систем.* – М.: Мир, 1984. – 264 с.
8. РУТКОВСКАЯ Д., ПИЛИНСКИЙ М., РУТКОВСКИЙ Л. *Нейронные сети, генетические алгоритмы и нечеткие системы* – 2-е изд. – М: Горячая линия-Телеком, 2008. – 452 с.
9. ЮДИЦКИЙ С.А., МАГЕРГУТ В.З. *Логическое управление дискретными процессами. Модели, анализ, синтез.* – М.: Машиностроение, 1987. – 176 с.

RESOURCE ALLOCATION IN PRODUCTION SYSTEM USING PETRI NETS AND GENETIC ALGORITHM

Alexey Sochnev, Siberian Federal University, senior lecturer at department «Robotics and Technical Cybernetics» (lesek@mail.ru)

Abstract: The method is described for determining the optimal volume of resources using a simulation model based on Petri nets, whose parameters are determined by the genetic algorithm. The example is provided of using the proposed method for pre-production planning. Recommendations are given on the use of genetic algorithms for the problems of resource allocation.

Keywords: Petri net, genetic algorithm, resource allocation.

*Статья представлена к публикации
членом редакционной коллегии А. К. Погодаевым*

УДК 004.896
ББК 30.604

МУЛЬТИАГЕНТНАЯ СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПО ОПЕРАТИВНОМУ ПЛАНИРОВАНИЮ И ТЕХНОЛОГИЧЕСКОЙ КООРДИНАЦИИ СЛОЖНОСТРУКТУРИРОВАННЫХ ПРОИЗВОДСТВЕННЫХ СИСТЕМ

Цуканов М. А.¹, Боева Л. М.²

*(Старооскольский технологический институт (филиал)
ФГОУ ВПО «Национальный исследовательский техноло-
гический университет «МИСиС» (СТИ НИТУ МИСиС),
Старый Оскол)*

Обоснована целесообразность и перспективность разработки СППР оперативного управления и технологической координации сложноструктурированных распределенных производственных систем на основе мультиагентных технологий. Представлены модели и алгоритмы оптимизации оперативного планирования с использованием имитационного моделирования и методов искусственного интеллекта.

Ключевые слова: мультиагентная система, агенты, иммунный алгоритм, сети Петри, агрегат Бусленко.

1. Введение

Основной целью оперативного планирования производства является составление согласованных планов цехов предприятия и обеспечение их выполнения. Оперативное управление выполнением производственных планов осуществляется на основе

¹ Михаил Александрович Цуканов, ассистент (tsukanov_m_a@mail.ru)

² Людмила Михайловна Боева, кандидат технических наук, доцент (boeva@inbox.ru)

технологической координации, заключающейся в согласовании работы технологического оборудования и транспорта, движения материальных потоков, взаимодействия производственного персонала цеха при отклонении фактического хода производства от запланированного [1].

Задача существенно усложняется для производств, характеризующихся широкой номенклатурой выпускаемой продукции, многообразием оборудования, многостадийностью технологических процессов, многовариантностью технологических маршрутов и, как следствие, сложными перекрестными материально-транспортными связями.

Помимо перечисленных трудностей задачу оперативного планирования таких производств усложняет необходимость составления и корректировки производственных графиков в темпе производства.

2. Мультиагентная система оперативного управления и технологической координации

2.1. ПРОИЗВОДСТВЕННОЕ РАСПИСАНИЕ КАК ОСНОВА ОПЕРАТИВНОГО УПРАВЛЕНИЯ

Принятой формой разработки оперативных планов на предприятиях является календарное планирование, в основе которого лежит построение месячных графиков производства готовой продукции, детализация их в недельные планы для цехов и дальнейшее их разбиение на сменно-суточные задания (СЗЗ) или графики.

Инструментом текущей реализации СЗЗ и одновременно основой ТК является контактный график (КГ) – производственное расписание, регламентирующее работу основного технологического оборудования по горизонтали и вертикали [2].

Основная проблема ТК заключается в необходимости оперативного принятия решений по вопросам перестроения КГ в связи с рассогласованием текущей производственной ситуации относительно запланированной в темпе производства.

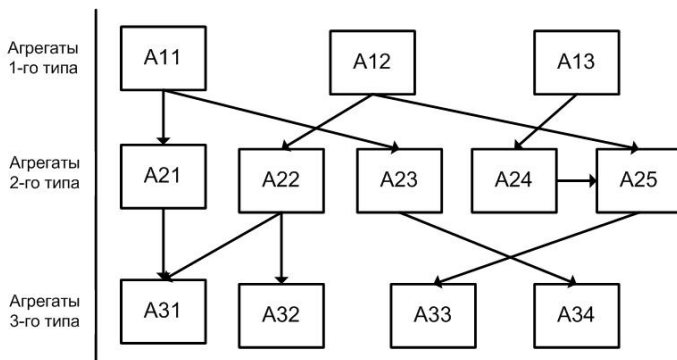


Рис. 1. Структура связей сложноструктурированного производства

Для решения этой проблемы предлагается система поддержки принятия решений (СППР) на основе мультиагентных технологий (МАТ), которая базируется на комплексе моделей и алгоритмов оптимизации производственных планов в режиме *on-line* (рис. 2).

Блок принятия решений концептуальной схемы МАСППР (рис. 2) представлен агентом-оптимизатором и агентом-супервизором, блок анализа проблем – агентом-реализатором, имитационная модель производства – агентами-исполнителями.

Агент-супервизор – интерфейсный агент, решающий задачу взаимодействия агентов МАС и связи с пользователем. Он выдает плановый КГ, отчет по анализу «узких мест» КГ и варианты его корректировки, формирует задания нижестоящим агентам МАС на обработку производственных заказов в соответствии с принятым КГ.

Агент-оптимизатор – гибридный агент, в задачу которого входит построение оптимального КГ на основе правил и ограничений производства.

Агент-реализатор – гибридный агент, который осуществляет проверку сформированного КГ на реализуемость.

Агент-исполнитель – партнерский агент, задачей которого является слежение за работой конкретного технологического агрегата, входящего в реализуемый технологический маршрут.

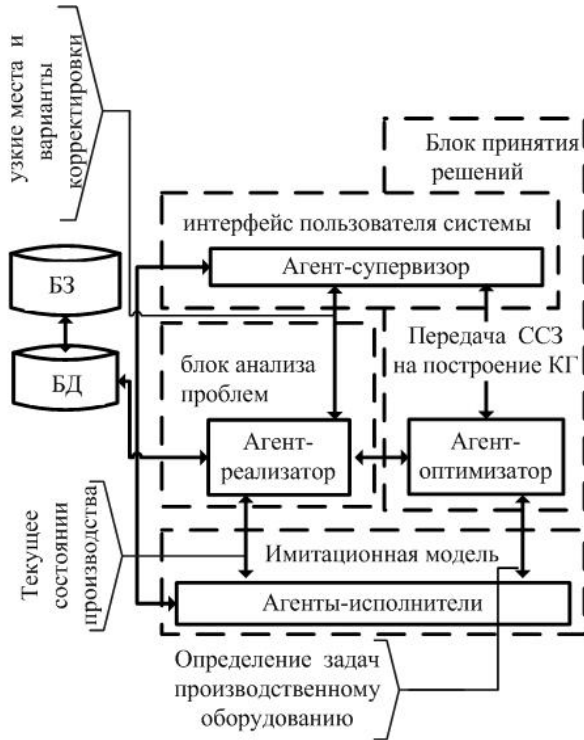


Рис. 2. Структура СППР на основе МАТ

Супервизор на основе информации о готовности агрегатов и их занятости в технологических маршрутах плана корректирует производственную программу и направляет откорректированный КГ агенту-реализатору для проверки возможности его выполнения.

2.2. ОПТИМИЗАЦИЯ ПРОИЗВОДСТВЕННОГО РАСПИСАНИЯ

Задача оптимизации производственного графика цеха относится к классу задач составления расписаний. Наличие нескольких однотипных агрегатов, многовариантность технологических маршрутов, последовательно-параллельные и перекрестные транспортно-технологические потоки определяют эту задачу как

NP-сложную. Время ее решения с использованием комбинаторных и эвристических методов оптимизации и известных методов искусственного интеллекта (генетического алгоритма, алгоритма муравьиных колоний) не удовлетворяет требованиям оперативного, в темпе производства, управления сложноструктурированными производствами.

Для оптимизации контактного графика и внесения в него необходимых корректировок в режиме *on-line* предлагается математический аппарат одной из модификаций алгоритма иммунных сетей (рис. 4), обеспечивающего эффективный параллельный поиск оптимального решения на основе принципа клональной селекции, предложенного Де Кастро [5].

Оптимизация расписания осуществляется путем выбора последовательности технологических маршрутов производства продукции согласно ССЗ, минимизирующей суммарные приведенные потери R_{ij} [3], связанные с переналадкой всех технологических агрегатов при переходе от обработки i -го заказа к $(i + 1)$ -му:

$$(1) \quad F = \sum_i \sum_j R_{ij} \rightarrow \min$$

$$(2) \quad R_{ij} = C_{cp} \cdot \rho_{ij} \cdot \Delta D_{ij},$$

где C_{cp} – средняя себестоимость продукции за смену; ρ_{ij} – производительность j -го технологического агрегата по выпуску i -го заказа; ΔD_{ij} – длительность переналадки j -го технологического агрегата при переходе от обработки i -го заказа к $(i + 1)$ -му.

2.3. ПРОВЕРКА ПРОИЗВОДСТВЕННОГО РАСПИСАНИЯ НА ВОЗМОЖНОСТЬ РЕАЛИЗАЦИИ

Алгоритм проверки составленного КГ на реализуемость осуществляется с использованием математического аппарата вложенных сетей Петри, в которой каждая позиция-вершина системной сети представлена как группа оборудования, соответствующая технологическому маршруту КГ [3].

Сеть описывается формально множествами переходов и вершин сети. Функционирование сети задается правилами срабатывания переходов.

Математически сеть описывается коротжем

$$(3) S = \langle P, T, F, \tau_T, C, \{V_S\}, K, M_0 \rangle,$$

где P – множество позиций, представленных моделями отдельных единиц технологического оборудования; T – множество переходов между смежными агрегатами; F – функция инцидентностей позиций и переходов, определяющая для рассматриваемого требования к обслуживанию агрегат-источник и агрегат-исполнитель; C – функция цвета маркера, сигнализирующая о принадлежности перехода к определенному технологическому маршруту; τ_T – модельное время, отнесенное ко всем компонентам сети $P, T, F, M_0; \{V_S\}$ – условия выполнения переходов, отнесенных к компонентам сети, входным и выходным позициям; K – емкость маркеров в позициях с учетом C ; M_0 – вектор начальной маркировки, компоненты которого помечают закрытые позиции при поступлении требования на обслуживание.

Срабатывание каждого перехода из множества $T \{t_1, t_2, \dots, t_{11}\}$ определяется наличием сигнала на выходе определенной технологической установки. Возможность осуществления перехода в одну из позиций P определяется с учетом значений параметров сети $F(A_{lk}, A_{(l+1)(k+1)})$, идентифицирующих агрегат-исполнитель следующего требования, и вектора M_0 , компоненты которого помечают закрытые позиции при поступлении требования на обслуживание. Аргументы функции $A_{lk}, A_{(l+1)(k+1)}$ представляют соответственно агрегат-источник и агрегат – приемник требования на обслуживание. Вектор M_0 характеризуется переменной размерностью, которая зависит от этапа обработки и определяет общее число агрегатов-приемников технологического требования. Закрытые позиции помечаются как 0, допустимые как 1.

Объединённые в группы агрегаты имеют соответствующее входное и выходное условие работы (переход), что представлено на сети множеством стрелок. Согласно этим условиям проверяется занятость агрегата на момент поступления требования на обслуживание. В случае успешной проверки, т.е., когда агрегат-

приемник свободен ($p_{lk} = 0$), на время D_Φ ему присваивается значение 1.

В случае если КГ признан невыполнимым, на основе недоступных позиций или неосуществимых переходов базой знаний формируется набор рекомендуемых корректировок с целью устранения узких мест при перепланировке КГ агентом-оптимизатором.

2.3. МОДЕЛИРОВАНИЕ ПРОИЗВОДСТВА

Для моделирования отдельных технологических и транспортных единиц оборудования и звеньев производства предлагается агрегативная модель Н.П. Бусленко [4]. Агрегат A_j (рис. 6) характеризуется набором координат x_m , $l = 1, \dots, M$, которые описывают его состояние: для основных агрегатов – простой, ожидание продукта, операция обработки и передачи, операция ожидания; для агрегатов-накопителей – простой, ожидание продуктов до обработки; для агрегатов – транспортных средств – простой, операция транспортировки, операция ожидания; $z^{(1)}, \dots, z^{(n)}$ – управляющие сигналы; $y_{l-1,k}$ – вход агрегата, поступающий с выхода предыдущего агрегата; $y_{l,k}$ – выход текущего агрегата и вход следующего.

Агрегат реализует алгоритм выходов G_n (окончание обработки на одном агрегате и передача другому) и алгоритм переходов H_n (изменение состояния агрегата в процессе работы). Параметры агрегата β_n характеризуют его работоспособность.

Агрегативные модели оборудования реализуются на нижнем уровне МАС агентами-исполнителями.

Связная агрегативная модель всего технологического процесса цеха представляется партнерской агентной системой (рис. 7).

Ее работа направлена на отслеживание этапов выполнения КГ. Последовательно анализируется возможность каждого из назначенных в КГ агрегатов обработать технологическое задание. В случае возникновения рассогласования планового и фактического расписания агентом-реализатором формируется запрос агенту-оптимизатору на построение нового КГ.

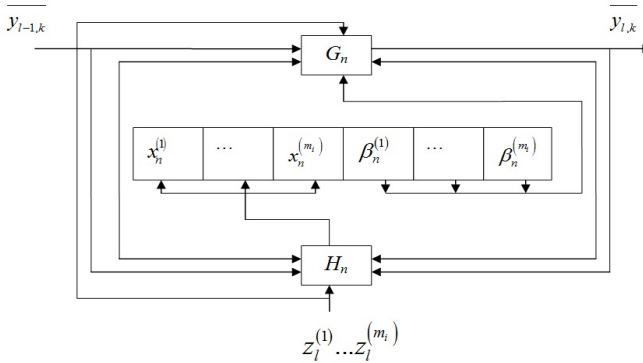


Рис. 6. Схема агрегата Бусленко

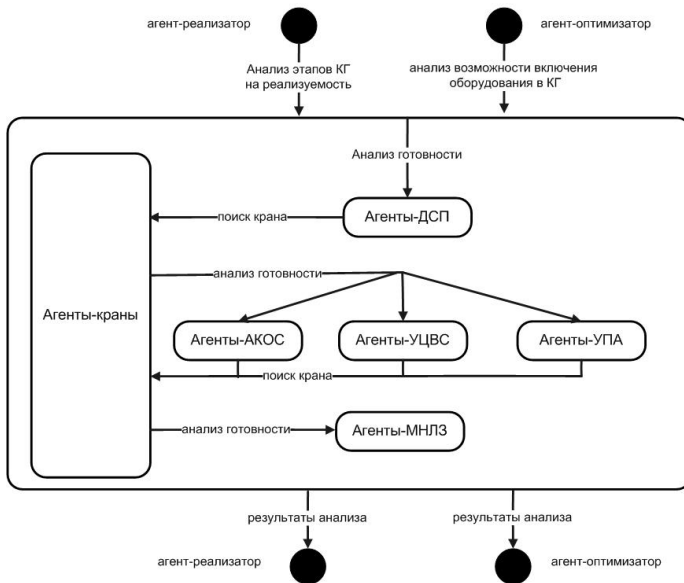


Рис. 7. Структурная схема взаимодействия партнерских агентов (на примере сталеплавильного производства)

2.4. КООРДИНАЦИЯ АГЕНТОВ МАС

Координация агентов в системе осуществляется на основе непрямого взаимодействия, которое соответствует распределению функциональных задач МАС. Агент-супервизор формирует технологическое задание по обслуживанию заказов с учетом изменившейся производственной ситуации, которое в виде входного сообщения поступает агенту-оптимизатору, идентифицирующему состояние производства и определяющему группы агентов-исполнителей. Степень готовности агентов-исполнителей принять задание анализируется агентом-реализатором, который формирует и посылает агенту-супервизору сообщение о возможности выполнения заданий каждым из членов рабочей группы.

3. Заключение

На основе разработанных моделей и алгоритмов возможна реализация интеллектуальной МАС ППР по ОУ и ТК производства, которая отслеживает текущее состояние производства и обеспечивает эффективное выполнение производственной программы.

Литература

1. ЛИТВИНЦЕВ П.И. *Методы организации вычислений в диалоговых системах планирования*: дис. канд. техн. наук. – М.: Выч. центр АН СССР, 1981. – 190 с.
2. ВЕРЕВКИН С.В. *Формирование контактного графика в параллельно-последовательных системах* // Информационные технологии в экономике, промышленности и образовании: Сб. науч. тр. – Кемерово. Изд-во НФИ КемГУ, 2000. – С. 18–24.
3. ЦУКАНОВ М.А., БОЕВА Л.М. *Моделирование технологической координации оборудования сталеплавильного цеха на основе аппарата вложенных сетей Петри* // Электротехнические комплексы и системы управления. – 2010. – № 2(18). – С. 51–54.
4. БУСЛЕНКО Н.П. *Моделирование сложных систем*. М.: Наука, 1968. – 355 с.

5. DE CASTRO L.N., VON ZUBEN F.J. *The Clonal Selection Algorithm with Engineering Applications* // Proc. GECCO'00, 2000. (http://www.dca.fee.unicamp.br/~vonzuben/research/lnunes_dout/artigos/gecco00.pdf)

**MULTIAGENT SYSTEM OF DECISION SUPPORT FOR
OPERATIONAL CONTROL AND TECHNOLOGICAL
COORDINATION OF DISTRIBUTED MANUFACTURE
WITH COMPLEX STRUCTURE**

Michael Tsoukanov, Oskol institute of technology branch of the “National University of Science and Technology “MISiS”, Stary Oskol, postgraduate (tsukanov_m_a@mail.ru).

Ludmila Boeva, Oskol institute of technology branch of the “National University of Science and Technology “MISiS”, Stary Oskol, Cand.Sc, assistant professor (boeva@inbox.ru).

Abstract: We apply multi-agent approach to design decision-support systems of operational control and technological coordination for distributed manufactures with complex structure. We develop the models and suggest optimization algorithms for operational planning, based on simulation and artificial intelligence methods.

Keywords: multi-agents system, agent, immune algorithm, Petri network, Buslenko unit.

*Статья представлена к публикации
членом редакционной коллегии Н. Н. Бахтадзе*

УДК 681.5
ББК 32.817

СЦЕПЛЕНИЕ КООРДИНАТ И ИЕРАРХИЧЕСКИЕ АЛГОРИТМЫ В ЗАДАЧЕ РАВНОУДАЛЕННОГО РАСПОЛОЖЕНИЯ АГЕНТОВ НА ОТРЕЗКЕ¹

Парсегов С. Э.²

(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Предлагаются и изучаются новые стратегии равноудаленного расположения агентов на отрезке. Модификации стратегии включают в себя повышение порядка агентов, образующих систему, и сцепление координат агентов. Кроме того, рассматривается двухуровневая иерархическая схема, позволяющая управлять группами агентов.

Ключевые слова: управление формациями, сцепление координат, иерархические алгоритмы.

Введение

Теория управления мультиагентными системами является дальнейшим развитием классической теории управления. Особенностью мультиагентных систем является наличие не одного объекта управления, как в классической теории, а целой группы объектов, которые называют агентами. Теоретический аппарат этого нового направления базируется как на классических принципах и методах теории управления [8, 13], так и на положениях спектральной теории графов, необходимой для исследования структуры мультиагентных систем [1, 9, 15].

Особенностью мультиагентных систем является их децентрализованность — управление всей системой осуществляется

¹ Автор признателен П.С. Щербакову за ценные обсуждения содержания статьи.

² Сергей Эрнестович Парсегов, аспирант (s.e.parsegov@gmail.com).

без центрального регулятора, действующего на каждого агента. Достижение цели управления достигается за счет локального взаимодействия агентов. Платой за отсутствие централизованного управления является более высокая сложность задач управления мультиагентными системами. Однако по ряду причин применение такого подхода является более оправданным и многообещающим. Прежде всего — это наличие реальных физических ограничений, действующих на агентов, таких как пропускная способность каналов связи, ограниченная зона видимости, большое число взаимодействующих подсистем. К преимуществам распределенного подхода можно отнести робастность к отказу отдельных подсистем, надежность и гибкость [21].

Адекватность моделей мультиагентных систем, и, соответственно, их сложность определяется многими факторами: сложностью моделей агентов, входящих в систему, структурой, связывающей агенты между собой, типами этих связей [25].

Насколько известно автору, на настоящий момент не очень много работ, где были бы структурированы и отражены все аспекты этого направления, литература по данному вопросу весьма обширна и разнородна, а терминология часто варьируется. В качестве примера следует привести [21], где сделана попытка обобщить и изложить имеющиеся результаты по управлению мультиагентными системами. Можно выделить основные направления/задачи в области мультиагентных систем: задачи достижения консенсуса, распределенной оптимизации, распределенного оценивания, управления формациями и многие другие.

Вопросы, связанные с управлением формациями, занимают особое место в этой тематике. Распределенное управление формациями подразумевает определенное поведение агентов, при котором они образуют некоторые геометрические конфигурации в пространстве посредством локального взаимодействия. Алгоритмы такого рода применяются для управления группами колесных роботов, беспилотных летательных/подводных аппаратов [2] и т.д. Управление системами, состоящими из большого числа мобильных объектов, является важной задачей из-за их ком-

мерческого и военного применения. В сравнении с автономными устройствами, выполняющими одиночные миссии, применение групп совместно работающих устройств открывает большие возможности. Можно привести лишь некоторые примеры потенциальных приложений таких групп автономных устройств: военное дело, системы наблюдения, системы метеослужб, погрузочно-разгрузочные операции опасных материалов, распределенные сети сенсоров и т.п. Управление формациями включает в себя такие задачи, как *получение формации (formation producing)* и *слежение формации (formation tracking)* [21, 23]. В рамках задачи получения/образования формации группа агентов в пространстве образует некоторую геометрическую фигуру без какого-либо указания для группы (без лидера, виртуального или реального); в задаче слежения группа агентов достигает желаемого геометрического расположения и следит за задающим воздействием (следует за лидером).

Одной из типовых задач в этой области является задача равноудаленного расположения агентов на некотором фиксированном отрезке посредством локального взаимодействия [7, 10, 24]. Стратегия управления устроена следующим образом: каждому агенту в группе предписывается расположиться в середине отрезка, соединяющего его двух крайних по номерам соседей. Существующие алгоритмы равноудаленного расположения агентов на отрезке описываются дифференциальными уравнениями первого порядка, не учитывают возможную связь между пространственными координатами. Особенность рассматриваемой стратегии состоит в том, что она не является задачей достижения консенсуса.

В работе предлагаются различные обобщения задачи равноудаленного расположения на отрезке: усложняются модели агентов, связи между ними, предлагается и изучается двухуровневый иерархический алгоритм.

1. Основные понятия и постановка задачи

1.1. РАВНОУДАЛЕННОЕ РАСПОЛОЖЕНИЕ НА ОТРЕЗКЕ

Одномерный случай

Рассмотрим линейный алгоритм перемещения n подвижных агентов для их равномерного расположения на заданном отрезке $[x_0, x_{n+1}]$. Пусть положение каждого агента при $t \geq 0$ обозначается через $x_i(t) \in \mathbb{R}$. Закон управления предполагает наличие информации о расстояниях между агентом и двумя его ближайшими по номерам соседями. Каждый агент движется в направлении середины отрезка, соединяющего его соседей. Векторно-матричное описание стратегии имеет вид [24]:

$$(1) \quad \begin{bmatrix} \dot{x}_0 \\ \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \\ \dot{x}_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ 0.5 & -1 & 0.5 & 0 & \dots & 0 \\ 0 & 0.5 & -1 & 0.5 & \dots & 0 \\ \vdots & & & & & \vdots \\ \vdots & 0 & \dots & 0.5 & -1 & 0.5 \\ 0 & \dots & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_n \\ x_{n+1} \end{bmatrix}$$

Анализируя систему (1), можно сказать, что неподвижные агенты x_0 и x_{n+1} являются *лидерами* формации. Целью этих лидеров является обеспечение правильного движения формации. Важно отметить, что крайние виртуальные агенты-лидеры не координируют движение формации централизованно, поскольку они не передают информацию о своем положении всем агентам. Подвижные агенты $x_i, i = 1, \dots, n$, получают информацию только о расстояниях до своих крайних по номерам соседей x_{i-1} и x_{i+1} . Такую схему обмена информацией можно представить в виде ориентированного графа рис. 1.

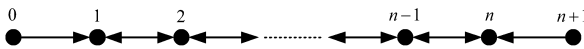


Рис. 1. Ориентированный граф алгоритма равномерного расположения на отрезке

Замечание 1. Легко заметить, что матрица системы (1) имеет два нулевых собственных числа, а граф системы, представленный на рис. 1, не содержит ориентированного остовного дерева. Следовательно задача равноудаленного расположения агентов на отрезке не является задачей достижения консенсуса (см. [9, 21]).

С другой стороны, агенты в пределе выстраиваются на равных расстояниях (длины отрезков между точками стремятся стать равными), т.е. задача расположения на отрезке напоминает задачу согласования характеристик (консенсуса). К такой задаче можно перейти при помощи несложных преобразований [6]. Первое преобразование заключается в переносе одного из неподвижных агентов (границ отрезка) в начало координат. Второе преобразование является заменой координат: расстояния между отрезками рассматриваются как новые координаты системы. Преобразованная система удовлетворяет условиям теорем о согласовании характеристик и консенсус достигается.

Особенностью перехода к задаче консенсуса является *понижение размерности* системы («потеря» одного из неподвижных агентов) и *утрата геометрического смысла* [6]. Обобщения задачи равноудаленного расположения на отрезке, предлагаемые в работе, строятся в исходной системе координат.

Представим иной способ описания стратегии, отличный от (1). Очевидно, что матрица системы (1) является вырожденной, работать с такими матрицами не всегда удобно. Кроме того, часто необходимо иметь независимое описание модели агента и управляющего закона (протокола управления), обеспечивающего требуемое движение мультиагентной системы.

С этой целью представим модель каждого агента в виде интегратора:

$$(2) \quad \dot{x}_i = u_i, \quad i = 1, \dots, n.$$

Тогда из (1) видим, что линейный протокол управления $u_i \in \mathbb{R}$ имеет вид:

$$(3) \quad u_i = \frac{x_{i+1} + x_{i-1}}{2} - x_i, \quad i = 1, \dots, n.$$

Пусть $x = [x_1, x_2, \dots, x_n]^\top$ — вектор состояния мультиагентной системы, тогда ее динамика может быть записана в компактной форме [7]

$$(4) \quad \dot{x} = Ax + b,$$

где матрица A и вектор b имеют вид

$$(5) \quad A = \begin{bmatrix} -1 & 0.5 & 0 & \dots & 0 \\ 0.5 & -1 & 0.5 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & 0.5 & -1 \end{bmatrix} \in \mathbb{R}^{n \times n},$$

$$(6) \quad b = [0, 5x_0, 0, \dots, 0, 5x_{n+1}]^\top \in \mathbb{R}^n.$$

Матрица A является трехдиагональной, ее собственные значения имеют вид [3]

$$(7) \quad \lambda_k = -2 \sin^2 \frac{k\pi}{2(n+1)}, \quad k = 1, \dots, n.$$

Поскольку $\lambda_k < 0, k = 1, \dots, n$, состояние $x^* \in \mathbb{R}^n : x^* = -A^{-1}b$ является устойчивым положением равновесия системы (4). При $t \rightarrow \infty$ имеем: $x_i \rightarrow x_0 + \frac{i}{n+1}(x_{n+1} - x_0), i = 1, \dots, n$. Это в точности означает, что вне зависимости от начальных условий агенты стремятся расположиться равноудаленно на отрезке с зафиксированными концами x_0 и x_{n+1} .

Легко показать, что для системы (4) справедлива оценка $\|x(t) - x^*\| \leq e^{\hat{\lambda}t} \|x(0) - x^*\|$, где $x(0)$ — вектор начального положения агентов. Величина $\hat{\lambda}$ характеризует скорость сходимости траекторий системы и равна

$$(8) \quad \hat{\lambda} = \max_k \lambda_k = -2 \sin^2 \frac{\pi}{2(n+1)}.$$

Двумерный случай

Покажем как можно обобщить стратегию равномерного расположения на двумерный случай. Произведем следующую группировку координат для каждого агента. Пусть положение i -го агента в момент времени $t \geq 0$ определяется вектором $\xi_i(t) = [x_i(t), y_i(t)]^\top \in \mathbb{R}^2, i = 1, \dots, n; \xi_0 = [x_0, y_0]^\top$ и $\xi_{n+1} = [x_{n+1}, y_{n+1}]^\top$ — координаты начала и конца отрезка соответственно. С учетом введенных обозначений динамика всей системы может быть записана в виде

$$(9) \quad \dot{\xi} = (A \otimes I_2)\xi + \hat{b},$$

где I_2 — единичная матрица размера 2×2 ; \otimes — символ произведения Кронекера; $\xi = [\xi_1^\top, \xi_2^\top, \dots, \xi_n^\top]^\top \in \mathbb{R}^{2n}$ — сборный вектор координат агентов системы; $\hat{b} = [0,5\xi_0^\top, 0, \dots, 0,5\xi_{n+1}^\top]^\top \in \mathbb{R}^{2n}$.

Можно сделать следующие выводы:

— как видно из формул (2)–(3), управление движением агентов происходит путем мгновенного изменения скорости. Такой алгоритм интересен с точки зрения изучения кинематики агентов, но, к примеру, *не может быть реализован физически* для механических систем. В соответствии со Вторым законом Ньютона, прикладывая силу к объекту, можно менять его ускорение, а не непосредственно скорость;

— в двумерном случае рассматривается по сути движение двух независимых подсистем — для координат x и y . В реальности координаты могут быть связаны между собой. Такое *сцепление* координат приводит к изменению математической модели системы и траекторий ее движения;

В работе предлагаются новые стратегии равноудаленного расположения агентов на отрезке, развивающие и обобщающие изложенные выше положения. Прежде чем переходить к изучению новых модификаций, обсудим, как может меняться сложность моделей мультиагентных систем и от чего она зависит.

1.2. СЛОЖНОСТЬ МУЛЬТИАГЕНТНЫХ СИСТЕМ

Как известно, хорошая математическая модель — это модель, которая наиболее полно отражает свойства реального объекта. Обычно улучшение/уточнение модели, появление новых требований к модели приводит к ее усложнению. Поскольку мультиагентные системы состоят из подсистем-агентов, связанных некоторой структурой, сложность задач управления ими можно условно определить тремя факторами [25]:

- 1) сложностью математической модели агента;
- 2) сложностью сетевой структуры, связывающей агентов;
- 3) сложностью типов связей между агентами.

Условно можно считать, что каждый из этих факторов характеризует некоторое измерение в пространстве сложности мультиагентных систем [25]. Схематическая иллюстрация этого понятия приведена на рис. 2.

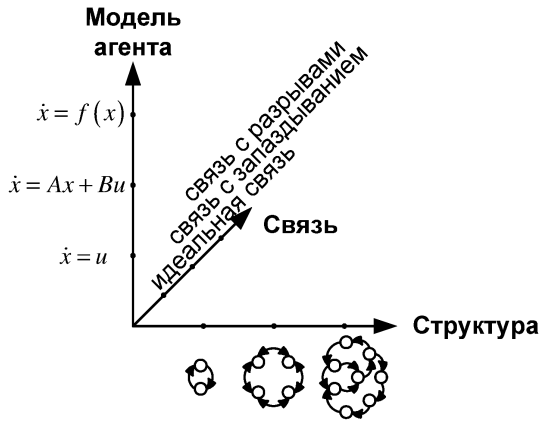


Рис. 2. Пространство сложности мультиагентной системы

Математические модели агентов могут быть как простыми линейными, так и сложными нелинейными, содержать запаздывания и нестационарные элементы. Наиболее простыми и, впрочем, наименее реалистичными моделями агентов являются интеграторы: в этом случае управление меняет непосредственно скорость агента. В этом случае решаются задачи кинематики. Самыми простыми динамическими моделями агентов являются двойные интеграторы.

Структура мультиагентной системы показывает, по каким направлениям распространяется информация в системе, т.е. какие выходы/состояния одних агентов доступны для измерения (поступают на входы) другим. Усложнение структуры может происходить по разным причинам, например, при введении дополнительных обратных связей, появлении уровней иерархии и т.д. Структура мультиагентной системы может быть представлена структурной схемой [8], графом [15] и т.п. Наиболее широкое применение теория графов получила для решения задач

консенсуса/синхронизации. Важные характеристики алгоритмов консенсуса связаны со спектральными свойствами лапласовских матриц графов, определяющих структуру системы [1, 15, 21].

Связи между агентами могут быть как идеальными (простейший случай), так и неидеальными, например, содержать запаздывание, нелинейности. Разрыв связей может приводить к тому, что некоторые агенты «покидают» систему, меняя структуру системы и уменьшая ее порядок. Если движение системы происходит в некотором m -мерном пространстве (например, на плоскости, $m = 2$), то изменение каждой из координат может происходить либо независимо — в этом случае мы имеем дело с m никак не связанными друг с другом подсистемами, — либо существуют связи между координатами. В линейном случае такая связь может определяться некоторой матрицей, например, матрицей поворота [19, 20]. В таких случаях говорят, что имеет место *сцепление координат (coordinate coupling)* [21].

Получение новых алгоритмов управления формациями предполагает как изменение модели агентов, так и изменение типов связей и структуры системы.

1.3. ЦЕЛЬ РАБОТЫ

В соответствии со схемой, представленной на рис. 2, предлагаются следующие обобщения задачи равноудаленного расположения агентов на отрезке:

1) усложнение модели агента: рассматриваются модели агентов в виде двойных интеграторов. Предлагается протокол управления, позволяющий управлять движением агентов посредством изменения ускорения;

2) усложнение структуры системы: предлагается двухуровневый иерархический алгоритм, позволяющий равноудаленно располагать на отрезке группы агентов;

3) усложнение типа связи: рассматриваются алгоритмы, учитывающие сцепление координат. Изучаются новые траектории движения агентов.

2. Равноудаленные расположения на отрезке со сцеплением координат

2.1. МОДЕЛИ АГЕНТОВ ПЕРВОГО ПОРЯДКА

Для начала рассмотрим стратегию равноудаленного расположения на отрезке со сцеплением координат для случая, когда модели агентов имеют вид интеграторов. Сцепление координат можно получить введением в выражение (9) матрицы более общего вида, чем единичная. Рассмотрим ситуацию, когда вектор скорости каждого агента смещен на одинаковый для всех агентов угол $\alpha \in [-\pi, \pi]$, тогда соответствующая матрица имеет смысл матрицы поворота.

Наличие такой связи может быть обусловлено физическими ограничениями на действия агентов, к примеру необходимостью обходить препятствия мобильными роботами, либо наличием внешних возмущений, отклоняющих вектор скорости агента, и т.п. Для задач циклического преследования алгоритм, использующий матрицу поворота, был впервые предложен в [17, 18]. В рамках задач консенсуса с моделями агентов в виде интеграторов первого и второго порядка похожий подход рассматривался в [19, 20]. Введение матрицы поворота позволяет получить новые траектории системы. Воспользуемся этими идеями для задачи равномерного расположения точек на отрезке.

Итак, управление на входе i -го агента будет определяться следующей формулой:

$$(10) \quad u_i = R(\alpha) \left(\frac{\xi_{i+1} + \xi_{i-1}}{2} - \xi_i \right), \quad i = 1, \dots, n,$$

где $R(\alpha)$ — матрица поворота, определяемая следующим выражением:

$$(11) \quad R(\alpha) = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}.$$

Система уравнений, описывающая динамику всей системы из n агентов, имеет вид

$$(12) \quad \dot{\xi} = (A \otimes R(\alpha))\xi + b^*,$$

$$(13) \quad b^* = (I_n \otimes R(\alpha))\hat{b} \in \mathbb{R}^{2n}.$$

Теорема 1 [4]. Система (12) устойчива тогда и только тогда, когда $0 \leq |\alpha| < \frac{\pi}{2}$. При этом $\xi_k \rightarrow \xi_0 + \frac{k}{n+1}(\xi_{n+1} - \xi_0)$, $t \rightarrow \infty$, $k = 1, \dots, n$, т.е. все агенты из любого начального положения стремятся выстроиться на отрезке с фиксированными концами ξ_0 и ξ_{n+1} на равном расстоянии друг от друга. При $n \rightarrow \infty$ оценка скорости сходимости (12) определяется $\hat{\lambda} \approx -\frac{\pi^2 \cos \alpha}{2n^2}$. Для значений угла $\frac{\pi}{2} \leq |\alpha| < \pi$ цель управления не достигается.

Доказательство теоремы использует свойство спектра произведения Кронекера $A \otimes R(\alpha)$, который равен прямому произведению множеств собственных чисел матриц A и $R(\alpha)$ [14].

Замечание 2. Результаты, приведенные выше, легко обобщаются на трехмерных случай с вращением агентов относительно оси $(0, 0, 1)^T$ [4].

Замечание 3. В общем случае в протоколе управления (10) может фигурировать не матрица поворота $R(\alpha)$, а матрица более общего вида. Введение в уравнение движения матрицы $R(\alpha)$ имеет определенный физический смысл: тем самым мы отклоняем вектор скорости. Кроме того, для $R(\alpha)$ известна аналитическая запись спектра, что облегчает анализ задачи и ее обобщение на многомерные случаи. Если для сцепления координат применяется матрица общего вида, то вопросы трактовки ее физического смысла и вычисления спектра остаются открытыми.

2.2. МОДЕЛИ АГЕНТОВ ВТОРОГО ПОРЯДКА

Усложним также и модели агентов, учитывая при этом сцепление координат при помощи матрицы поворота $R(\alpha)$. Пусть модель агента описывается двойным интегратором

$$(14) \quad \ddot{x}_i = u_i, \quad i = 1, \dots, n.$$

Выберем закон управления в следующем виде [16]:

$$(15) \quad u_i = R(\alpha) \left(\frac{\xi_{i+1} + \xi_{i-1}}{2} - \xi_i - a\dot{\xi}_i \right), \quad i = 1, \dots, n,$$

где a — некоторый настраиваемый параметр.

В таком случае система уравнений, описывающая динамику n агентов, будет иметь вид

$$(16) \quad \ddot{\xi} = (A \otimes R(\alpha))\dot{\xi} - aR(\alpha)\dot{\xi} + b^*,$$

где матрицы A и $R(\alpha)$ имеют вид (5) и (11) соответственно, $b^* = (I_n \otimes R(\alpha)) [0, 5\xi_0^\top, 0, \dots, 0, 5\xi_{n+1}^\top]^\top \in \mathbb{R}^{2n}$.

Справедлив следующий критерий устойчивости.

Теорема 2 [4, 16]. Система (16) устойчива тогда и только тогда, когда

$$a^2 \cos \alpha > 2 \sin^2 \frac{\pi n}{2(n+1)} \sin^2 \alpha, \quad a > 0.$$

При этом $\xi_k \rightarrow \xi_0 + \frac{k}{n+1}(\xi_{n+1} - \xi_0)$, $t \rightarrow \infty$, $k = 1, \dots, n$, т.е. все агенты из любого начального положения стремятся выстроиться на отрезке с фиксированными концами ξ_0 и ξ_{n+1} на равном расстоянии друг от друга.

Доказательство теоремы основано на использовании критерия устойчивости мультиагентных систем Поляка–Цыпкина и введенного в [8] понятия области устойчивости мультиагентной системы.

Замечание 4. Заметим, что протокол управления (15) отличается от (10) необходимостью знания скорости i -го агента. По сути, слагаемые $-a\dot{\xi}_i$, $a > 0$, имеют смысл отрицательных обратных связей, стабилизирующих систему.

Пример 1. В качестве примера рассмотрим систему из 5 подвижных агентов для двумерного случая. На рис. 3–7 представлены устойчивые траектории движения агентов для разных значений угла поворота α и настраиваемого параметра a (в случае алгоритма второго порядка). Начальные координаты агентов имеют вид:

$$x(0) = [0,31, 0,5, -0,83, 0,83, 0,65]^\top,$$

$$y(0) = [0,38, -0,1, -0,54, -0,7, 0,08]^\top.$$

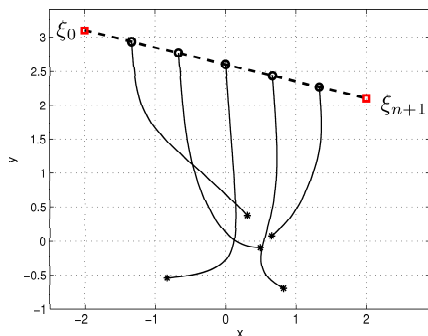


Рис. 3. Начальное положение и траектории агентов. Агенты первого порядка, угол $\alpha = 0$

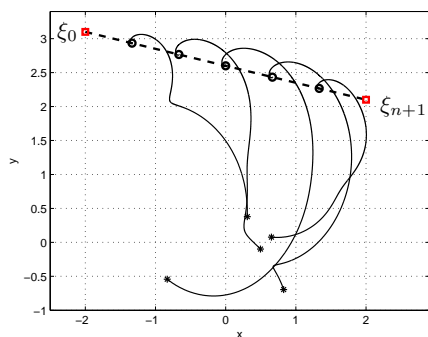


Рис. 4. Начальное положение и траектории агентов. Агенты первого порядка, угол $\alpha = \pi/4$

Проанализируем полученные результаты. Во-первых, сцепление координат меняет траектории движения системы, в нашем случае введение матрицы поворота с $\alpha > 0$ отклоняет траектории агентов вправо. С ростом угла смещения траектория искривляется сильнее. Во-вторых, скорость сходимости стратегий определяется выбором параметров α и a (в случае агентов второго порядка). Как и следовало ожидать, увеличение угла смещения α приводит к снижению скорости сходимости в обоих случаях. С другой стороны, системы (12) и (16) демонстрируют разную

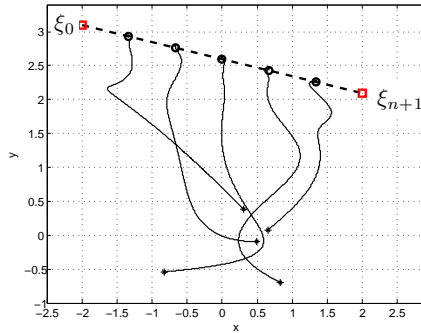


Рис. 5. Начальное положение и траектории агентов. Агенты второго порядка, $a = 0,6$, $\alpha = 0$

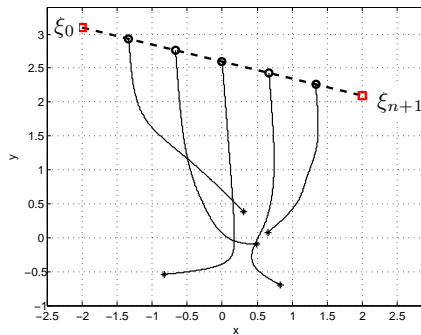


Рис. 6. Начальное положение и траектории агентов. Агенты второго порядка, $a = 2$, $\alpha = 0$

скорость сходимости при одинаковых углах α в зависимости от выбора параметра a . Например, при $\alpha = 0$ и $a = 0,6$ система (16) показывает более высокую скорость сходимости, чем (12), а при $a = 2$ — более медленную. Уменьшение параметра a приводит к увеличению колебательности переходных процессов в системе (16).

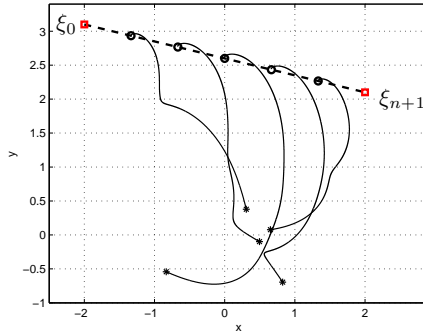


Рис. 7. Начальное положение и траектории агентов. Агенты второго порядка, $a = 2$, $\alpha = \pi/6$

3. Иерархический протокол управления

В данном разделе предлагается применение концепции иерархичности для алгоритма равноудаленного расположения групп агентов на отрезке. Рассмотрим следующую задачу: имеется n_g групп мобильных агентов, по n агентов в каждой (общее число агентов $N = n \times n_g$); необходимо получить такой закон управления, чтобы *центроиды* — центры масс каждой из групп — стремились равномерно расположиться на заданном отрезке $[x_0, x_{n_g}]$, а агенты каждой группы стягивались к ее центроиду.

Для решения поставленной задачи предлагается протокол управления в виде двухуровневой иерархической схемы. На втором (высоком) уровне иерархии рассматривается движение n_g групп агентов по закону равноудаленного расположения на отрезке, на первом (низком) — движение агентов внутри каждой из групп по алгоритму циклического преследования. Схема, иллюстрирующая идею алгоритма, представлена на рис. 8.

Идея линейного циклического преследования заключается в том, что скорость i -го агента равна (либо пропорциональна) расстоянию до следующего $(i + 1)$ -го по модулю n агента, а вектор скорости направлен в его сторону. Такой закон был впервые описан в работе Ж.Г. Дарбу [11]. Простейший линейный протокол

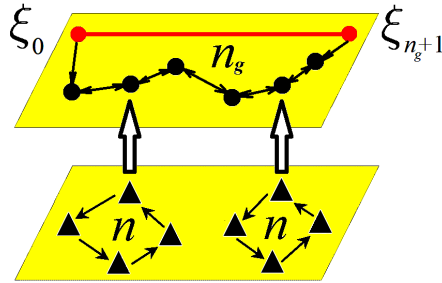


Рис. 8. Двухуровневый алгоритм равноудаленного расположения агентов на отрезке

циклического преследования для моделей агентов в виде одиночных интеграторов (2) в непрерывном времени для системы из n агентов имеет следующий вид [22]:

$$(17) \quad u_i = x_{i+1} - x_i, \quad x_i \in \mathbb{R}, \quad i = 1, \dots, n.$$

Легко заметить, что при таком алгоритме центр масс формации остается неподвижным и равен среднему начальных условий агентов.

Рассмотрим алгоритм циклического преследования в \mathbb{R} , в соответствии с которым каждый агент преследует следующего по модулю n агента со смещением [22]:

$$(18) \quad \dot{x}_i = (x_{i+1} + c_i) - x_i, \quad i = 1, \dots, n.$$

где c_i , $i = 1, \dots, n$, — смещения.

Обозначим центр масс группы агентов (18) через \bar{x} . Тогда справедливо следующее равенство [22]:

$$(19) \quad \dot{\bar{x}} = \frac{1}{n} \sum_{i=1}^n c_i.$$

Отсюда следует, что выбирая должным образом величины смещений c_i , можно обеспечить желаемое движение центроидов и прийти к иерархической схеме.

Введем следующие обозначения для агентов: $x_{p,q}$, где $p = 1, \dots, n_g$ является индексом группы, а $q = 1, \dots, n$ — индексом агента в группе. Тогда двухуровневая иерархическая схема может быть записана в виде [5]

$$\text{группа 1} \quad \left\{ \begin{array}{l} \dot{x}_{1,1} = x_{1,2} - x_{1,1} + d_{1,1} \\ \dot{x}_{1,2} = x_{1,3} - x_{1,2} + d_{1,2} \\ \vdots \\ \dot{x}_{1,n} = x_{1,1} - x_{1,n} + d_{1,n} \end{array} \right\}$$

(20)

⋮

$$\text{группа } n_g \quad \left\{ \begin{array}{l} \dot{x}_{n_g,1} = x_{n_g,2} - x_{n_g,1} + d_{n_g,1} \\ \dot{x}_{n_g,2} = x_{n_g,3} - x_{n_g,2} + d_{n_g,2} \\ \vdots \\ \dot{x}_{n_g,n} = x_{n_g,1} - x_{n_g,n} + d_{n_g,n} \end{array} \right\}$$

где $d_{p,q}$ являются смещениями.

Потребуем, чтобы центры каждой группы двигались по закону равноудаленного расположения на отрезке:

$$(21) \quad \dot{\bar{x}}_p = \frac{\bar{x}_{p+1} + \bar{x}_{p-1}}{2} - \bar{x}_p, \quad p = 1, \dots, n_g,$$

где \bar{x}_p — центр p -й группы.

По определению центроид p -й группы описывается выражением [22]

$$(22) \quad \bar{x}_p := \frac{1}{n} \sum_{q=1}^n x_{p,q}.$$

Возникает вопрос: как выбрать смещения для обеспечения желаемого движения центроидов (21)? Из (19) следует, что динамика центроида p -й группы имеет вид

$$(23) \quad \dot{\bar{x}}_p = \frac{1}{n} \sum_{q=1}^n d_{p,q}.$$

В сочетании с (21) получаем

$$(24) \quad \sum_{q=1}^n d_{p,q} = n \left(\frac{\bar{x}_{p+1} + \bar{x}_{p-1}}{2} - \bar{x}_p \right),$$

или

$$\sum_{q=1}^n d_{p,q} = n \left(\frac{\sum_{q=1}^n x_{p+1,q} + \sum_{q=1}^n x_{p-1,q}}{2} - \sum_{q=1}^n x_{p,q} \right).$$

Выберем смещения $d_{p,q}$ следующим образом:

$$(25) \quad d_{p,q} = \frac{x_{p+1,q} + x_{p-1,q}}{2} - x_{p,q}, \quad p = 1, \dots, n_g.$$

Последнее выражение означает, что q -й агент в p -й группе стремится расположиться в середине отрезка, соединяющего q -х агентов в $(p-1)$ -й и $(p+1)$ -й группах. Динамика всей системы из N агентов (20) с учетом (25) имеет следующий вид

$$(26) \quad \dot{x} = \bar{A}x + \bar{b},$$

где $x = [x_{1,1}, x_{1,2}, \dots, x_{1,n}, \dots, x_{n_g,1}, \dots, x_{n_g,n}]^\top$ – сборный вектор координат всех агентов, а матрица системы имеет вид

$$(27) \quad \bar{A} = I_{n_g} \otimes C + D \otimes I_n.$$

Матрица C является циркулянтной [12]:

$$(28) \quad C = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & \dots & 0 & -1 \end{bmatrix} \in \mathbb{R}^{n \times n},$$

Матрица $D \in \mathbb{R}^{n_g \times n_g}$ имеет вид

$$(29) \quad D = \begin{bmatrix} -1 & 0.5 & 0 & \dots & 0 \\ 0.5 & -1 & 0.5 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & 0.5 & -1 \end{bmatrix} \in \mathbb{R}^{n_g \times n_g},$$

и используется для описания движения центроидов групп. Вектор $\bar{b} = b \otimes \mathbf{1}$, где $\mathbf{1} = [1, 1, \dots, 1]^\top \in \mathbb{R}^n$, и $b \in \mathbb{R}^{n_g}$ определяется аналогично (6). Матрицы I_n, I_{n_g} – единичные матрицы соответствующих размерностей.

Учитывая тот факт, что матрица $\bar{A} \in \mathbb{R}^{N \times N}$ является кронеровой суммой матриц C и D : $\bar{A} = C \oplus D$, можно сформулировать следующую теорему.

Теорема 3. Система (26) устойчива, и множество собственных значений матрицы \bar{A} является прямой суммой множеств собственных значений матриц C и D . При этом $x_{p,q} \rightarrow x_0 + \frac{p}{n_g+1}(x_{n_g+1} - x_0)$, $t \rightarrow \infty, p = 1, \dots, n_g$, т.е. для любого начального положения каждый агент сходится к центроиду своей

группы, а центроиды групп в свою очередь стремятся расположиться равномерно на отрезке с фиксированными концами x_0 и x_{n_g+1} . Скорость сходимости алгоритма при $n_g \rightarrow \infty$ определяется $\hat{\lambda} \approx -\frac{\pi^2}{2n_g^2}$.

Доказательство. Как было показано ранее, собственные числа матрицы D имеют вид $\lambda_p = -2 \sin^2 \frac{p\pi}{2(n_g+1)}$, $p = 1, \dots, n_g$, и собственные числа матрицы C также известны и имеют вид $\bar{\lambda}_q = e^{2\pi jq/n} - 1$, $q = 1, \dots, n$ [12]. Собственные числа матрицы $\bar{A} = I_{n_g} \otimes C + D \otimes I_n = C \oplus D$ легко определяются из свойств кронекеровой суммы. Матрица C имеет единственное нулевое собственное число, остальные расположены в левой полуплоскости; все собственные числа матрицы D отрицательны. Отсюда следует, что матрица \bar{A} гурвицева, т.е. (26) устойчива. Очевидно, что скорость сходимости для системы (26) при $n_g \rightarrow \infty$ определяется $\hat{\lambda} \approx -\frac{\pi^2}{2n_g^2}$.

Замечание 5. Задача была рассмотрена в одномерном пространстве, но может быть легко обобщена на двумерный введением вектора координат агента ξ_i и выкладками, аналогичными (9).

Пример 2. Рассмотрим систему на плоскости, состоящую из 3 групп по 3 подвижных агентов в каждой. Начальные координаты агентов имеют вид:

$$x(0) = [0,31, 0,5, -0,83, 0,83, 0,65, 0,99, -0,11, 0,92, 0,55]^\top,$$

$$y(0) = [0,38, -0,1, -0,54, -0,7, 0,08, -0,84, -0,8, -0,99, 0,63]^\top.$$

На рис. 9 представлены траектории движения агентов.

Из полученных результатов видно, что агенты каждой группы быстро стягиваются к центроиду и, сильно сблизившись, стремятся равномерно выстроиться на отрезке. Отметим также, что скорость сходимости системы (26) зависит от числа групп агентов и не зависит от числа агентов в каждой из групп.

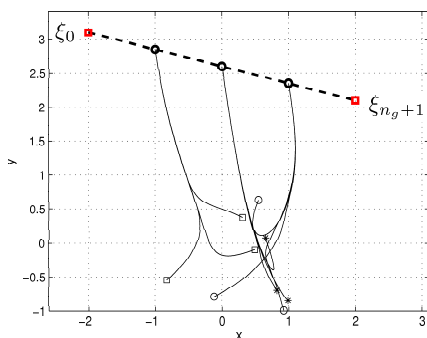


Рис. 9. Иерархический алгоритм. Начальные условия агентов каждой из трех групп отмечены разными символами

4. Заключение

В работе предложены и изучены новые обобщения типовой задачи управления формациями — задачи равноудаленного расположения агентов на отрезке, обобщающие и дополняющие существующие результаты. В качестве первого шага предложен протокол управления со сцеплением координат для мультиагентной системы с моделями агентов в виде интеграторов. В качестве матрицы, связывающей координаты, рассмотрена матрица поворота. Синтезирован протокол управления для агентов второго порядка, также учитывающий связь между координатами. Для каждой из полученных систем уравнений проведен анализ устойчивости и сделаны выводы по характеру движения агентов и скорости сходимости. В качестве последнего шага рассмотрена двухуровневая иерархическая схема управления, позволяющая располагать на отрезке группы агентов, внутри которых агенты движутся по закону циклического преследования. Таким образом, в соответствии со схемой на рис. 2 по каждому из направлений в «пространстве сложности» было сделано некоторое обобщение/усложнение стратегии равномерного расположения на отрезке.

Литература

1. АГАЕВ Р.П., ЧЕБОТАРЕВ П.Ю. *Сходимость и устойчивость в задачах согласования характеристик (Обзор базовых результатов)* // Управление большими системами. Специальный выпуск 30.1 «Сетевые модели в управлении», 2010. – С. 470–505.
2. АМЕЛИН К.С., ГРАНИЧИН О.Н. *Мультиагентное сетевое управление группой легких БПЛА* // Нейрокомпьютеры: разработка, применение. – 2011. – №6. – С. 64–72.
3. ВОЕВОДИН В.В., КУЗНЕЦОВ Ю.А., Матрицы и вычисления. – М.: Наука, 1984. – 320 с.
4. ПАРСЕГОВ С.Э. *Обобщенные линейные алгоритмы управления формациями* // Стохастическая оптимизация в информатике. – 2011. – Т. 7 (Вып. 1). – С. 186–203.
5. ПАРСЕГОВ С.Э. *Некоторые новые иерархические алгоритмы управления формациями* // Труды международной научно-практической конференции «Теория активных систем – 2011», Т. 3, Москва, 2011. – С. 238–241.
6. ПАРСЕГОВ С.Э. *Равноудаленное расположение на отрезке как задача достижения консенсуса* // Управление большими системами: Материалы IX Всероссийской школы-конференции молодых ученых. Том 1. Липецкий государственный технический университет. – Тамбов-Липецк: Изд-во Першина Р.В., 2012. – С. 79–84.
7. ПЕТРИКЕВИЧ Я.И. *Линейные алгоритмы управления геометрическим расположением объектов в многоагентной системе* // Управление большими системами. Специальный выпуск 30.1 «Сетевые модели в управлении», Москва, 2010. – С. 665–680.
8. ПОЛЯК Б.Т., ЦЫПКИН Я.З. *Устойчивость и робастная устойчивость однотипных систем* // Автоматика и телемеханика. – 1996. – №11. – С. 91–104.
9. ЧЕБОТАРЕВ П.Ю., АГАЕВ Р.П. *Согласование характеристик в многоагентных системах и спектры лапласовских*

- матриц орграфов* // Автоматика и телемеханика. – 2009 – №3 – С. 136–151.
10. ЦЕРБАКОВ П.С. *Управление формациями: схема Ван Лоуна и другие алгоритмы* // Управление большими системами. Специальный выпуск 30.1 «Сетевые модели в управлении», Москва, 2010. – С. 681–696.
 11. DARBOUX J.G. *Sur un problème de géométrie élémentaire* // Bulletin des Sciences Mathématiques et Astronomiques. – 1878. – №2(1). – P. 298–304.
 12. GRAY R.M. *Toeplitz and Circulant Matrices: A review* // Foundations and Trends in Communications and Information Theory. – 2006. – №2(3). – P. 155–239.
 13. HARA S., HAYAKAWA T. AND SUGATA H. *Stability Analysis of Linear Systems with Generalized Frequency Variables and Its Applications to Formation Control* // Proc. IEEE CDC. – 2007. – P. 1459–1466.
 14. HORN R.A., JOHNSON C.R. *Topics in matrix analysis*. – Cambridge. Univ. Press, 1991.
 15. OLFATI-SABER R., FAX J.A. AND MURRAY R.M. *Consensus and Cooperation in Networked Multi-Agent Systems* // Proceedings of the IEEE 2007. – №95(1). – P. 215–233.
 16. PARSEGOV S.E. *Allocation of agents on a line: simple algorithm and generalizations* // Proc. 14th Baltic olympiad on automatic control, Saint-Petersburg, Russia, Sep. 21–23, 2011. – P. 119–125.
 17. PAVONE M. AND FRAZZOLI E. *Decentralized policies for geometric pattern formation and path coverage* // ASME Journal of Dynamic Systems, Measurement, and Control. – 2007. – №129(5). – P. 633–643.
 18. RAMIREZ J.L., PAVONE M., FRAZZOLI E. AND MILLER D.W. *Distributed Control of Spacecraft Formations via Cyclic Pursuit: Theory and Experiments* // AIAA Journal of Guidance, Control, and Dynamics. – 2010. – №33(5). – P. 1655–1669.

19. REN W. *Collective motion from consensus with Cartesian coordinate coupling - Part I: Single-integrator kinematics* // Proc. IEEE CDC. – 2008. – P. 1006–1011.
20. REN W. *Collective motion from consensus with Cartesian coordinate coupling - Part II: Double-integrator dynamics* // Proc. IEEE CDC. – 2008 – P. 1012–1017.
21. REN W., CAO Y. *Distributed Coordination of Multi-agent Networks: Emergent Problems, Models, and Issues*. – London: Springer-Verlag, 2011.
22. SMITH S.L., BROUCKE M.E. AND FRANCIS B.A. *A Hierarchical Cyclic Pursuit Scheme for Vehicle Networks* // Automatica. – 2005. – №41(6). – P. 1045–1053.
23. VEERMAN J.J.P., LAFFERRIERE G., CAUGHMAN J.S., WILLIAMS A. *Flocks and Formations* // Journal of Statistical Physics. – 2005. – №121(5-6). – P. 901–936.
24. WAGNER I.A., BRUCKSTEIN A.M. *Row straightening via local interactions* // Circuits, Systems, and Signal Processing. – 1997. – №16(3). – P. 287–305,
25. WIELAND P. *From static to dynamic couplings in consensus and synchronization among identical and non-identical systems*: PhD Thesis. – Berlin: Logos Verlag, 2010.

COORDINATE COUPLING AND HIERARCHICAL ALGORITHMS IN UNIFORM ALLOCATION OF AGENTS ON A SEGMENT

Sergey Parsegov, Institute of Control Sciences of RAS, Moscow, Phd student (s.e.parsegov@gmail.com).

Abstract: New control strategies of uniform allocation of agents on a segment are proposed. Different modifications of this typical formation control problem include double-integrator models of agents and coordinate coupling. Moreover, a new two layer hierarchical scheme to provide equidistant allocation of groups of agents on a segment is designed.

Keywords: formation control, coordinate coupling, hierarchical algorithms.

Статья представлена к публикации членом редакционной коллегии П. С. Щербаковым

УДК 681.51:629.78

ББК 32.973.202

ПУТИ ПОВЫШЕНИЯ НАДЕЖНОСТИ И КАЧЕСТВА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ В КОСМИЧЕСКОЙ ОТРАСЛИ

Тюгашев А. А.¹

(ФГБОУ ВПО «Самарский государственный аэрокосмический университет им. академика С.П. Королева (Национальный исследовательский университет)», Самара)

Ильин И. А.²

(ЦНИИМАШ, Королев)

Ермаков И. Е.³

(НПО «Тесла», Орел)

Описывается текущая ситуация с созданием бортового программного обеспечения в аэрокосмической отрасли, описываются проблемы и намечаются пути их решения.

Ключевые слова: управляющее программное обеспечение, космический аппарат, верификация, тестирование программ.

1. Введение

В настоящее время в высокотехнологичных и наукоемких отраслях промышленности в нашей стране существует ряд серьезных проблем, среди которых – необходимость обновления основных фондов, оборудования и используемых технологий, проблема нехватки квалифицированных специалистов, особенно наиболее активного и продуктивного возраста – 30–50 лет, и т.д. Актуальной частной задачей при этом является повыше-

¹ Андрей Александрович Тюгашев, доктор технических наук, доцент (tau7@ssau.ru).

² Илья Александрович Ильин (ilyailyin@yandex.ru).

³ Илья Евгеньевич Ермаков (ilya@ertakov.net.ru).

ние надежности и качества космической техники, включая системы управления и их программное обеспечение [4].

2. Современное состояние в области создания программного обеспечения в космической отрасли

Как показывает практика, одной из главных причин аварий, потери дорогостоящих аппаратов, в которых воплощен труд многотысячных коллективов, являются сбои и ошибки функционирования систем управления. Часто к критическим последствиям приводят недостатки, ошибки и низкий уровень отказоустойчивости бортового программного обеспечения, на которое и возлагается в настоящее время, как правило, непосредственное решение задач управления. Рассматривая ситуацию через призму проблем в данной предметной области, можно отметить следующее. Еще в годы реализации программы «Аполлон» ее руководитель Дж. Ф. Ши отмечал [1]: «Хотя инженеры в любой области стремятся расценивать свои собственные разработки как наиболее ответственные, все же можно утверждать, что после создания необходимых ракетных двигателей основой надежного выполнения широких задач при длительных космических полетах являются вычислительные средства, имеющиеся на борту космического корабля». И сейчас [5] по многим оценкам, именно построение комплекса программного обеспечения – наземного и бортового, является зачастую критическим путем на сетевом графике работ по созданию современных аэрокосмических комплексов. Современному программному обеспечению свойство сложности присуще вообще, как неотъемлемая черта. Объем используемых программ достигает десятков миллионов строк исходного кода. Успешная и своевременная реализация проектов в сфере создания современного сложного программного обеспечения – весьма сложная проблема даже для опытных высококвалифицированных коллективов, вовсе не всегда успешно решаемая [6, 10, 12].

Еще более усложняется ситуация, когда речь идет об управляющих программах, функционирующих в режиме реального времени и от правильного функционирования которых напрямую зависит реализация важнейших миссий, безопасность, а иногда – человеческие жизни. К надежности подобного программного обеспечения, называемого критическим, предъявляются особо высокие требования. Именно такого рода программное обеспечение используется в аэрокосмической отрасли. К сожалению, ошибки в программах приводили и приводят к катастрофам и авариям.

Можно отметить следующие современные тенденции при создании систем управления космическими комплексами:

- реализация управления современными средствами выведения и космическими аппаратами с помощью бортовых цифровых вычислительных машин (БЦВМ) или комплекса БЦВМ – БВС (бортовой вычислительной системы);
- перенос «центра тяжести» в принятии решений с Земли на борт;
- возрастающий объем и сложность бортового программного обеспечения;
- необходимость согласованной работы большого числа приборов, устройств, функциональных программных модулей.

Наиболее широко на предприятиях отрасли применяется подход к обеспечению надежности, основанный на многоэтапной отладке и испытаниях управляющего программного обеспечения. Сначала проводится, как правило, автономная отладка, затем проверка функционирования управляющего программного обеспечения (ПО) во взаимодействии со специально созданными программными моделями бортовых систем и параметров движения летательного аппарата, и на заключительном этапе комплексной отладки – совместно с реальной БВС и бортовой аппаратурой [6, 7].

Несмотря на это, можно говорить о частном «кризисе» разработки ПО в космической отрасли. Трудоемкость, стоимость и сложность разработки ПО вносят значительный вклад в стои-

мость и сроки создания ракетно-космических комплексов в целом. Повторяющиеся ошибки в управляющем ПО приводят к потерям дорогостоящих аппаратов, утрате результатов длительного труда многотысячных коллективов, включая международные проекты. Существующие методы и технологии не позволяют гарантировать необходимый уровень надежности, отказоустойчивости и безопасности.

Говоря о причинах сложившейся ситуации, можно выделить следующие:

- зависимость от уникального опыта и квалификации конкретного разработчика;
- переход на новые вычислительные платформы (в том числе зарубежные);
- недопонимание между специалистами по бортовым системам, проектантами, программистами;
- большой объем программной документации, требующий актуализации наряду с ПО;
- противоречивость требований к управляющему ПО.

При этом следует сказать, что тестирование ПО не дает необходимых гарантий качества и надежности и не может дать их в принципе, что отмечалось еще классиками программирования [3]. Использование программных моделей бортовой аппаратуры (БА) на испытательных стендах в свою очередь требует верификации и ускорения разработки комплекса моделей. Часто одна и та же логика управления должна реализовываться на различных аппаратных и программных платформах.

Ясно, что нужна разработка и внедрение в аэрокосмической отрасли комплекса мер для преодоления сложившегося кризиса, качественного увеличения уровня надежности, безопасности и отказоустойчивости ПО, и на этой основе – отказоустойчивости и надежности ракетно-космических изделий в целом с соответствующим повышением вероятности успешного выполнения целевых задач.

3. Методы повышения надежности и качества программного обеспечения

На основании проведенного анализа, изучения отечественного и зарубежного опыта, современных достижений информатики и программирования, авторы выделяют следующие методы повышения качества и надежности ПО в отрасли:

- совершенствование процессов разработки ПО – обеспечение качества за счет использования современных методик, языков и инструментов;
- создание методик и языков для строгой, непротиворечивой и полной спецификации требований к ПО;
- автоматизация тестирования ПО с обеспечением требуемой полноты покрытия;
- автоматическая генерация фрагментов программ, заведомо соответствующих спецификации (для тех составляющих частей системы управления, где это возможно);
- формальные методы верификации со строгим доказательством свойств ПО;
- использование проверки моделей (*model checking*).

Кроме того, для решения проблемы своевременного создания комплекса бортового ПО и повышения производительности труда разработчиков целесообразна работа по следующим направлениям:

- использование современных методик, языков и инструментария;
- внедрение платформенно-независимого «промежуточного слоя» ПО;
- применение интуитивно понятных человеку визуальных представлений;
- автоматическая генерация фрагментов программ (где это возможно);
- автоматизация тестирования ПО;

- снижение зависимости от уникального опыта и квалификации конкретного разработчика за счет создания полной и качественной документации и использования проблемно-ориентированных языковых средств высокого уровня, удобных для человека.

Внедрение в жизнь перечисленных мер возможно путем разработки и внедрения на предприятиях аэрокосмической отрасли специализированных методик спецификации, проектирования, разработки, синтеза и верификации программного обеспечения, поддерживаемых специализированными инструментальными программными средствами. Набор подобных средств (индивидуально подбираемый на каждом предприятии в зависимости от специфики изготавливаемых изделий) должен создаваться со следующими главными целями:

- обеспечение необходимых показателей надежности, качества и безопасности ПО в отрасли;
- снижение сроков, стоимости и трудоемкости разработки управляющего ПО;
- снижение сроков, стоимости и трудоемкости разработки программных моделей БА, используемых при испытаниях;
- автоматизация построения актуальной, полной и точной программной документации.

Примером работ в направлении построения подобных средств автоматизации может служить программный комплекс ГРАФКОНТ/ГЕОЗ, созданный в Самарском государственном аэрокосмическом университете (национальном исследовательском университете), по заказу ГНПРКЦ «ЦСКБ-Прогресс» [5, 9]. Программный комплекс включает следующие основные модули и средства:

- транслятор с проблемно-ориентированного языка спецификаций управляющих алгоритмов реального времени;
- визуальный конструктор циклограмм управляющих алгоритмов;
- визуальный конструктор управляющей программы комплексного функционирования;

- генератор управляющей программы на языках ассемблер и Си (может параметрически настраиваться на другие языки программирования);
- генератор таблицы возможных вариантов исполнения управляющего алгоритма;
- генератор отладочных заданий для каждого варианта исполнения;
- генератор таблицы управляющих и информационных связей управляющей программы.

Среди других работ в данной области можно назвать технологию ГРАФИТ-ФЛОКС, используемую успешно в НПЦ АП им. Н.А. Пилюгина, на протяжении ряда лет, в том числе в проектах «Морской старт», РБ «Бриз-М», РБ «Фрегат» [2, 8]. Технология использует визуальный конструктор процедурной части управляющих программ ГРАФИТ и табличный процессор декларативной части ФЛОКС. Комплексную автоматизированную методику создания БПО внедряет в производственный процесс в настоящее время МОКБ «Марс».

Весьма значительных успехов в совершенствовании технологии создания управляющих программ для своих космических аппаратов достигло ОАО «ИСС», см. работы А.А. Колташева и В.В. Хартова [7, 11]. Важнейшими принципами при этом стали использование нескольких слоев программного обеспечения с выделением платформенно-независимого промежуточного слоя, использование и бортовых интеллектуальных интерпретаторов СЕАНС и ДКД, реализующих автономное управление КА. При написании бортовых программ используется язык Модуля-2 со строгой статической типизацией, жесткими нормами модульности и структурного программирования, позволяющие снизить количество ошибок в БПО.

В качестве направлений дальнейших работ в настоящее время целесообразным представляется следующее.

Первое направление связано с созданием теоретического и методологического фундамента для качественного совершенствования процессов разработки БПО за счёт внедрения и разви-

тия современных достижений науки и практики программной инженерии, таких как:

- высокоуровневые языки спецификаций, проектирования, моделирования, программирования, в том числе визуальные;
- формальные методы спецификации требований к ПО систем управления, верификации и валидации ПО, аттестации ПО на соответствие требованиям, гарантирующие качество ПО в противоположность традиционному тестированию;
- развитое технологическое ПО – инструменты автоматизации, поддерживающие ПО на всех этапах жизненного цикла.

Использование автоматизированных средств разработки ПО в отрасли даст возможность обеспечить гарантированный уровень качества ПО (его надёжности, отказоустойчивости и безопасности), снижение влияния человеческого фактора на качество ПО, автоматическую генерацию программной документации, автоматизацию генерации полного набора тестов с учетом внутренней структуры программ, технологичность и единообразие процессов создания ПО, сокращение сроков и трудозатрат.

Второе направление – это всемерное развитие и совершенствование нормативных основ в области процессов создания БПО в виде рекомендаций, требований, отраслевых стандартов. Подобного рода единый отраслевой нормативный базис призван обеспечить:

- единство основополагающей терминологии, определений и классификаций;
- использование на предприятии методологически проработанных подходов, моделей, инструментов при создании ПО;
- единую шкалу требований и критериев, определяющих качество процессов и технологий создания ПО и качество программной продукции;
- задание направлений совершенствования процессов и технологий разработки ПО в отрасли.

Примером успешно работающего аналога может служить международный стандарт DO-178 и его отечественная версия

ГОСТ Р 51904-2002 «Программное обеспечение встроенных систем. Общие требования к разработке и документированию», введенная в действие с июня 2002 года. Кстати, он поддерживается рядом инструментальных программных средств крупнейших производителей. Предусматривается сертификация процессов жизненного цикла БПО в авиации в соответствии с требованиями стандарта. Подобный же подход (несомненно, с весьма вдумчивым и постепенным его внедрением, со следованием принципу «не навреди»), как представляется, может быть применен и в космической отрасли. Можно привести еще пример «Стандарты кодированию на языке С++ при создании бортового ПО» истребителя пятого поколения F-35.

И, наконец, третьим направлением работ является создание, наполнение и развитие учебно-методической базы знаний в сфере создания БПО для обеспечения кадровой преемственности. Данная база знаний предназначена:

- для парирования кадровых проблем (естественный уход квалифицированных кадров старшего поколения и низкое качество подготовки молодых специалистов);
- для обеспечения необходимого уровня качества подготовки персонала и сокращения сроков такой подготовки;
- для аттестации персонала;
- для сохранения опыта, накопленного в отрасли за десятилетия (и предотвращения реальной угрозы его невосполнимой утраты) и пополнения этого опыта современными достижениями науки и практики;
- для обеспечения основы взаимодействия с системой высшего и специального образования.

Соответственно, можно выделить следующие виды работ по указанным направлениям:

- аналитическое – изучение отечественного и зарубежного опыта в области создания ПО в аэрокосмической отрасли, отслеживание перспективных тенденций и их оценка, изучение соответствующих процессов на предприятиях нашей страны;

- теоретическое – разработка формализованных методов и математических моделей, применимых на различных этапах жизненного цикла критического ПО;
- нормативное – разработка требований, рекомендаций и проектов стандартов, регламентирующих процессы разработки и этапы жизненного цикла БПО;
- технологическое – формулирование требований к средствам инструментального ПО их проектирование и разработка;
- информационно-методическое – создание электронной отраслевой базы знаний по теме разработки и использования ПО, планирование состава и содержания методического обеспечения подготовки специалистов соответствующего профиля для аэрокосмической отрасли.

Литература

1. *Авиастроение. Том 6.* (Итоги науки и техники, ВИНТИ АН СССР). – М., 1978.
2. БАЛТРУШАЙТИС В.В. «ГРАФИТ-ФЛОКС» – технология разработки программного обеспечения бортовых вычислительных машин // Тезисы докл. Межд. научно-техн. конф., посвященной 90-летию со дня рождения акад. Н.А. Пилюгина. – М.: НПЦ АП, 1998. – С. 79–81.
3. ДЕЙКСТРА Э.В. *Навстречу корректным программам.* – Grenoble: Mathematics Technological University, 1967.
4. ИЛЬИН И.А., ЕРМАКОВ И.Е., ТЮГАШЕВ А.А. *Создание единой отраслевой САПР ПО как основное условие предотвращения кризиса управляющего ПО в космической отрасли* // Актуальные проблемы российской космонавтики: Труды XXXVI Академических чтений по космонавтике, Москва, январь 2012 г. – С. 468.
5. КАЛЕНТЬЕВ А.А., ТЮГАШЕВ А.А. *ИППИ/CALS технологии в жизненном цикле комплексных программ управ-*

- ления. – Самара: Изд-во Самарского научного центра РАН, 2006.
6. КИРИЛИН А.Н., АХМЕТОВ Р.Н., СОЛЛОГУБ А.В., МАКАРОВ В.П. *Методы обеспечения живучести низкоорбитальных автоматических КА зондирования Земли: математические модели, компьютерные технологии.* – М.: Машиностроение, 2010. – 384 с.
 7. КОЛТАШЕВ А.А. *Эффективная технология управления циклом жизни бортового программного обеспечения спутников связи и навигации // Авиакосмическое приборостроение.* – 2006. – №12. – С. 20–25.
 8. ПАРОНДЖАНОВ В.Д. *Развитие системного обеспечения вычислительных комплексов при отработке систем управления ракет-носителей и космических аппаратов / В кн.: XXII научные чтения по космонавтике, посвященные памяти академика С.П. Королева и других выдающихся отечественных ученых – пионеров освоения космического пространства. Тезисы докл.* – М.: ИИЕТ РАН, 1998. – С.14–15.
 9. ТЮГАШЕВ А.А. *Интегрированная среда для проектирования управляющих алгоритмов реального времени // Известия российской академии наук: Теория и процессы управления.* – 2006. – №2. – С. 128–141.
 10. *Управление космическими аппаратами зондирования Земли: компьютерные технологии / Д.И. Козлов, Г.П. Аншаков, Я.А. Мостовой, А.В. Соллогуб.* – М.: Машиностроение, 1998. – 368 с.
 11. ХАРТОВ В. В. *Автономное управление космическими аппаратами связи, ретрансляции и навигации // Авиакосмическое приборостроение.* – 2006. – №6. – С. 29–33.
 12. BROOKSE F. *No Silver Bullet – Essence and Accident in Software Engineering // Proceedings of the IFIP 10th World*

Computing Conference, Dublin, Ireland, September 1-5,
1986. – P. 1069–1076.

WAYS TO IMPROVE QUALITY AND RELIABILITY OF SOFTWARE IN AEROSPACE INDUSTRY

Andrey Tyugashev, Samara State Aerospace University, professor
(tau7@ssau.ru).

Ilya Ilyin, TsNIIMASH (Korolyov, ilyailyin@yandex.ru).

Ilya Ermakov, NPO Tesla, Technical Director (ilya@ermakov.net.ru).

Abstract: The problem is considered of reliable control software development for Aerospace Industry. The state-of-the-art is described and possible ways to increase quality, reliability and safety of the critical software are proposed.

Keywords: control software, spacecraft, program verification, program testing.

*Статья представлена к публикации
членом редакционной коллегии М. В. Губко*