

Институт космических исследований РАН

На правах рукописи

БАХШИЯН Борис Цолакович

ИСПОЛЬЗОВАНИЕ МЕТОДОВ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ  
ДЛЯ РЕШЕНИЯ ОПТИМАЛЬНЫХ ЗАДАЧ  
ОЦЕНИВАНИЯ И КОРРЕКЦИИ

Диссертация на соискание ученой степени  
доктора физико-математических наук

(05.13.01 – системный анализ, управление и обработка информации)

Москва – 2001

# Оглавление

<b>Введение</b>	<b>7</b>
0.1 Общая характеристика работы . . . . .	7
0.2 Краткое содержание работы . . . . .	11
<b>1 Некоторые результаты в теории линейного оценивания</b>	<b>25</b>
1.1 Представление весовых матриц, определяющих заданную оценку наименьших квадратов. . . . .	25
1.1.1 Введение . . . . .	25
1.1.2 Некоторые сведения из теории матриц . . . . .	29
1.1.3 Основные результаты . . . . .	30
1.1.4 Пример . . . . .	37
1.1.5 О применении полученных результатов . . . . .	38
1.2 Выбор мешающих параметров в схеме линейной регрессии и множе- ство линейных несмещенных алгоритмов оценивания . . . . .	40
1.2.1 Модель оценивания . . . . .	40
1.2.2 Эквивалентность множеств всевозможных линейных несме- щенных оценок при различном выборе мешающих параметров	42
1.2.3 Эквивалентность множества всех оценок метода наименьших квадратов и линейных несмещенных оценок при различном выборе вектора мешающих параметров. . . . .	44
1.2.4 Ошибки линейного оценивания. . . . .	45
1.3 Вычисление гарантированных характеристик точности оценивания при наличии немоделируемых возмущений . . . . .	45

1.3.1	Метод наименьших квадратов и ошибка оценивания для линейного приближения . . . . .	49
1.3.2	Вычисление гарантированной ошибки линейного оценивания	56
<b>2</b>	<b>Простейшие задачи оптимального оценивания и коррекции и их сведение к задачам линейного программирования</b>	<b>61</b>
2.1	Классический и гарантирующий подходы к оптимизации оценщика, их преимущества и недостатки . . . . .	61
2.1.1	Классический подход к оптимизации оценщика и его практические недостатки . . . . .	62
2.1.2	Гарантирующий подход к вычислению точности оценивания.	63
2.2	Сравнение решений задач оптимального оценивания в двух простейших случаях при гарантирующем и классическом подходах . . . . .	65
2.2.1	Задача о выборе оптимального оценщика при возможности повторения измерений и ограничении на их общее число . . . . .	65
2.2.2	Оптимизация гарантированной дисперсии $D_1$ . . . . .	69
2.2.3	Минимаксная задача оценивания при ограниченных по модулю ошибках измерений. . . . .	70
2.3	Оптимальная задача линейной идеальной коррекции и обобщенное линейное программирование . . . . .	71
<b>3</b>	<b>Критерии оптимальности и монотонные алгоритмы решения вырожденной и обобщенной задач линейного программирования</b>	<b>75</b>
3.1	Теория решения вырожденной задачи линейного программирования	75
3.1.1	Введение . . . . .	75
3.1.2	Основные теоремы . . . . .	78
3.1.3	Описание алгоритма . . . . .	82
3.1.4	Эквивалентный критерий оптимальности и дополнения к алгоритму . . . . .	83
3.1.5	Практические результаты . . . . .	87
3.2	Обобщенная задача линейного программирования . . . . .	94

3.2.1	Виды обобщенных задач и соотношения между ними . . . . .	94
3.2.2	Критерий оптимальности для обобщенной задачи линейного программирования . . . . .	99
3.2.3	О сходимости алгоритма генерации столбцов. . . . .	100
3.2.4	Об эквивалентном критерии оптимальности и дополнениях к алгоритму . . . . .	101
<b>4</b>	<b>Теория и алгоритмы решения задач <math>L</math>- и <math>MV</math>-оптимального планирования эксперимента и их применение</b>	<b>102</b>
4.1	Задачи $L$ - и $MV$ -оптимального планирования и исторический комментарий . . . . .	102
4.1.1	Постановка задачи . . . . .	102
4.2	Сведение $L$ -задачи к задаче оптимальной линейной импульсной коррекции и алгоритм ее решения . . . . .	108
4.2.1	Необходимое и достаточное условие оптимальности $L$ -задачи	110
4.2.2	Получение оптимального плана с минимальным числом положительных компонент . . . . .	113
4.2.3	Алгоритм решения задачи (2.4) . . . . .	114
4.3	Сведение $MV_s$ -задачи к параметрической задаче оптимальной линейной импульсной коррекции и алгоритм ее решения . . . . .	116
4.3.1	Редукция задачи (3.2) к задаче многомерной максимизации и алгоритм ее решения . . . . .	118
4.4	Нахождение аналитического решения для случая для случая полиномиальной регрессии. . . . .	120
4.5	Оптимальное планирование лазерных наблюдений спутников ЛАГЕОС-1,2 . . . . .	127
4.5.1	Введение . . . . .	127
4.5.2	Математическая постановка задачи . . . . .	130
4.5.3	Проверка алгоритмов на примере определения координат полюса Земли . . . . .	131

4.5.4	Заключение . . . . .	136
<b>5</b>	<b>О возможности решения проблемы моментов методами линейного программирования и применение к задачам робастного и минимаксного оценивания</b>	<b>137</b>
5.1	Введение. . . . .	138
5.2	Решение проблемы моментов в случае минимизации опорной функции и применение к задаче минимаксного оценивания при наличии немоделируемых возмущений. . . . .	139
5.2.1	Обоснование алгоритма. . . . .	139
5.2.2	Решение минимаксной задачи с немоделируемыми возмущениями. . . . .	143
5.3	О решении задачи робастного оценивания. . . . .	147
5.3.1	Задача робастного оценивания. . . . .	147
5.3.2	Решение задачи робастного оценивания с помощью обобщенного линейного программирования. . . . .	148
5.4	Решение проблемы моментов в случае известной двойственной нормы и применение к задаче минимизации гарантированной дисперсии. . . . .	152
5.4.1	Обоснование алгоритма. . . . .	153
5.4.2	Минимизация гарантированной дисперсии. . . . .	155
5.5	Приложение. . . . .	157
<b>6</b>	<b>Об оптимальной линейной идеальной коррекции при ограничениях на корректирующие импульсы</b>	<b>159</b>
6.1	Постановка задачи и невозможность ее решения методом, используемым для задачи без ограничений на импульсы . . . . .	160
6.1.1	Постановка задачи . . . . .	160
6.1.2	Геометрическое описание алгоритма решения задачи (1.6) . . . . .	165
6.1.3	Аналитическое описание алгоритма решения задачи (1.6) и его обоснование . . . . .	166

6.1.4	Неприменимость описанного алгоритма для задачи обобщенной линейной импульсной коррекции с ограничениями . . . . .	169
6.2	Преобразование задачи с ограничениями к обобщенной задаче линейного программирования . . . . .	170
6.3	Алгоритм поиска оптимального базисного решения . . . . .	172
6.3.1	Описание алгоритма . . . . .	172
6.3.2	Решение подзадачи (3.1) в некоторых важных частных случаях	173
6.3.3	Дополнения к предлагаемому алгоритму . . . . .	175
6.4	Некоторые численные результаты . . . . .	176
6.4.1	Постановка задачи . . . . .	176
6.4.2	Плоская задача коррекции . . . . .	177
6.4.3	Пространственная коррекция . . . . .	180
	<b>Заключение</b>	<b>184</b>
	<b>Список литературы</b>	<b>185</b>

# Введение

## 0.1. Общая характеристика работы

**Актуальность работы.** Настоящая работа посвящена решению оптимальных задач определения и коррекции движения системы. Обе эти задачи тесно связаны между собой, являясь составными частями так называемого дискретного управления движением, при котором управляющие воздействия подаются не непрерывно, а в виде дискретных корректирующих импульсов, скачкообразно изменяющих характер движения управляемой системы. При этом каждой коррекции предшествует определение фактического движения, на основе которого вычисляется требуемое значение корректирующего импульса. Классическим примером подобного управления может служить коррекция орбиты космического аппарата с использованием корректирующего двигателя большой тяги. Подобный способ управления может быть использован при решении других прикладных задач. Кроме того, задачи определения движения различных реальных систем по результатам измерений имеют самостоятельное значение. Необходимость в их решении возникает при обработке данных наблюдений и результатов различных экспериментов, определении физических констант и т. п.

Методы решения рассматриваемых задач существенным образом зависят от принятых допущений об ошибках используемых исходных данных (математической модели движения и его коррекции, априорных сведений о параметрах этой модели, измерений). При этом в литературе обычно используется классический вероятностный подход, предполагающий знание характеристик распределений вероятностей ошибок исходных данных. В работах других исследователей, в том числе и автора, используется также гарантирующий подход, при котором

функции распределения вероятностей ошибок исходных данных считаются неизвестными, а задаются лишь некоторые множества, которым могут принадлежать указанные функции. При этом отыскивается способ решения рассматриваемых задач, гарантирующий достижение поставленной цели при наихудшем возможном распределении ошибок исходных данных. Оптимизация этого способа приводит к минимаксной или максиминной задачам и соответствующий подход называют минимаксным.

При решении многих задач рассматриваемого типа влияние сколь угодно малых отклонений от принятого распределения вероятностей ошибок неограниченно возрастает с увеличением числа используемых измерений. Это явление обычно называют неустойчивостью получаемых результатов по основным допущениям. Поэтому результаты, найденные на основе гарантирующего подхода, в ряде случаев оказываются значительно ближе к реальным условиям решения прикладных задач, чем выводы, получаемые при вероятностном подходе.

В связи с изложенным актуальными являются задачи оптимального выбора алгоритма оценивания и программы коррекции при каждом из указанных подходов. При использовании гарантирующего подхода для широкого класса таких задач оценивания оптимальным является линейный алгоритм. При этом минимаксную задачу часто удается свести к вырожденным или обобщенным задачам линейного программирования. С другой стороны, рассмотренные ниже задачи  $L$ - и  $MV$ -оптимального планирования эксперимента, задача робастного оценивания при неточно заданной функции распределения, а также оптимальная задача идеальной линейной коррекции с ограничениями на импульсы также сводятся к обобщенным задачам линейного программирования. При их решении возникают следующие проблемы.

1. Обобщенная задача линейного программирования уже не всегда решается так же эффективно, как обычная задача, поскольку проверка условий оптимальности представляет собой отдельную подзадачу. Если эта подзадача решается достаточно просто, например, аналитически, то и обобщенная задача решается так же эффективно, как и обычная задача линейного программирования тех же раз-



меров. В диссертации получено решение указанных подзадач, соответствующих ряду важных задач оценивания и коррекции параметров.

2. В процессе решения как обобщенной, так и обычной задачи линейного программирования могут возникать (и возникают, как показывает опыт) большое количество вырожденных итераций, при проведении которых целевая функция не изменяется. Это связано с тем, что для вырожденного плана обычно используемые достаточные условия оптимальности не являются, вообще говоря, необходимыми. При этом резко снижается эффективность симплекс-метода. Для обобщенной задачи появление таких итераций является регулярным случаем и может быть причиной сходимости к неоптимальному значению функционала. Вместе с тем, до последнего времени не был разработан алгоритм, основанный на критерии оптимальности. Решение этого вопроса и создание теории линейного программирования, включающего вырожденный случай, проведено в диссертации.

**Целью работы** является развитие методов линейного программирования и решение с их помощью оптимальных задач оценивания и коррекции параметров системы, а также получение на их основе некоторых качественных и аналитических результатов.

**Новизна работы.** В работе впервые создана теория линейного программирования (в том числе и обобщенного), основанная на необходимых и достаточных условиях оптимальности и пригодная для вырожденного случая. Впервые показано, что  $L$ -задача оптимального планирования эксперимента может быть сведена к задаче идеальной линейной коррекции и эффективно решена как обобщенная задача линейного программирования. Впервые получен критерий оптимальности для  $MV$ -задачи, и с помощью его найдены алгоритм решения в общем случае и аналитическое решение для случая полиномиальной регрессии. Данные результаты использованы для оптимального выбора программы измерений при оценивании геодинимических параметров Земли. Впервые показано, что задача робастного оценивания с линейными ограничениями на оцениваемые параметры может быть эффективно решена как обобщенная задача линейного программирования. То же самое показано для некоторых минимаксных задач оценивания

и для задачи оптимальной идеальной линейной коррекции с ограничениями на корректирующие импульсы. Проведены массовые расчеты для спутника Земли и показано, что оптимальная программа коррекции при достаточно малом пороге, ограничивающем величины импульсов, очень хорошо моделирует оптимальную непрерывную коррекцию двигателя с малой тягой. В работе также развита теория линейного несмещенного оценивания – установлено точное описание множества весовых матриц метода наименьших квадратов, определяющих заданную несмещенную оценку, и обобщены теоремы эквивалентности, позволяющие в схеме с мешающими параметрами рассматривать только линейные оценки при условиях несмещенности и без мешающих параметров.

**Методы исследования.** В работе используются методы математического программирования, линейной алгебры, матричного анализа, теории вероятностей, теоретической астродинамики.

**Практическая ценность** полученных результатов состоит в том, что они являются основой для создания эффективного программного обеспечения на ЭВМ. Для всех рассмотренных в диссертации задач построены эффективные алгоритмы, которые реализованы на ЭВМ при решении конкретных задач навигации космических объектов (в частности, для проектов "Вега" и "Лагос").

**Апробация работы.** Основные результаты диссертационной работы докладывались и обсуждались на

- Всесоюзном совещании-семинаре "Проблемы оптимизации и управления динамическими системами" (Владивосток, 1987);
- семинаре центрального экономико-математического института под рук. Е.Г.Гольштин (Москва, 1987);
- семинаре института математики и механики УНЦ СССР (г.Свердловск, 1988);
- VII Всесоюзной конференции "Управление в механических системах" (Свердловск, 1990);

- III Всесоюзной школе по навигации и управлению движущимися объектами (Феодосия, 1990);
- Всероссийской конференции "Общие проблемы управления и их приложения к математической экономике" (Тамбов, Россия, 2000 г.);
- научном семинаре "Механика, управление и информатика" ИКИ РАН (Москва, 2000);
- семинарах под руководством акад. Ф.Л. Черноусько (ИПМех РАН); акад. Я.З. Цыпкина (ИПУ РАН); профессоров В.Н. Афанасьева, В.Б. Колмановского, В.Р. Носова (МИЭМ); проф. А.И. Кибзуна (МАИ).

По теме диссертации опубликованы 24 печатные работы, в том числе одна монография. Результаты диссертационной работы являются основой проектов "Решение оптимальных задач оценивания и коррекции параметров системы методами линейного программирования" (поддержан РФФИ, проект N ° 95-01-00807) и "Разработка эффективных методов решения задач планирования эксперимента и коррекции движения и их использование в космической навигации" (поддержан РФФИ, проект N ° 98-01-00384), в которых автор был научным руководителем.

**Объем и структура работы.** Диссертация состоит из введения, шести глав, заключения и списка литературы. Общий объем работы - 260 м.п.с., напечатанных в текстовом редакторе TeX. Библиография - 102 названия. Текст содержит 14 рисунков, 9 таблиц.

## 0.2. Краткое содержание работы

**Во введении** описано состояние проблемы, обоснована актуальность тематики диссертационной работы, изложены основные идеи диссертации, приведена общая характеристика работы.

**В главе 1** рассматриваются некоторые теоретические аспекты используемой далее теории линейного оценивания, которые применяются в дальнейших

исследованиях и имеют самостоятельное методологическое значение. Пусть модель измерений есть

$$y = H^T \theta + \xi,$$

где векторы  $y$  и  $\xi \in \mathbb{R}^n$  — соответственно векторы измерений и ошибок,  $\theta \in \mathbb{R}^m$  — вектор оцениваемых параметров,  $H$  — известная матрица. Пусть  $L = C\theta$  — векторная линейная функция. Здесь  $C$  — известная матрица. Рассмотрим метод наименьших квадратов для определения оценок  $\hat{\theta}, \hat{L}$ :

$$\begin{aligned} \hat{\theta} &= \arg \min_{\hat{\theta}} (y - H^T \hat{\theta})^T W (y - H^T \hat{\theta}), \\ \hat{L} &= C \hat{\theta} = C (HWH^T)^{-1} HWy, \end{aligned}$$

где  $W$  — заданная неотрицательно определенная весовая матрица, дающая оценку наименьших квадратов. Как видно, каждой весовой матрице  $W$  соответствует линейный алгоритм, удовлетворяющий условию его несмещенности: если  $\xi = 0$ , то  $\hat{L} = L$  при любых  $\theta$ . В разделе 1.1 рассмотрена обратная задача: найти множество всех неотрицательно определенных матриц  $W$ , дающих линейную оценку  $\hat{L} = Xy$  ( $X$  —  $k \times n$ -матрица), удовлетворяющую указанному условию несмещенности, которое имеет вид  $XH^T = C$ . Другими словами отыскиваются все неотрицательно (или положительно) определенные решения матричного уравнения

$$X = C (HWH^T)^{-1} HW.$$

Задача нахождения этого множества рассматривалась различными авторами. Нами дано конкретное и наиболее подробное описание этого множества и проанализированы различные случаи подбора весовой матрицы с наибольшим количеством нулей вне диагонали. Указаны возможности применения полученных выражений при оценивании параметров системы.

На практике при оценивании параметров обычно некоторый подвектор  $\theta_2$  вектора  $\theta$  принимают равным своему априорному значению  $\tilde{\theta}_2$  ( $\theta_2 \neq \theta$ ), а оцениваемым вектором считают совокупность  $\theta_1$  остальных компонент вектора  $\theta$ . При указанном подходе вектор  $\theta_2$  называют мешающим параметром. Согласно разделу 1.2, справедливы следующие результаты: в силу теоремы 1.3, если

априорное значение оцениваемого вектора включить в число измерений, то множество линейных несмещенных оценок, получаемое при всевозможных матрицах  $X_1$ , удовлетворяющих условию несмещенности, и соответствующее некоторому выбору вектора  $\theta_2$  мешающих параметров, не зависит от этого выбора, совпадая тем самым с множеством линейных оценок, соответствующим отсутствию вектора мешающих параметров, т.е. случаю  $\theta_1 = \theta$ ; в силу теоремы 1.4 такое же утверждение имеет место для метода наименьших квадратов с мешающими параметрами. Эти результаты имеют более общий вид, чем аналогичные теоремы (так называемые теоремы эквивалентности), полученные другими авторами. Указанные выводы позволяют при построении оптимальных алгоритмов оценивания анализировать лишь линейные несмещенные алгоритмы.

Во вспомогательном разделе 1.3 приведены выражения для характеристик ошибок линейного оценивания при наличии действующих на систему сил, не учитываемых в модели оценивания.

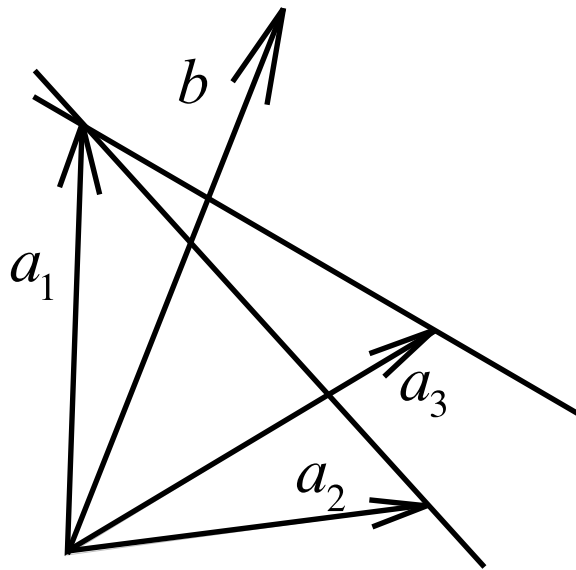
**В главе 2** рассмотрены некоторые простые задачи оптимального оценивания и коррекции траектории, сводящиеся к обычной или обобщенной задаче линейного программирования, для которых могут быть даны ясные интерпретации решения и получены важные методологические выводы. Более общие задачи рассмотрены в главах 4 – 6. В разделе 2.1 рассмотрен классический вероятностный и гарантирующий подходы при решении оптимальных задач оценивания. Показано, что классический подход к вычислению точности оценивания может давать слишком оптимистические результаты. Приведены выражения для гарантирующих характеристик точности, когда средние и ковариации лежат в заданных параллелепипедах. В разделе 2.2 показано, что задача  $C$ -оптимального оценивания (т.е. оценивания скалярного параметра) в схеме линейной регрессии при возможности повторения измерений (задача планирования эксперимента) сводится к той же задаче линейного программирования, что и задача минимизации гарантирующей дисперсии, полученная при возможности произвольной корреляции между измерениями, т.е. оптимальные программы измерений совпадают для наиболее оптимистического и пессимистического допущений. Этот результат удивителен и

методологически значим.

В разделе 2.2 рассматривается задача оптимальной идеальной линейной коррекции без ограничений. Затраты на каждую коррекцию есть некоторая норма корректирующего импульса, для которого известна двойственная норма. Показывается, что задача может быть решена, как и в известном случае евклидовой нормы, методом генерации столбцов для обобщенного линейного программирования. При этом число импульсов не превосходит числа корректируемых параметров.

**В главе 3** получены критерии оптимальности для вырожденной и обобщенной задач линейного программирования и на их основе разработаны монотонные алгоритмы решения этих задач. Обычно считается, что решение задач линейного программирования симплексным методом всегда эффективно. Именно поэтому мы и стремились свести оптимальные задачи оценивания и коррекции траектории к задачам линейного программирования. Однако если вектор правых частей  $b \in \mathbb{R}^m$  задачи линейного программирования всегда есть линейная комбинация векторов условий с положительными коэффициентами и число  $k$  этих векторов равно  $m$ , то при отсутствии вырожденных базисов решение достигается за конечное число итераций, так как достаточное условие оптимальности является и необходимым. Геометрически для случая  $m = 2$ , изображенного на рис.1, невырожденность задачи означает, что вектор  $b$  не совпадает ни с одним из векторов  $a_i$ .

При  $k < m$  имеет место вырожденный случай (рис.2). При этом использование обычного симплексного метода может привести к большому числу вырожденных итераций, при которых целевая функция не изменяется, или даже к возникновению цикла. Действительно, при вводе в базис согласно симплексной итерации вектора  $a_3$ , значение целевой функции не изменяется и это может повторяться много раз. Обобщенная задача линейного программирования характеризуется континуумом векторов условий; при этом вырожденный случай не является исключительным, а число вырожденных итераций может стать бесконечным (рис.3). Предлагаемые алгоритмы выхода из вырожденной итерации основаны на некотором (иногда виртуальном) изменении целевой функции. Наиболее эффек-

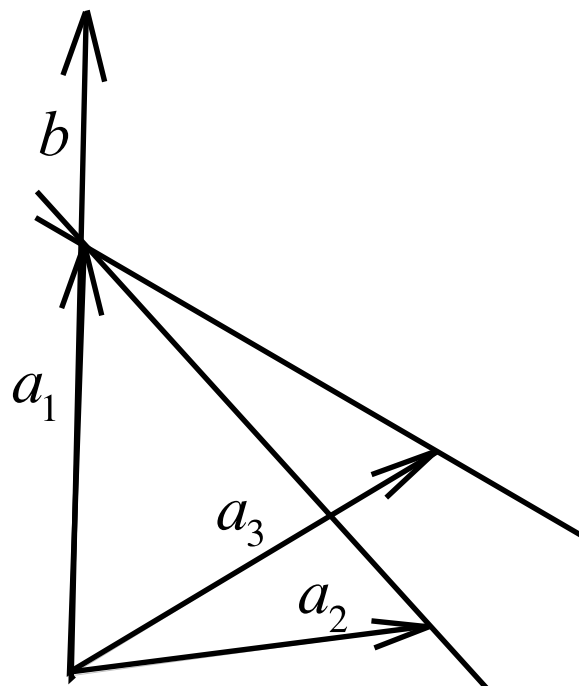


$$\min \sum x_i$$

$$\sum x_i a_i = b$$

$$x_i \geq 0$$

Рис.1. Симплекс-метод для невырожденной задачи



$$\min \sum x_i$$

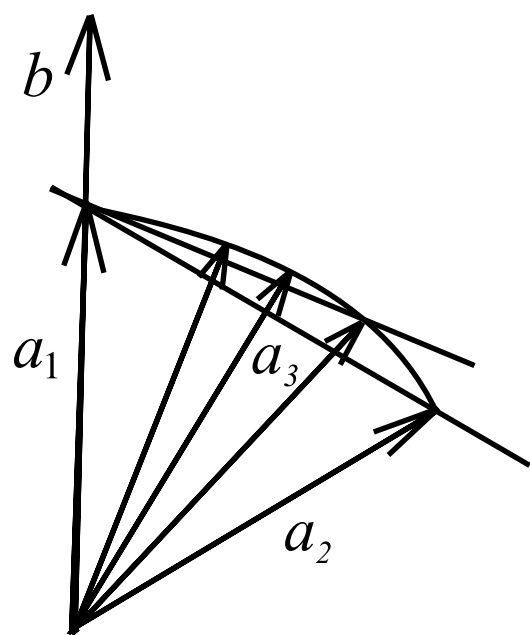
$$\sum x_i a_i = b$$

$$x_i \geq 0$$

$$b = k \cdot a_1$$

$$k > 0$$

Рис.2. Симплекс-метод для вырожденной задачи



$$\min \sum x_i$$

$$\sum x_i a_i = b$$

$$x_i \geq 0$$

$$a_i \in \mathcal{A}$$

Рис.3. Симплекс-метод для вырожденного случая в обобщенной задаче линейного программирования

тивным из этих алгоритмов является, по видимому, алгоритм Вулфа, который, однако, не применим для обобщенной задачи линейного программирования.

В разделе 3.1 нами получен критерий оптимальности задачи линейного программирования, заключающийся в следующем. Пусть задача линейного программирования имеет вид

$$\min_x \left\{ c^T x : \sum_{i=1}^n x_i a_i = b, \quad x = (x_1, \dots, x_n)^T \geq 0 \right\}$$

( $c \in \mathbb{R}^n$ ,  $b, a_i \in \mathbb{R}^m$ ) и текущий базисный вектор  $x$  содержит  $k$  ненулевых компонент, причем  $k < m$ . Пусть  $\mathcal{I}_+$  – множество индексов этих ненулевых компонент. Множество  $\{a_i, i \in \mathcal{I}_+\}$  назовем строгим базисом. Пусть  $\mathcal{U}$  – подпространство, порожденное векторами строгого базиса,  $\mathcal{V}$  – некоторое подпространство, такое что  $\mathbb{R}^m = \mathcal{U} \oplus \mathcal{V}$ . Например, можно взять в качестве  $\mathcal{V}$  ортогональное дополнение к  $\mathcal{U}$  в  $\mathbb{R}^m$ .

Согласно теореме 3.1, вектор  $x$  оптимален тогда и только тогда, когда имеет конечный оптимум задача линейного программирования с  $n - k$  переменными



и  $m - k$  независимыми ограничениями-равенствами

$$\min_{y_i} \left\{ \sum_{i \in \mathcal{I}_0} \Delta_i y_i : \sum_{i \in \mathcal{I}_0} y_i v_i = d, \quad y_i \geq 0, \quad i \in \mathcal{I}_0 \right\},$$

где  $\mathcal{I}_0 \doteq \{1, \dots, n\} \setminus \mathcal{I}_+$ ;  $\Delta_i = c_i - \pi^T a_i$ ;  $\pi$  – любое решение системы уравнений  $c_i - \pi^T a_i = 0$ ,  $i \in \mathcal{I}_+$ ; векторы  $v_i \in \mathcal{V}$  определяются разложением  $a_i = u_i + v_i$ ,  $u_i \in \mathcal{U}$ ;  $d \in \mathcal{V}$  – любой вектор, для которого совместны ограничения во вспомогательной задаче. При случайном выборе  $d$  вспомогательная задача не вырождена с вероятностью 1. Число ограничений-равенств задачи линейного программирования будем называть размерностью этой задачи, так как оно есть порядок обращаемой матрицы и определяет в основном эффективность решения задачи. Таким образом, для проверки условия оптимальности нужно решить невырожденную вспомогательную задачу размерности меньшей исходной. При  $k = m$ , т.е. в частном случае невырожденного базисного решения, условия теоремы 3.1 превращаются в известные достаточные условия оптимальности:  $\Delta_i \geq 0 \forall i$ .

Если критерий оптимальности не выполняется, то согласно теоремам 3.2 и 3.3 можно уменьшить целевую функцию путем изменения строгого базиса  $\{a_i, i \in \mathcal{I}_+\}$ . А именно, согласно теории линейного программирования для невырожденного случая, в процессе решения вспомогательной задачи симплекс-методом находятся линейно независимые векторы  $a_i$ ,  $i \in \mathcal{S}$ , такие что при некоторых  $\lambda_i$  выполняются условия

$$\sum_{i \in \mathcal{S}} \lambda_i v_i = 0, \quad \sum_{i \in \mathcal{S}} \Delta_i \lambda_i < 0; \quad \lambda_i > 0, \quad i \in \mathcal{S}.$$

Здесь  $\mathcal{S}$  – некоторое множество индексов из  $\mathcal{I}_0$ ,  $|\mathcal{S}| \leq m - k + 1$ . Эти векторы вводятся в строгий базис. Выводимые векторы определяются следующим образом. Первое из вышеприведенных условий эквивалентно равенству

$$\sum_{j \in \mathcal{S}} \lambda_j a_j = \sum_{i \in \mathcal{I}_+} \mu_i a_i,$$

которое означает, что вектор  $\sum_{j \in \mathcal{S}} \lambda_j a_j$  однозначно разлагается по векторам строгого базиса. Тогда из строгого базиса выводятся векторы  $a_j$ ,  $j \in \mathcal{R}$ , где  $\mathcal{R} = \{r : x_r / \mu_r = \theta\}$ ,  $\theta \doteq \min_j \{x_j / \mu_j : \mu_j > 0\}$ .

Алгоритм метода, обоснованного теоремами 3.1 — 3.3, зависит от выбора матрицы  $V$ , дополняющей матрицу строгого базиса  $U$  до базисной матрицы  $B$ . В пункте 3.1.3 описан алгоритм для случая, когда  $B$  найдена в результате обычных симплексных итераций, приведших к текущему вырожденному решению. Указан другой способ выбора базисной матрицы, особенно эффективный при  $k \ll m$ .

В пункте 3.1.4 дан эквивалентный критерий оптимальности текущего базисного решения, с помощью которого найдены полезные оценки минимума, которые позволяют повысить эффективность алгоритма.

Таким образом, приведенная теория линейного программирования и соответствующий симплексный алгоритм естественно обобщают классическую теорию и алгоритм, известные ранее для невырожденного случая. В пункте 3.1.5 показано, что новый алгоритм может позволить решать большие вырожденные задачи линейного программирования, что практически невозможно сделать методами, основанными на достаточных условиях оптимальности. Алгоритм опробован на примере решения задачи составления расписания авиаперевозок, для которой задача линейного программирования имеет 80 строк и 3135 столбцов. Показано, что новый алгоритм в большинстве случаев не уступает алгоритму Вулфа, и становится эффективнее его при небольших значениях  $(m - k)/m$ .

В разделе 3.2 развиты некоторые аспекты теории обобщенной задачи линейного программирования. Записаны различные формы этой задачи. Выявлены условия, при которых обобщенная задача может быть решена методом генерации столбцов. Показано, что эффективность решения обобщенной и обычной задач линейного программирования равной размерности не отличаются, если легко решается подзадача проверки условия оптимальности.

Рассмотрен вырожденный случай, который для обобщенной задачи является регулярным. Для нее выписан аналогичный выведенному в разделе 3.2 критерий оптимальности, сводящийся к решению вспомогательной обобщенной задачи линейного программирования размерности  $m - k$ . Однако, в отличие от обычной задачи линейного программирования, эта вспомогательная задача уже не может

быть, вообще говоря, сделана невырожденной путем выбора вектора  $d$  правых частей ограничений. Мы можем лишь решать вспомогательную задачу тем же методом, прибегая в случае вырожденности вспомогательной задачи к рекурсии и понижая порядок возникающих вырожденных задач линейного программирования.

Установлены условия сходимости алгоритма по функционалу в следующем смысле. Показано, что если алгоритм сходится к неоптимальному допустимому базисному решению, то это — вырожденное решение. Если взять это решение за текущее, то, согласно изложенному выше алгоритму, можно строго уменьшить целевую функцию на конечную величину.

**Глава 4** посвящена разработке теории и алгоритмов решения задач  $L$ - и  $MV$ -оптимального планирования эксперимента и их применению. Рассмотрены классические задачи планирования эксперимента, в которых в схеме линейной регрессии с  $m$  неизвестными параметрами имеется возможность повторения измерений одной и той же функции параметров при заданном общем числе измерений. Нормированное количество повторных измерений есть вероятностная мера, которая называется непрерывным планом эксперимента и принадлежит симплексу (при этом пренебрегается целочисленность числа повторных измерений). В зависимости от критерия оптимальности рассматриваются две задачи нахождения оптимального плана при использовании метода наименьших квадратов:

$L$ -задача — минимизируется сумма дисперсий  $s$  оценок контролируемых параметров из  $m$  неизвестных параметров,

$$s \leq m ;$$

$MV$ -задача — минимизируется максимальная из этих дисперсий.

Задачи ставятся путем использования вариационного подхода, при котором метод наименьших квадратов заменяется в соответствии с теоремой Гаусса-Маркова оптимальным линейным несмещенным оцениванием. При этом спецификой решения этих задач является то обстоятельство, что не каждый план измерений из симплекса является допустимым для несмещенного линейного оценивания.

Это обстоятельство ранее не учитывалось, что послужило не вполне строгому рассмотрению задачи рядом авторов.

В разделе 4.1 дана строгая постановка задачи. В разделе 4.2 показано, что  $L$ -задача сводится к задаче оптимальной линейной коррекции  $ms$  параметров, эффективно решаемой методом генерации столбцов. При этом оптимальный план, получаемый из решения обобщенной задачи линейного программирования, содержит не более  $ms$  ненулевых компонент. В то же время известно, что существует оптимальный план, содержащий не более чем  $m_0 = ms - s(s - 1)/2$  компонент. Показано, что такой план можно получить из оптимального плана с  $ms$  компонентами путем решения обычной задачи линейного программирования с  $m_0$  ограничениями-равенствами. В этом же разделе получены необходимые и достаточные условия оптимальности плана.

В разделе 4.3 рассмотрена  $MV$ -задача. Она представлена как минимаксная задача, где максимум берется по дополнительно введенному вектору  $\mu$  размерности  $s$ , принадлежащему симплексу  $\Sigma_s$ . Показано, что минимаксную задачу можно представить в виде максиминной. Это позволило получить необходимые и достаточные условия оптимальности  $MV$ -задачи, а также интерпретировать эту задачу как проектную задачу оптимальной линейной импульсной коррекции, в которой оптимальный корректируемый вектор  $b(\mu^*)$  отыскивается из условия, что минимальные затраты на коррекцию при каждом  $\mu$  будут максимальными среди всех  $\mu$  из симплекса. Это позволило предложить покоординатный спуск по  $\mu$ , причем на каждом шаге значение целевой функции не убывает.

В разделе 4.4 при помощи построенной теории найдено аналитическое решение  $MV$ -задачи для случая полиномиальной регрессии при  $m \leq 11$ . При этом минимизируется максимальный коэффициент при степенях полинома (т.е.  $s = m$ ). Оказывается, что при  $m \neq 5$  оптимальный план совпадает с оптимальным планом для задачи оценивания скалярного параметра, имеющего максимальную дисперсию. При  $m = 5$  две дисперсии имеют одинаковое максимальное значение и аналитическое решение получено при помощи найденного нами критерия оптимальности.

В разделе 4.5 решаются практически важные задачи об оптимальном по  $L$ - и  $MV$ -критериям использовании наблюдательных пунктов при уточнении геодинамических параметров Земли по наблюдениям ее искусственного спутника сетью станций. При этом использование оптимальной методики позволяет уменьшить дисперсии определяемых значений координат полюса Земли на 30-40%.

**В главе 5** рассмотрен класс задач, в которых в конечном евклидовом пространстве минимизируется некоторая норма  $p(x)$  при линейных ограничениях (проблема моментов). Постановки задач даны в разделе 5.1. В разделе 5.2 рассмотрен случай, когда норма может быть представлена как опорная функция некоторого симметричного относительно нуля выпуклого компакта  $\Gamma$ , т.е.

$$p(x) = \max \{x^T \gamma : \gamma \in \Gamma\} = x^T \gamma(x),$$

и явно или достаточно просто находится функция  $\gamma(x)$ , реализующая этот оптимум. Показано, что получаемая минимаксная задача может быть сведена к обобщенной задаче линейного программирования с  $n + 1$  ограничениями-равенствами и  $n + m$  переменными ( $n$  — число переменных оптимизации,  $m$  — число линейных ограничений в исходной задаче). Доказано, что если  $(\pi^T \in \mathbb{R}^n, \rho \in \mathbb{R}^1)$  — текущий двойственный вектор, то существует оценка для оптимума:  $\rho \leq p(x^*) \leq p(\pi)$  и метод генерации столбцов улучшает эту оценку на каждой итерации (при отсутствии вырождения). В качестве приложения теории показана возможность эффективного решения минимаксной задачи оценивания с немоделируемыми возмущениями.

В разделе 5.3 рассматривается задача робастного оценивания  $s$ -вектора  $\theta$  параметров по  $n$  измерениям  $z_i = H_i^T \theta + \xi_i, i = 1, \dots, n$ , при  $m$  линейных ограничениях на оцениваемые параметры

$$A\theta \geq b,$$

где  $b$  — заданный  $m$ -вектор,  $A$  —  $m \times s$  матрица. В теории робастного оценивания предполагается, что ошибки измерений есть независимые между собой случайные величины, функция распределения каждой из которых представляется в виде

выпуклой комбинации двух функций распределения:  $F_\varepsilon(x) = (1 - \varepsilon)\Phi(x) + \varepsilon G(x)$ , где  $\varepsilon$  – некоторое положительное число, не большее единицы,  $\Phi(x)$  – функция стандартного нормального распределения, вид функции  $G(x)$  распределения вероятности не известен. Для такой модели в предположении, что  $n$  достаточно велико и может быть применен закон больших чисел, Хьюбером обоснован устойчивый к возможно большим ошибкам измерений (робастный) метод оценивания. В этом методе оценка  $\hat{\theta}$  вектора  $\theta$  определяется следующим образом:

$$\theta = \arg \min_{\theta} \{F(\theta) : A\theta \geq b\},$$

где  $F(\theta) = \sum_{i=1}^n f(x_i(\theta))$ ,  $x_i(\theta) = z_i - H_i^T \theta$ , функция  $f(x)$  имеет вид

$$f(x) \doteq \begin{cases} x^2/2 & , \quad |x| \leq C, \\ |x|C - C^2/2 & , \quad |x| > C, \end{cases}$$

а величина  $C$  однозначно определяется по  $\varepsilon$ . При этом целевая функция может быть представлена уже как максимум квадратичной функции от  $\gamma$ :

$$F(\theta) = \max_{\gamma} \left\{ x(\theta)^T \gamma - \frac{1}{2} \gamma^T \gamma : |\gamma_i| \leq C, i = 1, \dots, n \right\},$$

где  $x(\theta)$  – вектор с компонентами  $x_i(\theta)$ . Тем не менее, это позволило свести задачу к обобщенной задаче с  $s + 1$  ограничениями и  $n + s$  переменными (число ограничений  $m$  не влияет на размеры задачи). При этом, аналогично предыдущему разделу, имеет место оценка для оптимума  $\rho \leq F(\hat{\theta}) \leq F(\pi)$ , где  $\pi \in \mathbb{R}^s$  и скаляр  $\rho$  — компоненты двойственного вектора на текущей итерации метода генерации столбцов. Вычисления показали, что этот метод при не слишком малых  $\varepsilon$  превосходит обычно используемый итеративный метод уже при  $s \geq 4$ . Кроме того, на эффективность симплексного алгоритма практически не влияет число ограничений  $m$ .

В разделе 5.4 рассмотрена задача минимизации нормы  $p(x)$  с известной двойственной нормой  $p^*(\gamma) \doteq \max\{\gamma^T \alpha : p(\alpha) = 1\}$ . Таким образом, решается задача раздела 5.2 при условии, что можно взять  $\Gamma = \{\gamma : p^*(\gamma) \leq 1\}$ . Тогда задача сводится к задаче обобщенного линейного программирования с  $m$  ограничениями-равенствами, что при  $m \ll n$  позволяет существенно улучшить

алгоритм. На каждом шаге также имеет место оценка для оптимума, а проверка условия оптимальности сводится к вычислению двойственной нормы  $p^*(A^T\pi)$ , где  $A$  — матрица линейных ограничений  $Ax \geq b$ . Показано, что к этому случаю сводится практически важная задача гарантирующего оценивания при условии, что все измерения разбиты на несколько групп, так что измерения внутри каждой группы могут быть произвольно коррелированы, а измерения из разных групп не коррелированы.

**В главе 6** задача идеальной линейной коррекции рассмотрена при дополнительных условиях принадлежности корректирующих импульсов  $u_i$  выпуклым множествам  $\mathcal{U}_i$  :

$$L = \min_{u_i} \left\{ \sum_{i=1}^n p_i(u_i) : \sum_{i=1}^n B_i u_i = b, \quad u_i \in \mathcal{U}_i \right\}.$$

Здесь  $u_i \in \mathbb{R}^{n_i}$  — векторы корректирующих импульсов, а убывающие положительные функции  $p_i(u_i)$  характеризуют энергетические затраты на исполнение  $i$ -го импульса. Фактически это есть задача минимизации нормы из определенного класса при линейных ограничениях и дополнительных нелинейных условиях.

В разделе 6.1 дана постановка задачи и показана невозможность ее решения методом, используемым для задачи без ограничений на импульсы. Здесь также теория, изложенная в главе 2, развита на случай, когда целевая функция есть произвольная норма.

В разделе 6.2 задача с ограничениями сведена к обобщенной задаче линейного программирования с  $m + 1$  ограничениями-равенствами ( $m$  — размерность корректируемого вектора). Показано, что если импульсы ограничены по норме  $p(\cdot)$ , т.е.

$\mathcal{U}_i = \{u_i \in \mathbb{R}^{n_i} : p_i(u_i) \leq c_i, \quad i = 1, \dots, n\}$ , то решение подзадачи сводится к вычислению двойственной нормы  $p_i^*(\cdot)$ , которая для известных приложений ( $p_i(x) = \|x\|$  — евклидова норма или  $p_i(x) = \|x\|_1$ ) вычисляется по известным формулам.

В разделе 6.4 приведены численные результаты выбора оптимальной коррекции на одном витке кеплеровской орбиты спутника Земли. В качестве корректируемых параметров выбраны следующие: радиусы апогея и перигея  $r_a$  и

$r_\pi$ , аргумент перигея  $\omega$ , наклонение  $i$ , долгота восходящего узла  $\Omega$ . Возможные моменты приложения импульсов распределяются равномерно по эксцентрической аномалии, а все нормы векторов в задаче — евклидовы и ограничены одной константой  $c$ . При достаточно большом значении  $c$  количество ненулевых импульсов не превосходит числа корректируемых параметров. При уменьшении величины  $c$  количество импульсов увеличивается до тех пор, пока при некотором критическом значении этой константы задача коррекции не перестанет иметь решения из-за несовместности ограничений этой задачи. Это означает возможность проведения непрерывной коррекции с двигателем малой тяги, столь же эффективной, как и импульсная коррекция, которая моделирует коррекцию с двигателем большой тяги на малом интервале времени.

**В заключении** отражены основные результаты диссертационной работы.

#### **Замечания.**

Глава 3 является основополагающей для диссертации. Однако некоторые результаты и определения из этой главы мы для удобства повторяем применительно к конкретным задачам почти в каждой главе. Это практически не увеличило объема работы.

Нумерация формул, теорем, лемм, замечаний в работе двойная, отдельная для каждой главы. Нумерация рисунков — сквозная.



## Глава 1

# Некоторые результаты в теории линейного оценивания

### 1.1. Представление весовых матриц, определяющих заданную оценку наименьших квадратов.

#### 1.1.1. Введение

Пусть нужно получить по данным измерений оценку  $\hat{L}$  некоторого вектора  $L \in \mathbb{R}^k$  параметров системы для случая, когда модель оценивания является линейной, т.е.

$$z = A\theta + \xi, \quad L = C\theta, \quad (1.1)$$

где  $z \in \mathbb{R}^n$  — вектор результатов измерений;  $\theta \in \mathbb{R}^m$  — вектор неизвестных параметров;  $\xi$  — вектор ошибок измерений;  $A$  и  $C$  — заданные матрицы размерности  $(n \times m)$  и  $k \times m$  соответственно.

Рассмотрим случай, когда оценка  $\hat{L}$  находится методом наименьших квадратов, в котором из минимизации некоторой неотрицательно определенной формы компонент вектора  $Z - A\theta$  определяется оценка  $\hat{\theta}$  вектора  $\theta$ . Матрицу этой квадратичной формы будем называть весовой матрицей.

Нас будет интересовать описание множества всех неотрицательно и положительно определенных весовых матриц, каждой из которых при известном  $d$

соответствует одна и та же для всех этих матриц оценка метода наименьших квадратов. Ниже мы покажем, что описание множества положительно определенных весовых матриц, обладающих указанным свойством, для случая  $C = A$  можно получить из [94], а для несколько более общего случая равенства рангов матриц  $A$  и  $C$  — из [97, 98]. Однако мы предпочли не использовать результаты указанных работ для нахождения искомого представления в общем случае, а привести иной вывод. Кроме того, мы рассмотрим возможности практического применения полученных результатов и связанные с этим вопросы, например о нахождении весовых матриц с максимальным числом нулевых элементов вне диагонали.

**Формулировка задачи.** Указанная выше оценка метода наименьших квадратов находится по формуле

$$\hat{L} = C\hat{\theta}, \quad \hat{\theta} \doteq \arg \min \{(d - A\theta)^T W(d - A\theta)\}; \quad (1.2)$$

где  $W \geq 0$  — заданная матрица. Пусть выполнено условие

$$A^T W A > 0, \quad (1.3)$$

из которого следует, что  $R(A) = m \Rightarrow m \leq n$  [96]. Тогда метод наименьших квадратов приводит к линейному алгоритму [96]

$$\hat{L} = Xd, \quad (1.4)$$

где

$$X = C(A^T W A)^{-1} A^T W. \quad (1.5)$$

Будем рассматривать только случай (1.3) (по терминологии [84, 18] — случай полноты системы измерений), чтобы не усложнять без особой надобности формулировку задачи. Если (2.3) не выполняется, то можно перейти к эквивалентной (1.1) модели измерений, удовлетворяющей условию (1.3) [84, 18]. С другой стороны, отметим, что изложенные ниже формулировку и решение задачи можно

распространить на общий случай  $A^T W A \geq 0$ , если воспользоваться соответствующими выражениями из [96] для оценки метода наименьших квадратов.

Матрица (1.5) удовлетворяет условию

$$X A = C. \quad (1.6)$$

Это условие, записанное для произвольного линейного алгоритма (1.4) (т.е. с произвольной матрицей  $X$  в (1.4)), эквивалентно условию несмещенности этого алгоритма в следующем смысле [84, 18]: если  $\xi = 0$ , то оценка  $\hat{L}$  совпадает с искомым параметром  $L$ .

Отметим, что в математической статистике условие несмещенности алгоритма оценивания определяется так: если  $E(\xi) = 0$ , то  $E(\hat{L}) = L$ , где  $E(\cdot)$  — вектор математического ожидания аргумента. Такое условие несмещенности эквивалентно условию (1.6) лишь при невырожденности матрицы ковариаций вектора  $\xi$  [96].

Задача, которую мы будем рассматривать, состоит в нахождении множества всех весовых матриц (положительно или неотрицательно определенных), для которых метод наименьших квадратов (1.2) приводит к одному и тому же линейному несмещенному алгоритму (1.4), т.е. отыскивается явное представление

$$\mathcal{W} \doteq \left\{ W > 0 : X = C (A^T W A)^{-1} A^T W \right\}, \quad (1.7)$$

$$\mathcal{W}' \doteq \left\{ W \geq 0 : X = C (A^T W A) A^T W \right\} \quad (1.8)$$

матриц  $W > 0$  и  $W \geq 0$ , которые определяют по формуле (1.5) одну и ту же заданную матрицу  $X$  (удовлетворяющую условию (1.6)). Будем отыскивать также явное описание множества

$$\mathcal{V} = \left\{ V : V = W^{-1}, W \in \mathcal{W} \right\} \quad (1.9)$$

матриц  $V$ , обратных матрицам  $W$  множества (1.7).

Решение поставленной задачи в общем виде с использованием изящной теории  $g$ -обратных матриц дано в теореме 1.1 раздела 1.1, (необходимые сведения о  $g$ -обратных матрицах даны в пункте 1.1.2). Читатель, для которого более привычны операции обычного обращения матриц, может сразу после введения перейти к теореме 1.2, где решение приведено для случая независимости строк матрицы  $C$  (к этому случаю всегда можно свести задачу). Иллюстрация этой теоремы и эквивалентности различных допущений о статистических характеристиках ошибок измерений дана в пункте 1.1.4 на простом примере. В пункте 1.1.5 приведены соображения о применении полученных результатов.

**ЗАМЕЧАНИЕ 1.1.** В работах [94, 97, 98] задачи типа (1.9) рассматриваются с несколько иной точки зрения, а именно предполагается, что известна матрица  $V \geq 0$  ковариаций вектора  $\xi$ . Кроме того, не требуется выполнения условия (1.3). В связи с этим обосновывается единая теория наименьших квадратов [?, 99] (обобщается алгоритм (1.2), в котором минимизация в (1.2) заменяется нахождением стационарной точки  $\hat{\theta}$  квадратичной формы

$$(z - A\theta)^T \tilde{V}^{-1} (z - A\theta), \quad (1.10)$$

где  $\tilde{V} = V + AUA^T$ ;  $U$  — произвольная  $m \times m$ -матрица, удовлетворяющая условию  $R(\hat{V}) = R(V, A)$ ;  $\tilde{V}$  — любая  $g$ -обратная матрица. При этом если вектор  $L \doteq C\theta$  может быть несмещенно оценен при  $E(\xi) = 0$ , то вектор оценки  $\hat{L} \doteq C\hat{\theta}$  имеет минимальные дисперсии своих компонент среди всех несмещенных линейных оценок (обобщение теоремы Гаусса-Маркова), а сама оценка одинакова для всех стационарных точек  $\hat{\theta}$  (при условиях (1.3) и  $V > 0$  стационарная точка единственна и совпадает с оценкой, полученной из (1.2) при  $W = V^{-1}$ ).

В теоремах 4.1 из [97], 6.2 из [98] для случая  $R(C) = R(A)$  (а также в [94] для случая  $C = A$ ) приведено явное описание множества (ср. с (1.9)):

$$\mathcal{V} \doteq \{V \geq 0 : Xz = C\theta\}. \quad (1.11)$$

Здесь  $C\hat{\theta}$  — оценка, соответствующая экстремуму в (1.10),  $Xz$  — линейная оценка с заданной матрицей  $X$ , удовлетворяющая условию (1.6) (которое является

достаточным для того, чтобы вектор  $L$  был оцениваем). В связи с этим можно сделать следующий вывод, упомянутый выше. Для случая  $R(C) = m$  (который переходом к эквивалентной модели сводится к случаю  $L = \theta$ ) множество (1.9) можно найти, выделяя из множества (1.11) подмножество, содержащее только матрицы  $V > 0$ .

### 1.1.2. Некоторые сведения из теории матриц

Приводимые ниже сведения являются известными результатами [96, 99] или их простыми следствиями.

А. Пусть  $A$  —  $n \times m$ -матрица ранга  $m$ , а  $W$  —  $n \times n$ -матрица,  $W > 0$ . Тогда  $A^T W A > 0$ .

*Сводка решений систем линейных уравнений*

Б. Если в соотношении (1.6)  $R(C) = k$ , то  $R(X) = k$ .

В.  $g$ -обращение и решение линейных уравнений для заданной  $(n \times m)$ -матрицы  $A$  произвольного ранга.  $g$ -обратной матрицей называется любая  $m \times n$ -матрица, удовлетворяющая соотношению  $AA^-A = A$ . Матрица  $A^-$  единственна и совпадает с  $A$  только в том случае, когда  $A$  невырождена. Способы нахождения  $A^-$  указаны в [96, 99]. Например, если ранг каждой из матриц  $A$  и  $Y$  размерностей  $n \times m$  и  $m \times n$  равен  $m$ , то их  $g$ -обратные матрицы определяются из условий

$$A^-A = I_m, \quad YY^- = I_m. \quad (1.12)$$

В таблице представлены условия совместности и общее решение систем уравнений, в которых  $X$  и  $Y$  — неизвестные матрицы,  $Z$  — произвольная матрица в общем решении, а остальные матрицы заданы. Условия совместности даны в двух эквивалентных формах: одна из них заимствована из [99] и использует  $g$ -обращение, другая выражена в виде традиционных равенств рангов и получена, например, из первой.

*Сводка решений систем линейных уравнений*

Система уравнений	Условия совместности	Общее решение
$BX = C$	$BB^{-1}C = C$	$X = B^{-1}C + (I - B^{-1}B)Z$
$BXA = C$	$\left. \begin{aligned} R(B, C) = R(B) \\ R(A^T, C^T) = R(A) \end{aligned} \right\} \Leftrightarrow$ $\Leftrightarrow C = BB^{-1}CA^{-1}A$	$X = B^{-1}CA^{-1} + Z - B^{-1}BZA A^{-1}$
$YA = B,$ $CY = D$	$DA = CB$ $R(A^T, B^T) = R(A) \Leftrightarrow BA^T A = B$ $R(C, D) = R(C) \Leftrightarrow CC^{-1}D = D$	$Y = C^{-1}D + BA^{-1} - C^{-1}CBA^{-1} + (I - C^{-1}C)Z(I - AA^{-1})$

### 1.1.3. Основные результаты

ТЕОРЕМА 1.1 (ПРЕДСТАВЛЕНИЯ МНОЖЕСТВ (1.7), (1.9)).

а) Для непустоты множества (1.9) необходимо и достаточно, чтобы система

$$YA = I, \quad CY = X \tag{1.13}$$

была разрешима относительно  $n \times t$ -матрицы  $Y$ , что эквивалентно условию

$$R(X) = R(C); \tag{1.14}$$

б) Для того чтобы матрицы  $W, V$  принадлежали множествам (1.7), (1.9), необходимо и достаточно, чтобы эти матрицы могли быть представлены в виде

$$W = Y^T M Y + S^T N S, \tag{1.15}$$

$$V = A \tilde{M} A^T + \tilde{S} \tilde{N} S^T, \tag{1.16}$$

где  $M, \tilde{M}, N, \tilde{N}$  — некоторые положительно определенные матрицы (порядок  $M$  и  $\tilde{M}$  равен  $t$ , а  $N$  и  $\tilde{N}$  —  $n$ );

$$S \doteq I - AA^-, \quad (1.17)$$

$$\tilde{S} \doteq I - Y^-Y, \quad (1.18)$$

$Y$  — некоторое решение системы (1.13), определяемое выражением

$$Y = C^-X + (I - C^-C) \{A^- + Z(I - AA^-)\}, \quad (1.19)$$

$Z$  — некоторая  $m \times n$ -матрица.

б) Множество (1.8) описывается текстом пункта б) с заменой условия  $N > 0$  на  $N \geq 0$ .

*Доказательство.* Докажем сначала, что выполнение равенства (1.14) является критерием совместности системы (1.13). Для этого отметим, что первые два условия совместности для этой системы из таблицы пункта 1.1.2 всегда выполнены в виду справедливости (1.6) и (1.3). Поэтому критерием совместности системы (1.13) является третье условие

$$R(X, C) = R(C). \quad (1.20)$$

Рассматривая теперь (1.6) как систему совместных уравнений относительно матрицы  $A$  и пользуясь таблицей, имеем  $R(X, C) = R(X)$ . Последнее равенство доказывает, что критерий совместности (1.20) записывается в виде (1.14).

Выпишем теперь необходимые в дальнейшем соотношения для матрицы (1.17). Используя сведения из пункта 1.1.2, легко получить, что

$$SS = S, \quad SA = 0. \quad (1.21)$$

Отсюда, учитывая (1.13), находим

$$(S - AY)S = (I - AY), \quad (1.22)$$

$$(S - AY)(S - AY) = I \Rightarrow R(S - AY) = n. \quad (1.23)$$

Перейдем к доказательству необходимых и достаточных условий из пунктов а) и б) теоремы для множества (1.7). Утверждение пункта в) доказывается аналогично; относительно множества (1.9) — см. ниже замечание к теореме.

Необходимость. Пусть множество (1.7) непусто и  $W$  — некоторый его элемент. Тогда легко проверить, что системе (1.13) удовлетворяет матрица

$$Y = (A^T W A)^{-1} A^T W, \quad (1.24)$$

т.е. выполнено необходимое условие пункта а).

Определим далее матрицы

$$M \doteq A^T W A, \quad N \doteq (S - AY)^T W (S - AY), \quad (1.25)$$

где  $y$  задано в (1.24). Покажем, что  $M > 0$ ,  $N > 0$  и при подстановке этих матриц в правую часть соотношения (1.15) получается значение, равное  $W$ . Тем самым будет доказано необходимое условие пункта б) теоремы. Применяя утверждение А из пункта 1.1.2, (с учетом (1.3)), получаем, что  $M > 0$ . Аналогично, учитывая (1.23), находим, что  $N > 0$ . Подставляя матрицы в правую часть (1.15) и используя (1.22), получим

$$Y^T M Y + S^T N S = W + Y^T A^T W (AY - I) + (Y^T A^T - I) W AY. \quad (1.26)$$

Равенство (1.24) эквивалентно соотношению

$$A^T W (AY - I) = 0.$$

Поэтому правая часть в (1.26) равна  $W$ . Необходимые условия теоремы доказаны.

Достаточность. Нужно показать, что если  $Y$  - некоторое решение системы (1.13), то при любых  $M > 0$  и  $N > 0$  выражение (1.15) определяет элемент множества (2.7). Легко проверить, используя (1.13) и второе соотношение в (1.21), что матрица  $W$ , определенная согласно (1.15), удовлетворяет уравнению (2.5). Остается показать, что  $W > 0$ . Из утверждения А пункта 1.1.2, следует, что



оба слагаемых в (1.15) есть положительно определенные матрицы. Поэтому для любого вектора  $p \in \mathbb{R}^n$  выполняется неравенство

$$p^T W p = 0. \quad (1.27)$$

При этом равенство в (1.27) возможно лишь при условиях  $Yp = 0$ ,  $Sp = 0$ , следствием которых является равенство  $(S - AY)p = 0$ . Согласно (1.23), последнее равенство, а следовательно и равенство в (1.27), возможны лишь при  $p = 0$ . Отсюда  $W > 0$ . Достаточные условия теоремы доказаны.

Для завершения доказательства теоремы отметим, что общее решение системы (1.13), определяемое выражением (1.19) (при произвольной матрице  $Z$  в этом выражении), найдено с помощью таблицы из пункта 1.1.2,

**ЗАМЕЧАНИЕ 1.2.** При доказательстве необходимых условий пункта б) теоремы 1.1 для множества (2.9) вместо матрицы (1.25) надо взять матрицы

$$\tilde{M} = (A^T V^{-1} A)^{-1}, \quad \tilde{N} = (\tilde{S} - AY) V (\tilde{S} - AY)^T. \quad (1.28)$$

**ЗАМЕЧАНИЕ 1.3.** Если соотношение между матрицами параметров представлений (1.15) и (1.16) имеет вид

$$\tilde{M} = M^{-1}, \quad \tilde{N} = T N^{-1} T^T \quad \Leftrightarrow \quad N = T^T \tilde{N}^{-1} T,$$

где  $T \doteq (\tilde{S} - AY)(S - AY)$ , то соответствующие этим параметрам элементы множеств (1.7) и (1.9) удовлетворяют условию  $V = W^{-1}$ . Это нетрудно получить, используя выражения (1.25), (1.28).

Приведенный в пункте а) теоремы 1.1 критерий непустоты множества (1.7) относится к самому общему случаю. Покажем, что в наиболее естественном случае, когда компоненты вектора  $L$  независимы, множество (1.7) непусто, и к этому случаю всегда можно свести задачу при ее корректной постановке. А именно, имеют место следующие два утверждения.

1. Если строки матрицы  $C$  линейно независимы, т.е.

$$R(C) = k \quad (1.29)$$

(и, следовательно,  $k \leq m$ ), то множество (1.7) не пусто.

Это следует из утверждения Б пункта 1.1.2 и пункта а) теоремы 1.1.

2. Пусть выполнено условие (1.14) и

$$k > \tilde{k} \doteq R(C) = R(X).$$

Тогда задачу нахождения множества (1.7) можно записать в эквивалентной форме, для которой выполнены условия (1.29). А именно, можно заменить в (1.5) и (1.6) матрицы  $C$ ,  $X$  на матрицы

$$\tilde{C} = GC, \quad \tilde{X} = GX, \quad (1.30)$$

где  $G$  —  $\tilde{k} \times k$ -матрица, такая, что  $R(\tilde{C}) = \tilde{k}$  ( $G$  может быть построена как произведение любой невырожденной  $\tilde{k} \times \tilde{k}$ -матрицы на матрицу перехода от  $C$  к любой ее  $\tilde{k} \times m$ -подматрице ранга  $\tilde{k}$ ).

Действительно, умножая определяющие постановку задачи (1.7) соотношения (1.5), (1.6) на  $G$  слева и учитывая (1.30), приходим к аналогичным соотношениям

$$\tilde{X}A = \tilde{C}, \quad \tilde{X} = \tilde{C}(A^TWA)^{-1}A^TW. \quad (1.31)$$

При этом условие (1.29) выполнено для  $\tilde{X}$  и  $\tilde{C}$  согласно утверждению Б из пункта 1.1.2, Осталось показать обратное, что из (1.31) следует (1.5) и (1.6). Умножая оба соотношения в (1.31) слева на любую матрицу  $G^{-1}$  и учитывая, что  $GG^{-1} = I$  (согласно (1.12)), приходим к соотношениям (1.5), (1.6). Утверждение 2 доказано.

Согласно последнему результату, задачу нахождения множеств (1.7), (1.8) можно без ограничения общности рассматривать для случая (1.29). Приведем для этого случая вариант теоремы 1.1, в котором не используется  $g$ -обращение и, кроме того, дается представление искомых множеств в зависимости от минимального числа параметров.

ТЕОРЕМА 1.2. Пусть выполнено условие (1.29).

а) Представим матрицы  $A$  и  $C$  в виде

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad C = [C_1, C_2]$$

где  $A_2$  и  $C_2$  - матрицы размерностей  $t \times t$  и  $k \times k$ , которые без ограничения общности можно считать невырожденными (это условие достигается нужной нумерацией строк матрицы  $A$  и столбцов матриц  $C$  и  $X$ ). Тогда элементы множества (1.7) описываются формулой

$$W = Y^T M Y + \begin{bmatrix} N_1 & N_1 B \\ B^T N_1 & B^T N_1 B \end{bmatrix}. \quad (1.32)$$

Здесь  $B \doteq A_1 A_2^{-1}$ ,  $M$  и  $N_1$  — произвольные положительно определенные матрицы порядков  $t$  и  $n \times t$ ,  $Y$  — произвольное решение системы (1.13), которое может быть представлено в виде

$$Y = \begin{bmatrix} Z_1 & Z_1 B + F \\ D Z_1 & D (Z_1 B + F) \end{bmatrix} + \begin{bmatrix} C \\ C_2^{-1} X \end{bmatrix}; \quad (1.33)$$

$D \doteq -C_2^{-1} C_1$ ,  $F$  -  $(t - k) \times t$ -матрица, составленная из первых  $t - k$  строк матрицы  $A_2^{-1}$ ,  $Z_1$  — произвольная матрица. Множество (1.8) также описывается в виде (1.32) с заменой условия  $N_1 > 0$  на  $N_1 \geq 0$ .

б) Представим произвольное решение системы (1.13) в виде

$$Y = (Y_1, Y_2),$$

где  $Y_2$  -  $t \times t$ -матрица, которую без ограничения общности можно считать невырожденной (это условие достигается при известной матрице  $Y$  нужной нумерацией ее столбцов и соответствующих строк матрицы  $A$  и столбцов матриц  $C$  и  $X$ ). Тогда элементы множества (1.9) описываются формулой

$$V = A \tilde{M} A^T + \begin{bmatrix} N_1 & \tilde{N}_1 \tilde{B}^T \\ \tilde{B} \tilde{N}_1 & \tilde{B} \tilde{N}_1 \tilde{B}^T \end{bmatrix}. \quad (1.34)$$

Здесь  $\tilde{B} \doteq -Y_2^{-1}Y_1$ ,  $\tilde{M}$  и  $\tilde{N}$  — произвольные положительно определенные матрицы порядков  $m$  и  $n \times m$ .

б) Представление (1.32) (или (1.34)) определяет зависимость элементов множества (1.7) (или (1.9)) от минимально возможного числа параметров, которыми являются неповторяющиеся элементы матриц  $M$ ,  $N_1$  (или  $\tilde{M}$ ,  $\tilde{N}_1$ ) и элементы матрицы  $Z_1$ .

*Доказательство.* Зависимости (1.32) — (1.34) получаются из (1.15), (1.19) и (1.16), если подставить туда для матриц  $A$ ,  $C$ ,  $Y$  их  $g$ -обратные

$$A^- = [0, A_2^{-1}], \quad C^- = \begin{bmatrix} C \\ C_2^{-1} \end{bmatrix}, \quad Y^- = \begin{bmatrix} C \\ Y_2^{-1} \end{bmatrix},$$

которые получены из условия (1.12). Для доказательства пункта в) теоремы достаточно показать, что соотношения (1.32), (1.33) определяют взаимно однозначное соответствие между элементами множества (1.7) и матрицами  $M$ ,  $N_1$ ,  $Z_1$ . Каждым  $M > 0$ ,  $N_1 > 0$  и  $Z_1$ , согласно утверждению а) теоремы, соответствует матрица из множества (1.7). Обратно, для каждой  $W$  из (1.32) однозначно определяются матрица  $M > 0$  по формуле (1.25), матрицы  $A_1 > 0$  и  $Z_1$  как верхние угловые подматрицы матриц  $W - Y^T M Y$  и  $Y$ , где  $Y$  определяется из (1.24). Теорема доказана.

**ЗАМЕЧАНИЕ 1.4.** Минимальное число параметров, от которых зависит матрица  $W$  в (1.32), согласно доказанной теореме равно

$$\begin{aligned} \frac{1}{2} \left[ (n-m)(n-m+1) + m(m+1) \right] + (m-k)(n-m) &\doteq \\ &\doteq \frac{n(n+1)}{2} - (nk - mk). \end{aligned} \quad (1.35)$$

Поясним этот факт следующими соображениями. Симметрическая матрица  $W$  содержит  $\frac{n(n+1)}{2}$  неповторяющихся элементов и связана  $nk$  уравнениями (1.5) с матрицей  $X$ , которая сама удовлетворяет  $mk$  уравнениям (1.6). Поэтому число параметров, от которых зависит общее решение уравнений (1.5) в классе симметрических матриц, должно быть равно правой части в (1.35).

**ЗАМЕЧАНИЕ 1.5.** В пункте 1.1.5, будет рассмотрена задача, где матрица из множества (1.7) используется как весовая в методе наименьших квадратов для модели, параметры  $A$  и  $C$  которой, вообще говоря, не совпадают с параметрами в (1.1). При этом для упрощения вычислений при нахождении соответствующей матрицы  $X$  желательно выбрать матрицу  $W$  с наибольшим количеством нулей вне диагонали. Мы не исследовали вопрос, может ли это число нулей быть равно числу (1.35). Однако легко отыскать матрицу из множества (1.7), у которой угловая  $((n - m) \times (n - m))$ -подматрица является заданной диагональной матрицей  $D > 0$ . Для этого достаточно выбрать матрицу  $M$  в (1.32) такой, чтобы наибольшее число матрицы  $Y^T M Y$  было меньше наименьшего элемента матрицы  $D$ , а затем положить в (1.32)  $N_1$  равной угловой подматрице матрицы  $D - Y^T M Y$ .

#### 1.1.4. Пример

Пусть требуется определить скалярную величину  $\theta$  по результатам измерений ее же самой. Для этой задачи модель (1.1) принимает вид

$$z = \theta e_n + \xi, \quad L = \theta \text{ — скаляр} \Rightarrow C = 1, \quad (1.36)$$

где  $e_n \in \mathbb{R}^n$  — вектор, все компоненты которого равны 1.

Рассмотрим для модели (1.36) метод наименьших квадратов (1.5) при  $W = I_n$ . Для данного случая алгоритм (1.3) имеет вид

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n d_i \Rightarrow X = \frac{1}{n} e_n^T \quad (1.37)$$

( $d_i$  — компоненты вектора  $d$ ). При этом формула (1.34) дает следующее представление множества (1.9) всех матриц  $V$ , для которых метод наименьших квадратов (1.5) при  $W = V^{-1}$  определяет алгоритм оценивания (1.37):

$$V = \begin{bmatrix} M e_{n-1} e_{n-1}^T & (M I_{n-1} - N) e_{n-1} \\ e_{n-1}^T (M I_{n-1} - N) & M + e_{n-1}^T N e_{n-1} \end{bmatrix}. \quad (1.38)$$

Здесь  $M > 0$  — произвольная постоянная,  $N > 0$  — произвольная матрица порядка  $n - 1$ . Положим в (1.38)

$$M = k + \frac{1 - k}{n}, \quad N = \frac{1 - k}{I_{n-1} + \frac{\epsilon_{n-1}\epsilon_{n-1}^T}{n}},$$

где  $k$  — любое число из отрезка  $[0, 1]$ . Тогда получим следующую матрицу из множества (1.38):

$$L = (1 - k)I_n + ke_n e_n^T. \quad (1.39)$$

Рассмотрим теперь трактовку метода наименьших квадратов (1.2), при которой под  $V = W^{-1}$  понимается матрица ковариаций вектора  $\xi$  (такая трактовка принята в математической статистике). В этой трактовке матрица (1.39) соответствует случаю, когда все ошибки измерений имеют единичную дисперсию, а коэффициент корреляции между любыми ошибками равен  $k$  [84, 18]. Метод наименьших квадратов (1.2) с матрицей (1.39) в качестве  $W^{-1}$  дает нам, как показано, одну и ту же оценку (1.37) при любом значении коэффициента корреляции  $k$ . Дисперсия этой оценки равна  $k + \frac{1 - k}{n}$  и при  $n \rightarrow \infty$  стремится к  $k$ . Поэтому даже при несмещенных ошибках измерений оценка (1.37) не является состоятельной при  $k > 0$ .

### 1.1.5. О применении полученных результатов

Рассмотрим сначала следующую задачу. Пусть метод наименьших квадратов используется для определения  $\hat{\theta}$  в нелинейной модели

$$z = f(\theta) + \xi, \quad L = L(\theta),$$

где  $f(\cdot)$ ,  $L(\cdot)$  — некоторые нелинейные функции. Оценку  $\hat{\theta}$  вычисляют обычно методом последовательных приближений путем линеаризации модели на каждом приближении около оценки  $\hat{\theta}^{(s)}$  на  $s$ -м шаге [84, 18]:

$$\hat{\theta}^{(s)} = \hat{\theta}^{(s-1)} + \hat{r}^{(s)}, \quad s = 0, 1, 2, \dots,$$

где  $\hat{r}^{(s)}$  — оценка поправки  $r^{(s)}$ , определяемая из метода наименьших квадратов с заданной весовой матрицей  $W$  при использовании линейной модели:

$$d - f(\theta^{(s)}) = A^{(s)}r^{(s)} + \xi, \quad A^{(s)} \doteq \left. \left( \frac{\partial f}{\partial \theta} \right) \right|_{\theta^{(s)}}.$$

Вычисления в этом методе значительно упрощаются, если заменить матрицу  $W$  на матрицу, имеющую много нулевых элементов. Это можно сделать следующим образом. На первом шаге вычислим

$$X^{(0)} = \left( A^{(0)\top} W A^{(0)} \right)^{-1} A^{(0)\top} W \quad (1.40)$$

и найдем множество (1.7) для этой матрицы (при  $C^{(0)} = I_m$ ). В соответствии с замечанием 1.3 к теореме 1.1 выберем из этого множества матрицу, содержащую угловую диагональную подматрицу порядка  $n \times m$ . Эту матрицу будем использовать для вычисления поправок  $r^{(s)}$  при  $s > 0$  на нескольких шагах, после чего повторяем указанную процедуру замены  $W$ , изменив индекс 0 в (1.40) на текущий индекс  $s$ . При  $n \gg m$  и большом числе итераций предлагаемая замена матрицы  $W$  может существенно сократить вычисления.

Приведем еще некоторые соображения относительно использования зависимости между множествами оценщиков и весовых матриц. Найдя из каких-либо соображений матрицу  $X$  (далее называемую оценщиком), определяющую оценку  $\hat{L} = Xy$ , мы можем подобрать весовую матрицу, которой соответствует в линейной модели та же оценка МНК. Однако для линейной модели нет смысла искать весовую матрицу, так как уже есть оценщик  $X$  и МНК даст этот же оценщик. Пусть мы используем линейную модель для итеративного решения задачи оценивания с нелинейной моделью измерений вида

$$y = f(\Theta) + \xi$$

( $\Theta$  — вектор состояния). При этом  $\theta = \Theta - \Theta^0$ , где  $\Theta^0$  — известное (номинальное) значение вектора состояния. А в нелинейной модели при использовании МНК желательно иметь одну весовую матрицу хотя бы на нескольких итерациях. Так как при переходе от нелинейной модели к линейной матрица  $H^\top$  есть производная от  $f(\Theta)$  по  $\Theta$ , то обычно эта производная мало меняется от итерации

к итерации. Поэтому возможно вычислить оптимальный по какому-либо критерию оценитель  $X$ , найти соответствующую ему весовую матрицу (желательно, почти диагонального вида, чтобы упростить вычисления в МНК) и использовать эту матрицу в течение нескольких итераций. Через несколько итераций, когда матрица  $H$  уже существенно изменится за счет того, что изменится номинальный вектор  $\Theta^0$  (он станет равным текущей оценке), мы можем заново найти оптимальный оценитель и соответствующую ему весовую матрицу.

## 1.2. Выбор мешающих параметров в схеме линейной регрессии и множество линейных несмещенных алгоритмов оценивания

### 1.2.1. Модель оценивания

Рассмотрим линейную модель оценивания

$$z = A\theta + \xi, \quad L = C\theta + \eta, \quad (2.1)$$

где  $z \in \mathbb{R}^n$  — вектор результатов измерений,  $\theta \in \mathbb{R}^m$  — вектор неизвестных параметров,  $\xi$  — суммарный вектор ошибок измерений,  $L \in \mathbb{R}^k$  — вектор интересующих нас параметров,  $\eta$  — вектор ошибок для параметра  $L$ ,  $A$  и  $C$  — заданные матрицы. Эта модель на практике получается линейризацией реальной нелинейной модели. При этом  $\theta$  есть отклонение вектора некоторых неизвестных параметров относительно их априорных значений. Поэтому можно считать, что априорное значение вектора  $\theta$  также известно и равно нулю. Мы однако будем обозначать это априорное значение через  $\tilde{\theta}$ , допуская, что линейризация может быть проведена не обязательно относительно априорных значений определяемого вектора. В соответствии с изложенным к уравнениям (2.1) следует добавить ещё  $m$  уравнений:



$$\tilde{\theta} = \theta + \delta\theta. \quad (2.2)$$

На практике обычно некоторый подвектор  $\theta_2$  вектора  $\theta$  принимают равным своему априорному значению  $\tilde{\theta}_2$  ( $\theta_2 \neq \theta$ ), а оцениваемым вектором считают совокупность  $\theta_1$  остальных компонент вектора  $\theta$ . Это обусловлено в частности следующими причинами:

- размерность вектора  $\theta$  достаточно велика для того, чтобы определить все его компоненты, так что неоправданное увеличение вектора состояния приводит к увеличению трудоёмкости вычислений при оценивании параметров траектории;
- вектор  $\theta_2$  априорно известен достаточно точно и включение его в вектор состояния может увеличить ошибки оценивания ввиду неточного знания статистических характеристик ошибки  $\tilde{\theta}_2 - \theta_2$  и поэтому неоптимального выбора алгоритма оценивания.

При указанном подходе вектор  $\theta_2$  называют мешающим параметром.

В соответствии с концепцией мешающих параметров разобьём матрицы  $A$  и  $C$  на подматрицы  $A_1, A_2$  и  $C_1, C_2$ , соответствующие векторам  $\theta_1$  и  $\theta_2$ , т.е.

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad C = [C_1, C_2],$$

в качестве новых векторов измерений и оцениваемых векторов рассмотрим векторы

$$z_1 = z - A_2\tilde{\theta}_2, \quad L_1 = L - C_2\tilde{\theta}_2, \quad (2.3)$$

в качестве новых векторов ошибок — векторы

$$\begin{aligned} \xi_1 &= \xi + A_2(\tilde{\theta}_2 - \theta_2), \\ \eta_1 &= \eta - C_2(\tilde{\theta}_2 - \theta_2), \\ \delta\theta_1 &= \tilde{\theta}_1 - \theta_1. \end{aligned} \quad (2.4)$$

Тогда модель оценивания (2.1), (2.2) переписется в виде

$$\begin{cases} z_1 = A_1\theta_1 + \xi_1 \\ \tilde{\theta}_1 = \theta_1 + \delta\theta_1 \end{cases} \Leftrightarrow \begin{cases} z_1 = A_1\theta_1 + \begin{bmatrix} \xi_1 \\ \delta\theta_1 \end{bmatrix} \\ Z_1 \doteq \begin{bmatrix} z_1 \\ \tilde{\theta}_1 \end{bmatrix}, \quad \tilde{A}_1 \doteq \begin{bmatrix} A_1 \\ I_1 \end{bmatrix}, \end{cases} \quad (2.5)$$

где  $I_1$  — единичная матрица, а  $Z_1$  — расширенный вектор измерений.

Найдя оценку  $\hat{L}_1$  вектора  $L_1$ , мы получим оценку вектора  $L$ , исходя из (2.3):

$$\hat{L} = \hat{L}_1 + C_2\tilde{\theta}_2. \quad (2.6)$$

Для оценки вектора  $L_1$  будем использовать линейный несмещенный алгоритм. Согласно пункту 1.1.1 этот алгоритм характеризуется матрицей  $X_1$  (далее — оцениватель), так что

$$\hat{L}_1 = X_1 Z_1, \quad (2.7)$$

где оцениватель удовлетворяет условиям несмещенности

$$X_1 \tilde{A}_1 = C_1. \quad (2.8)$$

Отметим, что согласно результатам пункта 1.1.3 множество оценок (2.7) эквивалентно множеству оценок метода наименьших квадратов, которые мы рассмотрим ниже.

## 1.2.2. Эквивалентность множеств всевозможных линейных несмещенных оценок при различном выборе мешающих параметров

Эта эквивалентность формулируется следующим образом.

**ТЕОРЕМА 1.3.** *Множество линейных несмещенных оценок (2.7), получаемое при всевозможных матрицах  $X_1$ , удовлетворяющих условию (2.8) и соответствующее некоторому выбору вектора  $\theta_2$  мешающих параметров в модели (2.1), (2.2) ( $\theta_2 \neq \theta$ ), не зависит от этого выбора, совпадая тем самым с*

множеством оценок, соответствующим отсутствию вектора мешающих параметров, т.е. случаю  $\theta_1 = \theta$ .

*Доказательство.* Нам нужно показать, что каждая оценка (2.6), обусловленная выбором оценителя  $X_1$  в (2.7), является несмещенной оценкой для случая  $\theta_1 = \theta$ , т.е. может быть записана в виде

$$\hat{L} = XZ, \quad Z \doteq \begin{bmatrix} z \\ \hat{\theta} \end{bmatrix}, \quad (2.9)$$

где  $Z$  — расширенный вектор измерений, а оценитель  $X$  удовлетворяет условию несмещенности

$$X\tilde{A} = C, \quad \tilde{A} \doteq \begin{bmatrix} A \\ I \end{bmatrix} \quad (2.10)$$

(здесь  $I_1$  и  $I_2$  - единичные матрицы), и обратно, каждая оценка (2.9) с матрицей  $X$ , удовлетворяющей (2.10), есть оценка, получаемая из (2.6), (2.7) с некоторой матрицей  $X_1$ , удовлетворяющей (2.8).

Запишем (2.6), (2.7) в виде

$$\hat{L} = C_2\theta_2 + X_1 \begin{bmatrix} z - A_2\tilde{\theta}_2 \\ \tilde{\theta}_1 \end{bmatrix} = XZ, \quad (2.11)$$

$$X \doteq \left[ X_1, C_2 - X_1 \begin{bmatrix} A_2 \\ 0 \end{bmatrix} \right]. \quad (2.12)$$

Отсюда, используя (2.8) — (2.11), получаем

$$X\tilde{A} = \left[ X_1\tilde{A}_1, C_2 \right] = [C_1, C_2] = C. \quad (2.13)$$

Таким образом первая часть доказательства завершена.

Пусть теперь матрица  $X$  удовлетворяет условию (2.10). Представляя эту матрицу в виде  $X = [X_1, X_2]$ , где  $X$  содержит столько столбцов, сколько матрица  $I_2$ , получаем из (2.10), что  $X_1$  удовлетворяет условию (2.8), а

$$X_2 = C_2 - X_1 \begin{bmatrix} A_2 \\ 0 \end{bmatrix}. \quad (2.14)$$

Поэтому оценку (2.9) можно записать в виде (2.11). Теорема доказана.

### 1.2.3. Эквивалентность множества всех оценок метода наименьших квадратов и линейных несмещенных оценок при различном выборе вектора мешающих параметров.

Рассмотрим теперь для модели (2.5) оценку метода наименьших квадратов

$$\hat{L} = C_1 \hat{\theta}_1 + C_2 \tilde{\theta}_2, \quad (2.15)$$

где  $\hat{\theta}_1$  - оценка получаемая при использовании метода наименьших квадратов с некоторой весовой матрицей  $W$ , т.е., согласно разделу 1.1,

$$\theta_1 = \arg \min (\delta^T W \delta), \quad \delta = z_1 - \tilde{A}_1 \theta_1. \quad (2.16)$$

Из (2.15), (2.16) получаем, используя формулы из раздела 1.1

$$\hat{L} = XZ, \quad (2.17)$$

где  $X$  определяется согласно формуле (2.12), в которой

$$X_1 \doteq C_1 (\tilde{A}_1^T W \tilde{A}_1)^{-1} \tilde{A}_1^T W. \quad (2.18)$$

Здесь расширенный вектор измерений  $Z$  определен в (2.9), а расширенная матрица  $\tilde{A}_1$  — в (2.5). Оценка (2.17) есть линейная оценка (2.11), причем матрица  $X_1$  удовлетворяет условию (2.8). Обратно, согласно теореме 1.1, каждая матрица  $X_1$ , удовлетворяющая условию (2.8), представима в виде (2.18), где  $W$  может быть взята из некоторого множества, зависящего от  $X_1$ . Поэтому каждая оценка (2.11) представляет собой оценку метода наименьших квадратов (2.15), (2.16) с некоторой весовой матрицей. Учитывая изложенное и теорему 1.3, приходим к следующему результату.

**ТЕОРЕМА 1.4.** *При любом выборе вектора  $\theta_2$  ( $\theta_2 \neq \theta$ ) мешающих параметров множество всех оценок метода наименьших квадратов, получаемое при всевозможных весовых матрицах в (2.16), совпадает с множеством линейных несмещенных оценок (2.9), получаемым при всевозможных  $X$ , удовлетворяющих (2.10).*

#### 1.2.4. Ошибки линейного оценивания.

Теоремы 1.3 и 1.4 имеют большое значение при построении оптимальных алгоритмов оценивания. Из них следует, что если такую оптимизацию проводить на множестве всех линейных несмещенных оценок (2.9) и исходить из условия достижения максимальной точности оценки  $\hat{L}$ , то при использовании линейной модели (2.1) эквивалентными являются подходы, учитывающие мешающие параметры при линейном оценивании или в методе наименьших квадратов (в виду совпадения множеств оценок, и следовательно, ошибок оценивания для этих всех случаев).

Таким образом, в оптимальных задачах оценивания при оптимизации по множеству всех несмещенных оценок можно ограничиться выражением для ошибки оценки (2.9), которая равна (см. раздел 1.2)

$$\delta L = \hat{L} - L \stackrel{(2.9)}{=} XZ - C\theta - \eta \stackrel{(2.1)}{=} X(\tilde{A}\theta + \tilde{\xi}) - C\theta - \eta \stackrel{(2.10)}{=} X\tilde{\xi} - \eta, \quad (2.19)$$

где  $\tilde{\xi} = \{\xi, \tilde{\theta} - \theta\}$  — составной вектор ошибок измерений и априорных данных, а матрица  $X$  удовлетворяет условию (2.10).

### 1.3. Вычисление гарантированных характеристик точности оценивания при наличии немоделируемых возмущений

Пусть  $X = X(t) \in \mathbb{R}^m$  — вектор, полностью определяющий траекторию движения системы в заданной системе координат. Этот вектор, который мы будем называть вектором состояния, включает в себя для самолета, например, обычно двенадцать фазовых координат, характеризующих движение центра масс самолета и вращение его около центра масс, а также ряд постоянных, характеризующих аэродинамические свойства самолета, величину сопротивления воздуха и т.п.

Пусть изменение вектора  $X$  на интервале  $[0, T]$  описывается системой

дифференциальных уравнений

$$\dot{X} = f(X, u, t), \quad t \in [0, T], \quad (3.1)$$

где  $u = u(t) \in \mathbb{R}^r$  — вектор возмущений, значение которого мало, но неизвестно. Мы будем полагать известными лишь ограничения

$$\|u(t)\| \leq \gamma(t), \quad (3.2)$$

где  $\|\cdot\|$  означает некоторую норму (например,  $\|u(t)\| = \sqrt{u^T u}$  или  $\|u(t)\| = \max_i |u_i|$ ),  $\gamma(t)$  — заданная функция).

Возмущение  $u(t)$  далее будем называть немоделируемым возмущением. Ввиду неопределенности этого возмущения при решении задачи определения движения системы мы можем использовать лишь модельное уравнение движения, которое получается из (3.1) при нулевом возмущении:

$$\dot{X} = f(X, O_r, t), \quad (3.3)$$

где  $O_r$  — нулевой  $r$ -мерный вектор. Пусть

$$X(t_0) = X_0 \quad (3.4)$$

— начальное условие для некоторого заданного момента  $t_0$ , а

$$X(t) = X(t, t_0, X_0) \quad (3.5)$$

— решение системы (3.3) при этом начальном условии. Задача оценивания траектории по результатам измерений сводится к нахождению оценки  $\hat{X}_0$  вектора  $X_0$ , после чего оценка вектора состояния определяется согласно (3.5):

$$\hat{X}(t) = X(t, t_0, \hat{X}_0). \quad (3.6)$$

Если скалярный параметр траектории  $l$  есть некоторый функционал от  $X(t)$ , т.е.

$$l = l[X(t)], \quad (3.7)$$

то оценка скалярного параметра (3.7) находится при известной оценке (3.6) по формуле

$$\hat{l} = l[\hat{X}(t)]. \quad (3.8)$$

В частности, параметр  $l$  может быть функцией состояния в заданный момент  $\tau$ :

$$l = l(X(\tau)). \quad (3.9)$$

Подставляя (3.5) в (3.7) или (3.9), получаем в конечном итоге зависимость вида

$$l = l(X_0, t_0). \quad (3.10)$$

Поэтому соотношение (3.8) для оценки параметра  $l$  эквивалентно соотношению, получаемому при использовании функции (3.10):

$$\hat{l} = l(\hat{X}_0, t_0). \quad (3.11)$$

Перейдем теперь к описанию модели измерений. Пусть на некотором множестве  $\mathcal{T} \subset [0, T]$  производятся измерения

$$z(t) = \varphi(X(t), u(t)) + \xi_z(t), \quad t \in \mathcal{T}. \quad (3.12)$$

Здесь  $z(t)$  — скалярная функция измеренных значений,  $\varphi(\cdot)$  — функция известного вида,  $\xi_z(t)$  — ошибка измерения. Множество  $\mathcal{T}$  на практике состоит из конечного числа моментов времени  $t_i$ , которые мы будем полагать упорядоченными:

$$t_1 \leq t_2 \leq \dots \leq t_n. \quad (3.13)$$

Иногда нам будет удобно рассматривать общий случай, когда это множество может включать произвольные числовые множества из  $[0, T]$ . Если эти множества — интервалы, то будем говорить о непрерывных измерениях, а в случае (3.13) — о дискретных измерениях.

В уравнениях измерений (3.12) мы предполагаем наличие тех же немоделируемых возмущений, что и в уравнениях движения (3.1). Такой случай имеет место, например, при посадке самолета. В силу неопределенности возмущений мы вынуждены заменить уравнения измерений (3.12) модельными уравнениями

$$z(t) = \varphi(X(t), O_r) + \xi_z(t), \quad (3.14)$$

полагая, как и ранее в (3.3), возмущение равным нулю.

Вышеизложенное показывает, что определение траектории движения сводится к нахождению оценки вектора состояния  $X_0$ . На практике этот вектор имеет достаточно много компонент. По тем же соображениям, которые были изложены в пункте 1.2.1, некоторые компоненты этого вектора целесообразно положить равными своим априорным значениям. В связи с этим представим вектор  $X_0$  в виде составного вектора

$$X_0 = \{X_{10}, X_{20}\}, \quad (3.15)$$

где компоненты вектора  $X_{10} \in \mathbb{R}^{m_1}$  определяются по результатам измерений, а компоненты вектора  $X_{20} \in \mathbb{R}^{m_2}$  полагаются равными своим известным априорным значениям  $\tilde{X}_{20}$ . Отметим при этом, что вектор  $\tilde{X}_{10}$  также может быть известен, но его компоненты уже должны рассматриваться как измерения. В представлении (3.15) векторы  $X_{10}$  и  $X_{20}$  называют соответственно векторами определяемых и мешающих параметров.

В связи с указанной концепцией рассмотрим для получения оценки вектора  $X_{10}$  метод наименьших квадратов (с мешающим параметром  $X_{20}$ ):

$$\hat{X}_{10} = \arg \min_{X_{10}} \left\{ \sum_{i=1}^n p_i^2 s_i^2 : X_{20} = \tilde{X}_{20} \right\}, \quad (3.16)$$

где

$$s_i \doteq z(t_i) - \varphi \left( X(t_i), O_r \right),$$

$X(t_i)$  выражается через  $X_0$  согласно (3.5), веса  $p_i$  выбираются в соответствии с точностью измерений и допущениями о немоделируемых возмущениях в моменты  $t_i$ . Величины  $s_i$  называют невязками. Отметим что в (3.16) минимизация ведется только по  $X_{10}$ , а  $X_{20}$  полагается равным своему априорному значению.

Далее мы будем считать, что априорно известны максимальные значения ошибок измерений (такая постановка задачи называется схемой бортиков), т.е.

$$|\xi_z(t)| \leq \sigma_i, \quad i = 1, \dots, n. \quad (3.17)$$

В соответствии с этим предположением и допущением (3.2), можно задать веса в



(3.16) следующим образом:

$$p_i^{-1} = \sigma_i + \gamma(t)\|g(t)\|, \quad g(t) \doteq \frac{\partial \varphi(X(t), u(t))}{\partial u(t)} \Big|_{X(t)=\hat{X}(t), u(t)=0}. \quad (3.18)$$

Рассмотрим теперь, какую точность дает алгоритм (3.16) для оценивания параметра (3.10). Ошибка оценивания этого параметра есть

$$\delta l = \hat{l} - l. \quad (3.19)$$

Наихудшее, т.е. гарантированное значение модуля этой ошибки при ограничениях (3.2), (3.17) на немоделируемые возмущения и ошибки измерения есть

$$\delta l_{\text{гар}} = \max_{u(t), \xi_z(t)} \{ \delta l : |\xi_z(t)| \leq \sigma_i, \|u(t)\| \leq \gamma(t), \}. \quad (3.20)$$

Гарантированная ошибка (3.20) определена нами для случая использования метода наименьших квадратов (3.16). Однако точно так же она может быть определена и для любого другого алгоритма оценивания. Если  $\mathcal{L}$  — множество всех рассматриваемых алгоритмов оценивания, то может быть поставлена задача выбора оптимального алгоритма, дающего наименьшее значение гарантированной ошибки. Эта задача описывается в виде

$$\delta l_0 = \min \{ \delta l_{\text{гар}} : \hat{\delta}(\cdot) \in \mathcal{L} \}. \quad (3.21)$$

В качестве примера множества  $\mathcal{L}$  приведем совокупность всех алгоритмов обобщенного метода наименьших квадратов. Каждый такой алгоритм характеризуется положительно определенной матрицей  $(P_{ij})$ , называемой весовой матрицей, а вектор  $\hat{X}_0$  определяется из более общего, чем (3.16), соотношения

$$\hat{X}_0 = \arg \min_{X_{10}} \left\{ \sum_{i,j=1}^n P_{ij} s_i s_j : X_{02} = \tilde{X}_{02} \right\}, \quad P_{ii} = P_i^2, \quad (3.22)$$

где невязки  $s_i$  определяются в (3.16). Алгоритм (3.16) есть, очевидно, частный случай алгоритма (3.22).

### 1.3.1. Метод наименьших квадратов и ошибка оценивания для линейного приближения

Далее будем полагать, что уравнения движения и наблюдения (3.1), (3.12) могут быть линеаризированы около некоторой опорной траектории, характеризу-

емой фазовым вектором  $\overset{\circ}{X}(t)$ . Наилучшей была бы линеаризация около истинной траектории движения. Однако она неизвестна. Поэтому опорная траектория уточняется путем ее определения по результатам измерений. В результате ошибка линеаризации будет порядка  $\|\hat{X}(t) - X(t)\|$ , где  $\|\cdot\|$  — евклидова норма. Этой ошибкой линеаризации мы можем либо пренебречь, либо при линеаризации уравнений (3.1), (3.12) учесть эту ошибку соответствующим добавлением их к немоделируемым возмущениям и ошибкам измерений. При линеаризации выражения (3.9) для оцениваемого параметра добавляется ошибка линеаризации  $\eta_l$ .

В результате вместо соотношений (3.1), (3.12), (3.9) имеем для вектора

$$q(t) = X(t) - \overset{\circ}{X}(t) \quad (3.23)$$

линейные уравнения движения и измерений:

$$\dot{q}(t) = R(t)q(t) + F(t)u(t), \quad t \in [0, T], \quad (3.24)$$

$$z(t) = h(t)^T q(t) + g(t)^T u(t) + \xi_z(t), \quad t \in \mathcal{T} \quad (3.25)$$

и выражение для оцениваемого параметра:

$$l = a^T q(\tau) + \eta_l, \quad (3.26)$$

( $\eta_l = 0$ , если пренебрегаем ошибками линеаризации). Здесь  $R(t)$ ,  $F(t)$  — известные матричные функции размерности  $(m \times m)$ , и  $(m \times r)$  соответственно,  $h(t) \in \mathbb{R}^m$ ,  $g(t) \in \mathbb{R}^r$  — известные вектор-функции,  $a \in \mathbb{R}^m$  — известный вектор. Функции  $R(t)$  и  $F(t)$  вычисляются по формулам

$$\begin{aligned} R(t) &= \left. \frac{\partial f(X(t), u(t), t)}{\partial X(t)} \right|_{X(t)=\overset{\circ}{X}(t), u(t)=0}, \\ F(t) &= \left. \frac{\partial f(X(t), u(t), t)}{\partial u(t)} \right|_{X(t)=\overset{\circ}{X}(t), u(t)=0}, \\ h(t) &= \left. \frac{\partial \varphi(X(t), u(t))}{\partial X(t)} \right|_{X(t)=\overset{\circ}{X}(t), u(t)=0}. \end{aligned} \quad (3.27)$$

где функция  $f(\cdot)$  определена в (3.1), а  $g(t)$  определена в (3.18).

Представим модель измерений (3.25) как сумму двух слагаемых: первое зависит лишь от вектора состояния в момент  $t_0 = 0$ , второе есть суммарная ошибка в модели (3.25), возникающая вследствие ошибок измерения и немоделируемых возмущений. Для этого выпишем решение системы (3.24) при начальном условии  $q(0) = \theta$ . Имеем согласно формуле Коши:

$$\begin{aligned} q(t) &= Q(t)[\theta + s(t)], \\ s(t) &= \int_0^t \Psi(s)u(s) ds, \\ \Psi(s) &= Q^{-1}(s)F(s), \end{aligned} \quad (3.28)$$

$Q(t)$  — фундаментальная матрица  $m \times m$ , определяемая из системы дифференциальных уравнений

$$\dot{Q}(t) = R(t)Q(t), \quad Q(0) = I_m, \quad (3.29)$$

$I_m$  — единичная матрица  $m \times m$ ,  $m$  — размерность вектора состояния.

Подставляя (3.29) в (3.25), получаем новую запись для модели измерений и оцениваемого параметра:

$$z(t) = H(t)^T \theta + \xi(t), \quad (3.30)$$

$$H(t) \doteq Q(t)^T h,$$

$$l = b^T \theta + \eta, \quad b \doteq Q(\tau)^T a, \quad (3.31)$$

где

$$\xi(t) = \xi_z(t) + H(t)^T s(t) + g(t)^T u(t) \quad (3.32)$$

— суммарная ошибка исходных данных, обусловленная ошибками измерений и немоделируемыми возмущениями,

$$\eta = \eta_l + b^T g(\tau) \quad (3.33)$$

— аналогичная ошибка в модели оцениваемого параметра.

Для рассматриваемого случая дискретных измерений запишем модель (3.30) в векторном виде:

$$Z = H^T \theta + \xi, \quad (3.34)$$

где  $Z \in \mathbb{R}^n$  — вектор измерений с компонентами  $z(t_i)$ ,  $i = 1, \dots, n$ ;

$$H = (h(t_1), \dots, h(t_n)) \quad (3.35)$$

— матрица размерности  $m \times n$ , которую будем называть матрицей измерений.

Для модели (3.30) рассмотрим обобщенный метод наименьших квадратов. В соответствии с разбиением (3.15) на векторы определяемых и мешающих параметров в линейной модели (3.30), аналогично раскладывается вектор  $\theta$ :

$$\theta = \{q_{10}, q_{20}\}. \quad (3.36)$$

Вектор  $q_{20}$  принимается равным своему априорному значению  $\tilde{q}_{20}$ , а априорное значение  $\tilde{q}_{10}$  может рассматриваться как вектор измерений и включаться в вектор измерений (3.34). Заметим, что линеаризацию естественно проводить относительно опорной траектории, соответствующей известным априорным значениям вектора  $X_0$ . При этом  $\tilde{\theta} = 0_m \doteq (0, \dots, 0)^T$ .

В связи с представлением (3.36) запишем известные матрицы в (3.31), (3.34) в виде составных матриц:

$$b = \{b_1, b_2\}, \quad H = \{H_1, H_2\}. \quad (3.37)$$

Для модели (3.36) обобщенный метод наименьших квадратов (3.22) записывается следующим образом:

$$\hat{q}_{10} = \arg \min_{q_{10}} S^T P S, \quad (3.38)$$

где

$$S = Z - H_2^T \tilde{q}_{20} - H_1^T q_{10}, \quad (3.39)$$

— вектор невязок, рассчитанный при априорно заданном векторе  $q_{20}$ ,  $P$  — весовая матрица. При этом оценка интересующего нас параметра имеет вид

$$\hat{l} = b_1^T \hat{q}_{10} + b_2^T \tilde{q}_{20}. \quad (3.40)$$

Минимизация в (3.38) приводит к известному факту, что оценка метода наименьших квадратов есть линейная функция вектора измерений и априорного значения вектора мешающих параметров. А именно, решение задачи (3.38) согласно разделу 1.1 есть:

$$\hat{q}_{10} = (H_1 P H_1^T)^{-1} H_1 P (Z - H_2^T \check{q}_{20}). \quad (3.41)$$

Подставляя (3.41) в (3.40), получаем выражение для оценки (3.40) в зависимости от исходных данных:

$$\hat{l} = x^T Z + (b_2 - H_2 x)^T \check{q}_{20}, \quad (3.42)$$

где

$$x^T = b_1^T (H_1 P H_1^T)^{-1} H_1 P \quad (3.43)$$

— вектор коэффициентов, называемый в дальнейшем оценителем.

Таким образом, оценка наименьших квадратов (3.38) приводит для линейной модели (3.30) к линейной оценке относительно исходных данных  $Z$  и  $\check{q}_{20}$ . Это вывод, конечно, следует и из формул (2.17), (2.18), но мы предпочли здесь для ясности непосредственный вывод для скалярного параметра  $l$ . Также прямыми вычислениями (без использования соотношения (2.19)) найдем ошибку  $\delta l$  оценки (3.42). Для этого подставим в (3.42) выражение (3.34) для вектора измерений и вычтем из полученного выражения истинное значение (3.31) оцениваемого параметра. Имеем

$$\delta l = \hat{l} - l = x^T \xi + (b_2 - H_2 x)^T (\check{q}_{20} - q_{20}) + (H_1 x - b_1)^T q_{10} - \eta. \quad (3.44)$$

Согласно (3.43),

$$H_1 x = b_1. \quad (3.45)$$

Поэтому окончательно имеем

$$\delta l = x^T \xi + (b_2 - H_2 x)^T \delta q_{20} - \eta, \quad (3.46)$$

где

$$\delta q_{20} = \check{q}_{20} - q_{20} \quad (3.47)$$

— ошибка априорного значения вектора мешающих параметров или, как мы будем говорить, ошибка априорного измерения.

Согласно (3.46), ошибка оценки равна нулю, если равны нулю ошибки исходных данных  $\xi$ ,  $\delta q_{20}$ ,  $\eta$ . Поэтому алгоритм метода наименьших квадратов

является несмещенным. Несмещенным будет также любой линейный алгоритм (3.42) с произвольным оценителем, удовлетворяющим условию (3.43). В связи с этим условие (3.43) называется условием несмещенности.

Таким образом, ошибка линейной оценки (3.42) для модели (3.34) имеет вид (3.46). В заключение параграфа приведем еще одну форму записи выражения (3.46), которая будет использована в следующем параграфе. Введем расширенный вектор измерений

$$\tilde{Z} = \begin{pmatrix} Z \\ \tilde{q}_{20} \end{pmatrix}, \quad (3.48)$$

включающий наряду с измерениями и априорные значения вектора мешающих параметров, а также расширенный вектор ошибок

$$\tilde{\xi} = \begin{pmatrix} \xi \\ \delta q_{20} \end{pmatrix}, \quad (3.49)$$

соответствующий вектору (3.48). Напомним, что в вектор измерений  $Z$  могут входить априорные значения компонент вектора определяемых параметров  $q_{10}$ . Введем также расширенную матрицу измерений

$$\tilde{H}^T = \begin{pmatrix} H_1^T & H_2^T \\ O & I_{m_2} \end{pmatrix}, \quad (3.50)$$

соответствующую вектору (3.48).

При введенных обозначениях и условии, что вектор мешающих параметров полагается равным своему априорному значению, модель измерений (3.34) записывается в виде

$$\tilde{Z} = \tilde{H}^T \theta + \tilde{\xi} \iff \begin{cases} Z = H^T \theta + \xi, \\ \tilde{q}_{20} = q_{20} + \delta q_{20}, \end{cases} \quad (3.51)$$

а линейный алгоритм оценивания (3.42) — в виде

$$\hat{l} = \tilde{x}^T \tilde{Z}, \quad (3.52)$$

где составная вектор-строка

$$\tilde{x}^T = (x^T, b_2^T - x^T H_2^T) \quad (3.53)$$

есть обобщенный оценщик. Этот оценщик удовлетворяет условию несмещенности

$$\tilde{H}\tilde{x} = b, \quad (3.54)$$

которое следует из аналогичного условия (3.45). Подставляя (3.51) в (3.52) и вычитая из полученного выражения истинное значение (3.31), получаем с учетом условия (3.54) выражение для ошибки оценивания:

$$\delta l = \tilde{x}^T \tilde{\xi} - \eta. \quad (3.55)$$

Отметим, что это выражение есть просто другая запись равенства (3.46) при введенных обозначениях.

Итак, оценка наименьших квадратов с любой весовой матрицей является линейной несмещенной оценкой общего вида (3.52), (3.54), что также является следствием теоремы 1.4 из пункта 1.1.1. Согласно этой теореме справедливо и обратное утверждение: каждая линейная оценка (3.52), удовлетворяющая условию (3.54), есть оценка (3.40) метода наименьших квадратов с некоторой весовой матрицей, причем разбиение вектора  $\theta$  на вектор определяемых и мешающих параметров может быть выбрано произвольно.

Таким образом, имеет место совпадение множества линейных несмещенных оценок (3.52) (т.е. всех оценок, определяемых из (3.52) при всевозможных  $\tilde{x}$ , удовлетворяющих условию (3.54)) и множества оценок метода наименьших квадратов, соответствующих всевозможным весовым матрицам при произвольном выборе вектора  $q_{20}$  мешающих параметров ( $q_{20} \neq q_2$ ). Поэтому при решении оптимальных задач оценивания можно рассматривать только несмещенную оценку в виде (3.52) с ошибкой (3.55). После нахождения оптимального оценщика можно удобным способом выбрать вектор мешающих параметров с возможным последующим применением метода наименьших квадратов уже в нелинейной задаче.

### 1.3.2. Вычисление гарантированной ошибки линейного оценивания

В связи с изложенным в конце предыдущего параграфа будем рассматривать линейную оценку вида (3.52), где оценщик удовлетворяет условию несмещенности (3.54). Перепишем ошибку (3.55) этой оценки для случая, когда измерения могут быть непрерывными. При этом введем непрерывный оценщик  $x(t)$ , так что

$$\hat{l} = \int_{\mathcal{T}} x(t)z(t) dt, \quad (3.56)$$

где  $\mathcal{T}$ , как и ранее, — множество измерений, а под  $z(t)$  мы понимаем теперь в модели (3.25) обобщенную функцию измерений, включающую и априорный вектор измерений  $\tilde{q}_{20}$ , который можно считать проведенным в момент  $t_1$  ( $t_1 \leq t \in \mathcal{T}$ ). Если набор измерений дискретен, то в (3.56) следует положить

$$x(t) = \sum_{i=1}^n x_i \delta(t - t_i), \quad (3.57)$$

где  $t_i$  — моменты измерений, упорядоченные согласно (3.13), а  $\delta(\cdot)$  — дельта-функция. Условие несмещенности (3.54) для непрерывного случая записывается так:

$$\int_{\mathcal{T}} x(t)H(t) dt = c, \quad (3.58)$$

где для априорных измерений в модели (3.25)  $h^T(t) = (0, \dots, 0, 1, 0, \dots, 0)$  (единица соответствует заданной компоненте вектора  $\theta$ ).

Ошибка оценки (3.56) записывается в виде, аналогичном (3.55). При этом для ошибки модели используется выражение (3.33), где примем в (3.26)  $\eta_l = 0$ . Тогда получим

$$\delta l = \int_{\mathcal{T}} x(t)\xi(t) dt - b^T g(\tau), \quad (3.59)$$

где  $\xi(t)$  — суммарная ошибка (3.32),  $s(t)$  определена в (3.28).

Представим теперь предпоследнее слагаемое в (3.59) в виде

$$b^T g(\tau) = b^T \int_0^{\tau} \Psi(t)u(t) dt = \int_0^{\tau} b^T \Psi(t)u(t)I(\tau - t) dt, \quad (3.60)$$



где

$$I(y) = \begin{cases} 0, & y \leq 0, \\ 1, & y > 0, \end{cases} \quad (3.61)$$

— функция Хевисайда, функция  $\Psi(t)$  определена в (3.28).

Введем расширенный оценитель на  $[0, T]$ :

$$\bar{x}(t) = \begin{cases} x(t), & t \in \mathcal{T}, \\ 0, & t \notin \mathcal{T}. \end{cases} \quad (3.62)$$

Используя (3.60), подставляя  $s(t)$  из (3.28) в (3.32), меняя там пределы интегрирования и группируя слагаемые при  $u(t)$ , можно получить для ошибки оценивания следующее выражение:

$$\delta l = \int_0^{\tau} \bar{x}(t) \xi_z(t) dt + \int_0^{\tau} [\bar{x}(t)g(t) + \Psi(t)Y(t)]^{\top} u(t) dt, \quad (3.63)$$

где

$$Y(t) = \int_t^{\tau} \bar{x}(s)h(s) ds - bI(\tau - t). \quad (3.64)$$

Отсюда находится гарантированная ошибка (3.20):

$$\delta l_{\text{гар}} = \int_0^{\tau} |\bar{x}(t)|\sigma(t) dt + \int_0^{\tau} \|\bar{x}(t)g(t) + \Psi(t)^{\top}Y(t)\|\gamma(t) dt. \quad (3.65)$$

В случае дискретных измерений (3.13) используем подстановку (3.57) и получаем:

$$\delta l_{\text{гар}} = \delta l_{\text{гар}}^z + \delta l_{\text{гар}}^u,$$

где

$$\begin{aligned} \delta l_{\text{гар}}^z &\doteq \sum_{i=1}^n \sigma(t_i)|x_i|, \\ \delta l_{\text{гар}}^u &\doteq \sum_{i=1}^n \gamma(t_i)\|g(t_i)\| \cdot |x_i| + \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \gamma(t)\|\Psi(t)^{\top}Y_j(t)\| dt + \\ &\quad + \int_{\tau}^{\tau} \gamma(t)\|\Psi(t)^{\top}b\| dt. \end{aligned} \quad (3.66)$$

Здесь  $\delta l^z$  и  $\delta l^u$  — ошибки, обусловленные соответственно ошибками измерений и неучетом немоделируемых возмущений,

$$Y_j(t) = \sum_{i=j}^n x_i h(t_i) - bI(\tau - t). \quad (3.67)$$

Формулы (3.65) или (3.66) позволяют вычислить гарантированную ошибку оценивания (3.20) для рассматриваемой линейной модели (3.24) — (3.26).

**ЗАМЕЧАНИЕ 1.6.** Мы отмечали, что в (3.2) и, соответственно, в (3.66)  $\|\cdot\|$  означает некоторую норму. Если принято допущение, что немоделируемое возмущение обусловлено, например, истечением газов в космическом аппарате в неизвестном в каждый момент времени направлении, то целесообразно принять норму евклидовой, т.е.

$$\|u\|^2 = \sum_{i=1}^r u_i^2. \quad (3.68)$$

Если же в предыдущем примере возможно истечение газов в  $r$  заданных направлениях, то целесообразно принять

$$\|u\| = \max\{|u_1|, \dots, |u_r|\}. \quad (3.69)$$

**ЗАМЕЧАНИЕ 1.7.** Вместо допущений (3.2), (3.17) детерминированного характера о немоделируемых возмущениях и ошибках измерений можно рассматривать менее жесткие предположения, состоящие в следующем. Ошибки измерений есть также несмещенные случайные величины, удовлетворяющие ограничениям

$$E[\xi(t)\xi(s)] \leq \frac{\sigma(t)\sigma(s)}{k^2}, \quad (3.70)$$

где число  $k^2 = const$  задано,  $E[\cdot]$  — операция математического ожидания. Неравенство (3.70) означает, что дисперсия ошибки ограничена числом  $\frac{\sigma^2}{k^2}$ , а коэффициент корреляции между ошибками измерений в моменты  $t$  и  $s$  может быть любым. При этих допущениях ошибка  $\delta l$  уже должна рассматриваться как случайная величина, и в качестве ее характеристики точности (3.20) можно рассмотреть величину

$$\overline{\delta l_{\text{gap}}^2} = k^2 \max\{E[\delta l^z]^2 : (3.70)\} + \max\{(\delta l^u)^2 : (3.2)\}, \quad (3.71)$$

где  $\delta l^z$  и  $\delta l^u$  — соответствующие ошибки оценивания, обусловленные ошибками измерений и незнанием немоделируемого возмущения. Отметим, что согласно (3.70) мы принимаем отношение максимально возможной ошибки измерений к ее максимально возможной дисперсии не зависящим от времени, т.е. в обозначениях (3.17), (3.70)

$$\frac{\sigma^2(t)}{E[\xi^2(t)]} = k^2. \quad (3.72)$$

Для нормально распределенной ошибки можно взять  $k = 3$ . При этих допущениях величина (3.71) есть максимальное значение ошибки оценивания при допущении (3.70). Если принять предположение (3.72), то для величины (3.71) имеет место формула (сравни с (3.66))

$$\overline{\delta l_{\text{гар}}^2} = (\delta l_{\text{гар}}^z)^2 + (\delta l_{\text{гар}}^u)^2, \quad (3.73)$$

где  $\delta l_{\text{гар}}^z$  и  $\delta l_{\text{гар}}^u$  определены в формуле (3.66). Менее жесткие по сравнению с (3.17) допущения (3.70) уменьшают гарантированную ошибку:

$$\overline{\delta l_{\text{гар}}} \leq \delta l_{\text{гар}}. \quad (3.74)$$

**ЗАМЕЧАНИЕ 1.8.** Рассмотрим случай, когда физический смысл задачи оценивания позволяет еще более ослабить ограничения на характеристики точности измерений. Пусть имеются  $s$  составов измерений с соответствующими ошибками  $\xi^i(t)$ ,  $i = 1, \dots, s$ . При этом относительно каждого состава выполняется условие (3.70), а между собой ошибки измерений из разных составов не коррелированы. Тогда в (3.73)

$$(\delta l_{\text{гар}}^z)^2 = \sum_{i=1}^s (\delta^i l_{\text{гар}}^z)^2, \quad (3.75)$$

где  $\delta^i l_{\text{гар}}^z$  определено в (3.66) при  $\xi(t) = \xi^i(t)$ .

Пусть и немоделируемые возмущения представляются как несмещенные случайные величины, причем вектор этих возмущений разбит на  $p$  подвекторов  $u^i(t)$ ,  $i = 1, \dots, p$ , так что каждый вектор  $u^i(t)$  подчинен ограничениям типа (3.70) (т.е. допускается произвольная корреляция), а между собой эти векторы не коррелированы. Тогда в (3.73)

$$(\delta l_{\text{гар}}^u)^2 = \sum_{i=1}^p (\delta^i l_{\text{гар}}^u)^2, \quad (3.76)$$

где  $\delta^i l_{\text{гар}}^u$  определено в (3.66) при  $\gamma(t) = \gamma^i(t)$  и соответствующим выделением в матрицах  $\Psi(t)$  и других матрицах  $i$ -й части, относящейся к вектору  $u^i(t)$ .

## Глава 2

# Простейшие задачи оптимального оценивания и коррекции и их сведение к задачам линейного программирования

### 2.1. Классический и гарантирующий подходы к оптимизации оценивателя, их преимущества и недостатки

Пусть измерения  $y_i$  ( $i=1, \dots, n$ ) представляются в виде

$$y_i = H_i^T \theta + \xi_i,$$

где  $H_i \in \mathbb{R}^m$  — известные векторы ( $m \leq n$ ),  $\xi_i$  — ошибки измерений, а  $\theta$  — вектор оцениваемых параметров. В векторной форме модель измерений запишем в виде

$$y = H^T \theta + \xi,$$

где  $H$  — матрица, строки которой есть  $H_i$ , а векторы  $y, \xi \in \mathbb{R}^n$ . Нас будет интересовать значение заданной линейной функции от  $\theta$ :  $l = b^T \theta$ , где вектор  $b$  задан.

Рассмотрим линейную оценку параметра  $l$ :  $\hat{l} = x^T y$ , где  $x$  — вектор, определяющий эту оценку. Наложим условие несмещенности на алгоритм оценивания: если  $\xi = 0$ , то  $\hat{l} = l$  при любых  $\theta$ . Это условие несмещенности

запишется, очевидно, в виде

$$Hx = b,$$

при этом ошибка оценки есть линейная функция ошибок измерений:

$$\delta l = x^T \xi.$$

### 2.1.1. Классический подход к оптимизации оценщика и его практические недостатки

Пусть вектор ошибок есть несмещенный случайный вектор с известной матрицей ковариаций  $K \equiv E(\xi\xi^T)$ . Тогда дисперсия оценки  $\hat{l}$  скалярного параметра  $l$  равна

$$D(\hat{l}) = x^T K x.$$

Минимум этой дисперсии по оценщику — вектору  $x$  при условии несмещенности  $Hx = b$  достигается при  $x^T = b^T (HK^{-1}H^T)^{-1} HK^{-1}$ . Это устанавливается с помощью множителей Лагранжа. Мы видим, что оценщик получился такой же, как если бы использовался метод наименьших квадратов с весовой матрицей  $W = K^{-1}$ . Таким образом, с точки зрения минимизации дисперсии ошибки оценивания любого скалярного параметра оптимальным является метод наименьших квадратов с весовой матрицей, обратной матрице ковариаций. Этот известный результат есть теорема Гаусса-Маркова. Оптимальная дисперсия при этом равна

$$D_0 = b^T K_\theta b, \quad K_\theta \doteq (HK^{-1}H^T)^{-1}.$$

Здесь  $K_\theta$  — ковариационная матрица оценки вектора состояния. При довольно естественных условиях на матрицы  $K$  и  $H$  дисперсия  $D_0$  стремится к нулю с увеличением числа измерений  $n$  [84]. На практике однако такой картины не наблюдается. Простой пример приведен в книге [85]. Пусть измеряются показания различных часов одной точности и находится среднее арифметическое всех таких показаний. Такая оценка есть оценка метода наименьших квадратов при отсутствии корреляции. При увеличении числа показаний мы должны были бы получить сколь угодно точное время, но это не будет иметь место на практике,

так как существует корреляция между ошибками в точности хода, обусловленная преемственностью способов изготовления часов, и эта корреляция может меняться в зависимости от времени их изготовления. В связи с изложенным можно сказать, что классический подход к вычислению точности оценивания может давать слишком оптимистические результаты.

### 2.1.2. Гарантирующий подход к вычислению точности оценивания.

Этот подход может дать наоборот более пессимистические прогнозы точности, но для практики иногда важно перестраховаться. Поясним сущность этого подхода на рассмотренном выше примере вычисления дисперсии оценки. Пусть нам известно лишь множество  $\mathcal{K}$ , которому принадлежит матрица ковариаций. Тогда в наихудшем с точки зрения величины дисперсии случае мы получим гарантированную дисперсию

$$D_{\text{гар}} = \max_{K \in \mathcal{K}} D(\delta l).$$

Отметим, что гарантированная дисперсия вычисляется при заданном оценителе  $x$ , т.е., например, при заданной весовой матрице, которая уже не может выбираться в соответствии с теоремой Гаусса-Маркова, так как неизвестна матрица ковариаций ошибок. Вопрос об оптимизации оценителя рассмотрим ниже. Сейчас выпишем  $D_{\text{гар}}$  для простых случаев задания множества  $\mathcal{K}$ . Допустим, что множество  $\mathcal{K}$  определяется условиями

$$|k_{ij}| \leq k \leq 1 \quad (i \neq j),$$

где  $k_{ij}$  — коэффициенты корреляции между ошибками измерений, число  $k$  задано, а дисперсии ошибок известны и для простоты записи приняты далее единицами. Кроме того, следует наложить необходимое для корреляционной матрицы условие неотрицательной определенности, которое будем записывать в виде  $K \geq 0$ . Тогда

$$D_{\text{гар}} \leq D_k,$$

где  $D_k$  вычисляется без условия  $K \geq 0$ . В конце этого раздела мы поясним, что на самом деле это условие можно не учитывать. Тогда

$$\begin{aligned} D_{\text{гар}} = D_k &\doteq \max \left\{ x^T K x = \sum_i x_i^2 + \sum_{i \neq j} x_i x_j k_{ij} : |k_{ij}| \leq k \right\} = \\ &= (1 - k) \sum_i x_i^2 + k \sum_i x_i^2 + \max \left\{ \sum_{i \neq j} x_i x_j k_{ij} : |k_{ij}| \leq k \right\} \end{aligned}$$

или, окончательно,

$$D_{\text{гар}} = D_k = (1 - k)D_0 + kD_1, \quad D_0 = \sum_i x_i^2, \quad D_1 = \left( \sum_i |x_i| \right)^2.$$

Здесь  $D_0$  есть дисперсия оценки метода наименьших квадратов при некоррелированных измерениях,  $D_1$  — гарантирующая дисперсия при условии, что корреляция может быть произвольной, т.е. формально при условии  $|k_{ij}| \leq 1$ , которое не является ограничением на коэффициенты корреляции, так как выполняется всегда. Как указывалось выше, очень часто  $D_0 \rightarrow 0$  при увеличении числа измерений. Поэтому при большом числе измерений в этих случаях  $D_{\text{гар}} \approx kD_1$ . Наиболее оптимистическая и пессимистическая дисперсии удовлетворяют неравенству Коши-Буняковского

$$D_1 \leq nD_0.$$

Ниже мы покажем, следуя Элфвингу ([89]), что если возможно повторение измерений, то при наиболее рациональном повторении здесь будет наблюдаться равенство. Отметим более общий случай [18], когда коэффициенты корреляции принадлежат многомерному прямоугольному параллелепипеду:

$$-W \leq K - K^* \leq W,$$

где  $K^*$  — некоторая известная номинальная корреляционная матрица, а матрица  $W$  имеет неотрицательные элементы и задает разброс около  $K^*$  ( $k_{ii} = 1$ ,  $w_{ii} = 0 \forall i$ ). В этом случае

$$D_{\text{гар}} \leq D_w = x^T K^* x + x_+^T W x_+,$$

где вектор  $x_+$  состоит из компонентов  $|x_i|$ . Критерий равенства  $D_{\text{гар}} = D_w$  приведен в книге [18], но является трудно проверяемым. Однако имеется простое



достаточное условие этого равенства:

$$K^* - \Lambda \geq 0, \quad W + \Lambda \geq 0$$

при некоторой диагональной матрице  $\Lambda$ . Если  $K^* = \Lambda = I$ , где  $I$  — единичная матрица, то в случае  $W + I \geq 0$  имеем  $D_{\text{гар}} = D_w$ . Отсюда следует упомянутое выше равенство  $D_{\text{гар}} = D_k$ .

## **2.2. Сравнение решений задач оптимального оценивания в двух простейших случаях при гарантирующем и классическом подходах**

Здесь мы покажем удивительное совпадение оптимальных программ измерений при двух различных подходах к модели ошибок измерений. Сначала рассмотрим классический подход.

### **2.2.1. Задача о выборе оптимального оценщика при возможности повторения измерений и ограничении на их общее число**

Рассмотрим опять классический случай, когда все измерения не коррелированы, но допускается повторение измерений. Это означает, что возможно проведение нескольких измерений  $r_i$  с одним и тем же уравнением измерений, т.е.

$$y_{ij} = H_i^T \theta + \xi_{ij}, \quad j = 1, \dots, r_i, \quad i = 1, \dots, n.$$

Удобно считать, что измерения  $y_{ij}$  при заданном  $i$  есть измерения в заданный момент времени  $t_i$ . Тогда можно измерения в этот момент считать сеансом измерения. Например, космические измерения проводятся сеансами, соответствующими небольшим интервалам времени, в течение которого есть радиовидимость космического объекта. Эти интервалы часто можно при планировании космического

эксперимента заменить одним моментом  $t_i$ . Неограниченное повторение измерений на практике недопустимо из-за ограничений по времени или энергии для осуществления измерений. Поэтому естественно наложить условие  $n_1 + \dots + n_n \leq N$ , где  $N$  задано. При этом условии и условии несмещенности запишем задачу оптимизации дисперсии оценки. Согласно разделу 2.1.1

$$D_0^* = \min_{x_{ij}, r_i} \left\{ D_0 = \sum_{i=1}^n \sum_{j=1}^{r_i} x_{ij}^2 \right\}.$$

Отметим сначала, что минимум по  $x_{ij}$  суммы  $\sum_{j=1}^{r_i} x_{ij}^2$  при условии  $\sum_{j=1}^{r_i} x_{ij} = x_i$  достигается при  $x_{ij} = x_i/r_i$  для всех  $j$ , т.е. все повторяющиеся измерения в оптимальном случае оцениваются с одним оценителем (это ясно и из соображений симметрии). Поэтому можно записать

$$D_0 = \min_{x_i, r_i} \sum \frac{x_i^2}{r_i}.$$

Проведем минимизацию по  $r_i$ , пренебрегая тем, что  $r_i$  должно быть целым. Это оправдано при достаточно большом  $N$ . Получим, используя множители Лагранжа, что оптимальное повторение измерений есть

$$r_i = \frac{|x_i|}{\sum |x_k|} N.$$

При этом

$$D_0 = \frac{1}{N} \left( \sum |x_i| \right)^2.$$

Как будет отмечено ниже в замечании 4.4 из главы 4 указанная оптимизация по  $r_i$  здесь проведена не совсем строго (так же как и в основополагающей работе [89]). Оптимизация величины  $D_0$ , полученной при оптимальных  $r_i$ , по оценителю  $x$  приводит к задаче линейного программирования

$$\sigma_0^* \doteq \sqrt{D_0^*} = \frac{1}{\sqrt{N}} \min \left\{ \sum |x_i| : Hx = b \right\} = \min \left\{ \sum |x_i| : \sum x_i H_i = b \right\}.$$

Введем новые переменные  $y_i = |x_i| \geq 0$ . Тогда  $x_i H_i = y_i a_i$ , где

$$a_i = \begin{cases} H_i, & x_i \geq 0, \\ -H_i, & x_i < 0. \end{cases}$$

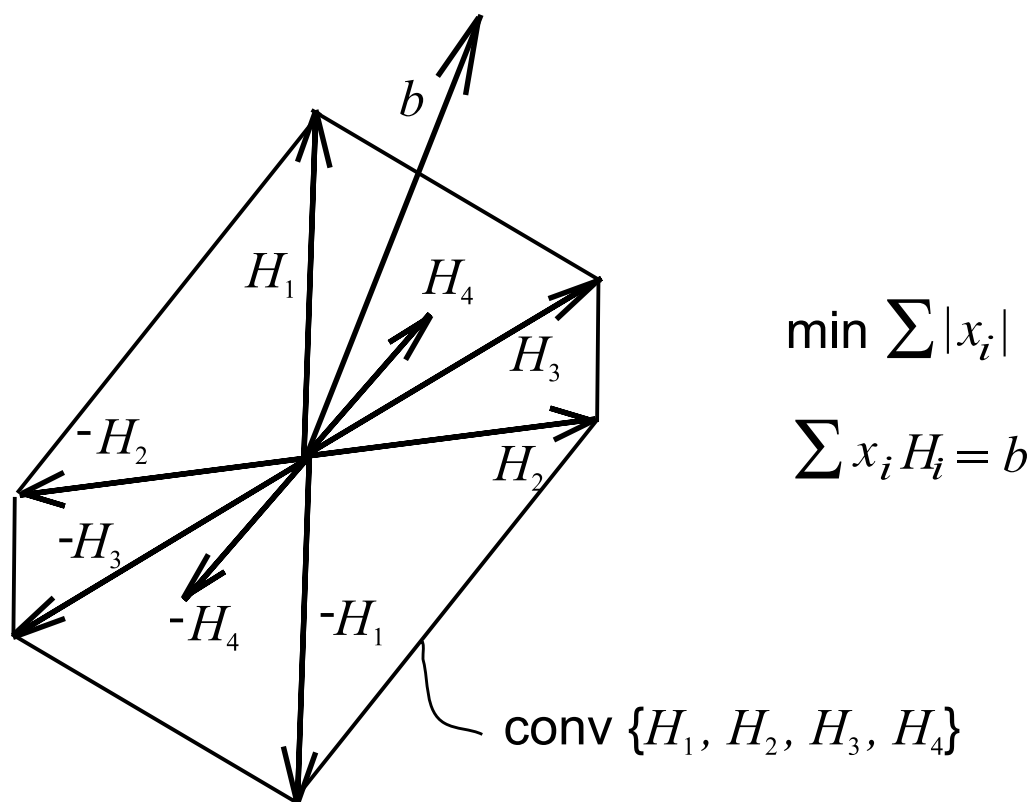


Рис.4.

Тогда оптимальную задачу можно записать в виде

$$\sigma_1^* = \min_{y_i, a_i} \left\{ \sum y_i : \sum y_i a_i = b, y_i \geq 0, a_i \in \mathcal{H}_i \right\},$$

где  $\mathcal{H}_i = \{H_i, -H_i\}$ , т.е. это множество состоит из двух точек (см. рис.4).

Эту задачу можно интерпретировать так: следует разложить вектор  $b$  по векторам  $\pm H_i$  с коэффициентами  $y_i > 0$  и минимизировать сумму коэффициентов. Такая задача представляет собой задачу линейного программирования. Из теории линейного программирования известно, что существует решение, содержащее не более  $m$  не равных нулю коэффициентов, т.е. вектор  $b$  лежит внутри многогранного угла, образованного некоторыми  $m$  векторами  $\pm H_i$  (базисом). Эти векторы таковы, что гиперплоскость, проведенная через их концы, отсекает от вектора  $b$  наибольшую часть по сравнению с другими комбинациями из  $m$  векторов. Таким образом мы делаем важный вывод. При наиболее неблагоприятной корреляции между ошибками измерений следует выбрать оптимальные  $m$  измерений и найти оценитель из соотношения

$$H^0 x^0 = b \Rightarrow x^0 = H^{0-1} b,$$

где  $H^0$  — матрица  $m \times m$ , составленная из оптимальных  $H_i$ . Это соответствует тому, что вектор  $\theta$  находится из  $m$  уравнений

$$y_i = H_i^T \theta, \quad i = 1, \dots, m,$$

где нумерация по  $i$  условна. Такой вывод получается потому, что допускается произвольная корреляция между измерениями, когда одни измерения могут "портить" другие. Опишем нетрадиционную геометрическую интерпретацию симплексного алгоритма нахождения оптимального базиса [44]. Этот базис удовлетворяет следующему условию: гиперплоскость, проходящая через концы векторов базиса, принадлежит выпуклой оболочке всех  $2n$  векторов  $\pm H_i$ ,  $i=1, \dots, n$  (см. рис.4).

Для нахождения оптимального базиса сначала берется некий базис из  $m$  векторов (на рис.5  $m = 2$  и базис есть  $\{H_1, H_2\}$ ). Потом проводится гиперплоскость через концы векторов базиса (прямая через концы  $H_1$  и  $H_2$ ). Если

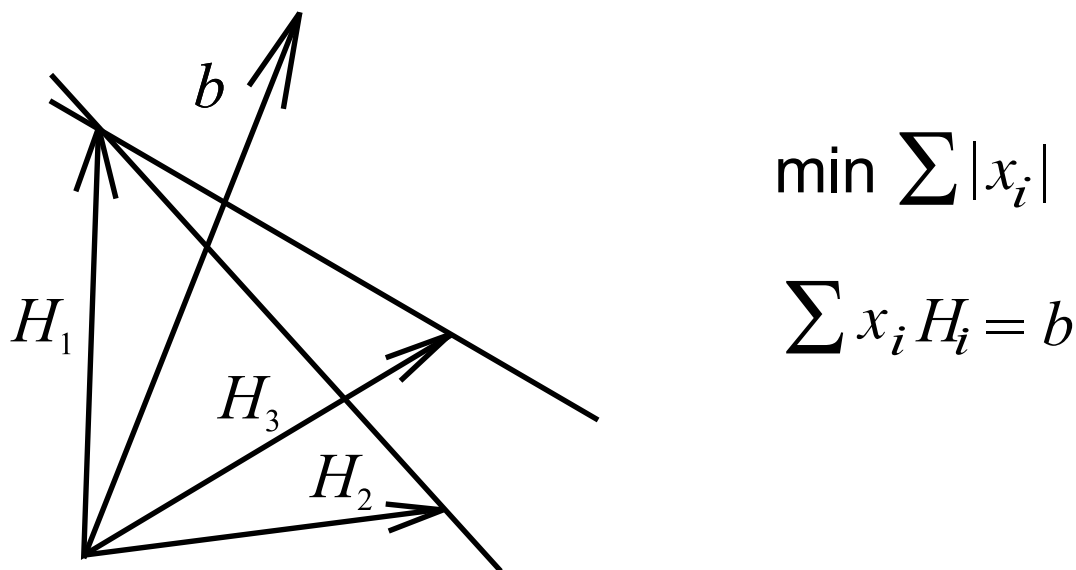


Рис.5.

нет векторов, пересекающих эту плоскость, то базис оптимален. На рис.5 вектор  $H_3$  пересекает эту плоскость. Если его ввести в базис и удалить оттуда вектор  $H_1$ , то новые коэффициенты  $y_1, y_2$  также положительны, и базис дает меньшее значение целевой функции. Этот упорядоченный перебор и составляет идею симплексного метода. Симплексный алгоритм хорошо разработан и эффективен.

### 2.2.2. Оптимизация гарантированной дисперсии $D_1$

Рассмотрим теперь задачу оптимизации алгоритма оценивания с точки зрения минимизации гарантированной дисперсии. Она имеет вид

$$\min\{D_{\text{гар}} : Hx = b\}.$$

Рассмотрим задачу минимизации  $D_1$  (напомним, что  $D_0$  соответствует известной ковариационной матрице и найдена выше). Имеем

$$\sigma_1^* = \min\left\{\sigma_1 \doteq \sqrt{D_1} : Hx = b\right\} = \min\left\{\sum |x_i| : \sum x_i H_i = b\right\}.$$

Введем новые переменные  $y_i = |x_i| \geq 0$ . Тогда  $x_i H_i = y_i H_i$ , где

$$H_i = \begin{cases} H_i, & x_i \geq 0, \\ -H_i, & x_i < 0. \end{cases}$$

Тогда оптимальную задачу можно записать в виде задачи линейного программирования

$$\sigma_1^* = \min_{y_i, H_i} \left\{ \sum y_i : \sum y_i H_i = b, y_i \geq 0, H_i \in \mathcal{H}_i \right\},$$

и величина  $D_1$  связана с величиной  $D_0$  соотношением

$$D_0 = \frac{D_1}{N},$$

о чем упоминалось в пункте 2.2.1. Отсюда мы приходим к важному методологическому выводу [2]: в случае повторяющихся некоррелированных измерений имеется  $m$  оптимальных точек повторения измерения, которые совпадают с оптимальными точками проведения измерений ( $x_i \neq 0$ ) для задачи минимизации гарантирующей дисперсии  $D_1$ , полученной при возможности произвольной корреляции между измерениями. При этом  $D_1 = ND_0$ , а число повторений  $r_i$  измерений пропорционально оценителю  $x_i$   $i$ -го измерения в случае произвольной корреляции.

### 2.2.3. Минимаксная задача оценивания при ограниченных по модулю ошибках измерений.

Как указывалось, такая задача рассматривалась Лидовым [43] и названа им "случай бортика". Предполагается, что ошибки измерений ограничены по модулю:

$$|\xi_i| \leq M_i,$$

где  $M_i$  — известные числа. Тогда гарантированная ошибка измерения равна

$$\delta l_{\text{гар}} = \max_{\xi_i} \sum x_i \xi_i = \sum M_i |x_i|,$$

а минимизация ее сводится к задаче линейного программирования того же типа, что и в п. 3:

$$\delta l^* = \min \left\{ \sum M_i |x_i| : Ax = b \right\}.$$

Рассмотрим предположение  $M_i = M \forall i$ , которое можно трактовать как допущение о пропорциональности максимально возможных ошибок измерений их дисперсиям (которые в пункте 2.2.1 приравнены к единицам). Тогда задачи линейного программирования из пункта 2.2.1 и из этого пункта дают один и тот же оптимальный оценщик, при этом

$$\delta l^* = M \sigma_1^*.$$

Отметим, что вместо условий на ошибки измерений мы могли бы рассматривать аналогичные условия на их математические ожидания:

$$|E(\xi_i)| \leq M_i.$$

При этом гарантированное математическое ожидание есть

$$M_{\text{гар}} = \max \{ |M(\delta l)| : |E(\xi_i)| \leq M_i \} = \sum M_i |x_i|,$$

а задача минимизации  $M_{\text{гар}}$  по линейному несмещенному оценщику приводит к той же по виду задаче линейного программирования, что и задача нахождения  $\delta l^*$ , выписанная в начале этого раздела.

### 2.3. Оптимальная задача линейной идеальной коррекции и обобщенное линейное программирование

Пусть  $l \in \mathbb{R}^k$  — некоторый вектор параметров системы, который может быть изменен путем импульсной коррекции траектории системы. Эта коррекция заключается в мгновенном изменении некоторых фазовых координат в моменты  $t_1, \dots, t_n$ . Такое изменение в каждый момент  $t_i$  будем характеризовать корректирующим вектором (импульсом)  $u_i$  размерности  $k_i$ , принадлежащим евклидову пространству  $\mathbb{R}^{k_i}$ . Будем рассматривать идеальную линейную коррекцию, т.е. предполагать, что импульсы производятся без ошибок и изменение

вектора  $l$  в момент  $t_i$  равно  $B_i u_i$ , где  $B_i$  — матрица  $k \times k_i$ , характеризующая влияние импульса. Тогда для изменения вектора  $l$  на векторную величину  $b$  нужно импульсы  $u_i$  произвольно выбирать из условия

$$\sum_i B_i u_i = b.$$

Пусть затраты на коррекцию в момент  $t_i$  характеризуются величиной  $p_i(u_i)$ , где  $p_i(\cdot)$  — некоторая норма в  $\mathbb{R}^{k_i}$ . Для коррекции траектории космического аппарата важны два случая (см. главу 6:  $p_i(u_i) = \|u_i\|_2 \equiv |u_i| \equiv \sqrt{u_i^T u_i}$ ,  $p_i(u_i) = \|u_i\|_1 = \sum_j |u_{ij}|$ , где  $u_{ij}$  — составляющие вектора  $u_i$ ). При сделанных допущениях задача оптимизации суммарных затрат на коррекцию имеет вид

$$W = \min_{u_i} \left\{ \sum_i p_i(u_i) : \sum_i B_i u_i = b \right\}.$$

Решив эту задачу, мы найдем и оптимальные моменты приложения импульса, так как если  $u_i = 0$ , то в момент  $t_i$  импульс не проводится. Представим каждый вектор  $u_i$  в виде

$$u_i = x_i \alpha_i, \quad x_i \equiv p_i(u_i) \geq 0,$$

где вектор  $\alpha_i \in \mathbb{R}^{k_i}$  принадлежит множеству  $\{\alpha_i : p_i(\alpha_i) = 1\}$ . Тем самым  $x_i$  есть длина вектора  $u_i$  в метрике, определяемой нормой  $p_i(\cdot)$ , а  $\alpha_i$  есть единичный вектор в этой метрике, направленный вдоль импульса. Используя введенные новые переменные, запишем оптимальную задачу коррекции в виде

$$W = \min_{x_i, u_i} \left\{ \sum_i x_i : \sum_i x_i a_i = b, x_i \geq 0, a_i \in \mathcal{A}_i \right\},$$

где  $\mathcal{A}_i = \{a_i : a_i = B_i \alpha_i, p_i(\alpha_i) = 1\}$ . Эта задача по виду напоминает задачу линейного программирования, однако здесь переменными оптимизации являются не только числа  $x_i$ , но и векторы  $a_i$  в линейных условиях-равенствах, выбираемые из множеств  $\mathcal{A}_i$ . Такая задача называется обобщенной задачей линейного программирования [31, 58]. Приведение задачи коррекции к обобщенной задаче линейного программирования и алгоритм ее решения даны в статье Лидова [44]. Там, а также в книге [18] рассмотрены более общие постановки задач коррекции. Далее изложение следует этим работам. Геометрическая интерпретация нахождения оптимального решения вытекает из интерпретации для



нахождения обычной задачи линейного программирования, данной выше. Симплексный алгоритм также сводится к вводу в базис вектора  $A_k$ , который пересекает гиперплоскость, проходящую через концы векторов базиса. Проверка достаточного условия оптимальности, таким образом, сводится к проверке условия непересечения этой гиперплоскости любого из бесконечного множества векторов  $a_i \in \mathcal{A}_i$ ,  $i=1, \dots, n$ . Если это условие проверяется эффективно, то алгоритм симплекс-метода также эффективен. Для того, чтобы понять, насколько эффективен этот измененный для обобщенной задачи линейного программирования симплексный алгоритм, запишем достаточное условие оптимальности базиса в формульном виде. Пусть  $B = (A_1, \dots, A_m)$  — матрица базиса (нумерация условна), а вектор  $z_i = B^{-1}a_i$ ,  $i = m+1, \dots, n$  состоит из координат вектора  $a_i \in \mathcal{A}_i$  в базисе  $A_1, \dots, A_m$ ;  $u = (1, \dots, 1) \in \mathbb{R}^m$  — вектор коэффициентов при  $x_i$  в целевой функции рассматриваемой задачи. Рассмотрим величины

$$\beta_i = \max\{u^T z_i : a_i \in \mathcal{A}_i\}, \quad i = 1, \dots, n.$$

Достаточное условие оптимальности базиса имеет вид

$$\Delta_i \equiv 1 - \beta_i \geq 0, \quad i = 1, \dots, n.$$

Если это условие не выполняется, т.е.  $\beta_k > 1$  для некоторого  $k$ , то в базис вводится любой вектор  $A_k \in \mathcal{A}_k$  с максимальным коэффициентом  $x_k > 0$ , при котором определенный коэффициент  $x_r$  при базисном векторе  $A_r$  ( $1 \leq r \leq m$ ) становится равным нулю, а остальные базисные коэффициенты  $x_i$  остаются неотрицательными. Это означает, что вектор  $A_r$  выводится из базиса. В отличие от обычной задачи линейного программирования, в обобщенной задаче линейного программирования изложенный алгоритм требует бесконечного числа итераций. Если полученный указанным алгоритмом оптимальный базис не содержит более одного вектора с ненулевым весом  $x_i$  из каждого множества  $\mathcal{A}_i$ , то этот базис дает решение задачи оптимальной коррекции. Это условие может не выполняться только в случае, когда некоторое множество  $\mathcal{A}_i$  содержит плоскость, проходящую через совокупность  $a_{ij}$ ,  $j=1, \dots, r_i$  векторов из  $\mathcal{A}_i$ , имеющих ненулевые

веса  $x_{ij}$  (подробнее см. в главе 6). Но в этом случае один из векторов указанной совокупности векторов заменяется некоторой их выпуклой комбинацией — вектором

$$a_i = \sum_{j=1}^{r_i} \lambda_j a_{ij}, \quad \lambda_j = x_j / \sum_{k=1}^{r_i} x_{ik}.$$

После этого множество  $\mathcal{A}_i$  содержит лишь один вектор  $a_i$  с ненулевым весом, т.е. новый базис дает решение оптимальной задачи коррекции. Изложенное показывает, что рассмотренный алгоритм эффективен, если достаточно просто решаются на каждой итерации задачи нахождения величин  $\beta_i$ . Легко видеть, что эту подзадачу можно записать в виде

$$\beta_i = \max \{ \pi^T a_i : a_i \in \mathcal{A}_i \},$$

где  $\pi^T = u^T B^{-1}$  — известный вектор. Для случая, когда  $p_i = \|u_i\|_2$ , множество  $\mathcal{A}_i$  есть  $k_i$ -мерный эллипсоид

$$\mathcal{A}_i = \{ a_i = B_i \alpha_i : \|\alpha_i\|_2 = 1 \},$$

который есть линейное преобразование  $m$ -мерной сферы из  $\mathbb{R}^m$  в  $\mathbb{R}^{k_i}$ . В этом случае подзадача решается аналитически [44]:

$$\beta_i = \|\pi^T B_i\|_2.$$

Для другого указанного выше случая, когда  $p_i = \|u_i\|_1$ , подзадача также решается аналитически [18] и  $\beta_i$  есть максимальная по модулю компонента вектора  $\pi^T B_i$ . Таким образом, для представляющих практический интерес случаев решение оптимальной задачи коррекции осуществляется эффективным симплексным алгоритмом, который эквивалентен алгоритму решения обычной задачи линейного программирования, размерность которой (т.е. число линейных уравнений в ограничениях) равна  $m$ . В этом алгоритме приходится оперировать с матрицей базиса  $m \times m$ .

## Глава 3

# Критерии оптимальности и монотонные алгоритмы решения вырожденной и обобщенной задач линейного программирования

Излагается критерий оптимальности и соответствующий симплексный алгоритм решения вырожденной задачи линейного программирования. Этот алгоритм сводится к решению невырожденной задачи линейного программирования, меньшей размерности. Приводятся результаты решения практической задачи, показывающие эффективность использования нового алгоритма, проводится его сравнение с аналогичными алгоритмами, известными ранее. Приводится аналогичная теория для решения обобщенной задачи линейного программирования, в которой вырожденные итерации не являются исключительным случаем. Кроме того обосновываются вопросы сходимости алгоритма.

### 3.1. Теория решения вырожденной задачи линейного программирования

#### 3.1.1. Введение

Настоящий раздел посвящен вырожденным задачам линейного программирования. При решении таких задач симплексным методом часто возникают

большие последовательности итераций, на которых целевая функция практически не изменяется. Вероятность появления таких вырожденных итераций резко возрастает с увеличением размерности задачи и зачастую делает применение симплексного алгоритма бесполезным. Большинство классических методов теории вырожденного линейного программирования было посвящено лишь борьбе с за-цикливанием, избежать вырожденных итераций при использовании таких методов не удавалось. Эти способы сводились либо к специальному выбору выводимого из базиса вектора (лексикографическое правило и правило случайного выбора [31]), либо к выбору вводимого в базис вектора (правило Данцига [90]), либо к одновременному выбору обоих этих векторов (правило Блэнда [88]).

По-видимому, наиболее эффективным методом борьбы с вырожденностью является метод Вулфа, предложенный впервые в [102] и модифицированный в [100]. Этот метод требует решения на каждой вырожденной итерации вспомогательной задачи линейного программирования. Ее решение позволяет либо сделать вывод об оптимальности текущего базиса основной задачи, либо заменить в нем сразу несколько векторов, что приводит к уменьшению целевой функции. Таким образом, процесс поиска оптимального решения исходной задачи становится строго монотонным. Однако метод Вулфа имеет два недостатка. Во-первых, размерность вспомогательной задачи совпадает с размерностью исходной. Во-вторых, при решении вспомогательной задачи также могут возникнуть вырожденные итерации. В этом случае в процедуре преодоления вырожденности необходимо применять рекурсии.

Нами предложен алгоритм, предусматривающий решение на каждой вырожденной итерации вспомогательной задачи линейного программирования, которая с вероятностью 1 не является вырожденной и имеет меньше строк и столбцов в матрице ограничений, чем исходная. В настоящей работе дано оригинальное и дополненное изложение нового алгоритма, устраняющее некоторые неточности статьи [9], а также впервые представлены результаты решения с его помощью задач большой размерности. Кроме того, получены полезные модификации алгоритма и оценки для оптимума. Сравнение предлагаемого алгоритма с алгоритмом

Вулфа в модификации [100] показывает, что во многих случаях новый алгоритм оказывается более эффективным, хотя и требует дополнительных вычислений при построении вспомогательной задачи.

Будем рассматривать задачу линейного программирования, в которой по заданным векторам  $a_i, b \in \mathbb{R}^m$ ,  $i = 1, \dots, n$ ,  $c \in \mathbb{R}^n$ , требуется найти вектор  $x^* \in \mathbb{R}^n$  такой, что

$$c^T x^* = \min_x \left\{ c^T x : \sum_{i=1}^n x_i a_i = b, \quad x = (x_1, \dots, x_n)^T \geq 0 \right\}. \quad (1.1)$$

Оптимальное значение  $c^T x^*$  целевой функции  $c^T x$  далее будем называть *значением задачи* (1.1). Будем также считать, что линейная оболочка векторов  $a_1, \dots, a_n$  совпадает с  $\mathbb{R}^m$ , иначе следует перейти к соответствующему подпространству. Пусть  $x$  — *допустимое базисное решение* задачи (1.1), характеризуемое тем, что линейно независимы все  $k$  векторов-условий  $a_i$ , отвечающих положительным компонентам вектора  $x$ . Очевидно,  $k \leq m$ .

В случае  $k < m$  текущее допустимое базисное решение называется *вырожденным*, а число  $m - k$  — *степенью вырожденности*. Матрицу  $U$ , состоящую из столбцов

$$a_i, \quad i \in \mathcal{I}_+ \doteq \{i : x_i > 0, \quad 1 \leq i \leq n\},$$

будем называть *матрицей строгого базиса* [9] в отличие от составной *базисной матрицы*  $B = (U, V)$ , где  $V$  — такая произвольная  $m \times (m - k)$ -матрица (не обязательно составленная из векторов  $a_i$ ), что матрица  $B$  невырождена. Тогда пространство  $\mathbb{R}^m$  разлагается в прямую сумму

$$\mathbb{R}^m = \mathcal{L}(U) \oplus \mathcal{L}(V)$$

подпространств  $\mathcal{L}(U)$  и  $\mathcal{L}(V)$  размерности  $k$  и  $m - k$ , базисами которых являются столбцы соответственно матриц  $U$  и  $V$ . Иными словами,

$$a_i = \begin{cases} u_i, & i \in \mathcal{I}_+, \\ u_i + v_i, & i \in \mathcal{I}_0 \doteq \{1, \dots, n\} \setminus \mathcal{I}_+, \end{cases} \quad u_i \in \mathcal{L}(U), \quad v_i \in \mathcal{L}(V). \quad (1.2)$$

Пусть  $\pi_0 \in \mathbb{R}^m$  — любое решение неоднозначно разрешимой относительно вектора двойственных переменных  $\pi$  системы уравнений

$$c_i - \pi^T a_i = 0, \quad i \in \mathcal{I}_+ \quad \Leftrightarrow \quad \pi^T U = c_u^T, \quad (1.3)$$

где  $c_u$  состоит из компонент  $c_i$  при  $i \in \mathcal{I}_+$ . Рассмотрим величины

$$\Delta_i = c_i - \pi_0^T a_i, \quad 1 \leq i \leq n, \quad (1.4)$$

которые носят название *относительных оценок* [59]. Согласно (1.3),  $\Delta_i = 0$  при  $i \in \mathcal{I}_+$ . Достаточным условием оптимальности текущего допустимого базисного решения является выполнение неравенства [31]

$$\Delta_{\min} \doteq \min_{i \in \mathcal{I}_0} \Delta_i \geq 0, \quad (1.5)$$

где множество  $\mathcal{I}_0$  определено в (1.2). Условие (1.5) является необходимым для невырожденного допустимого базисного решения [59]. Найдем необходимые и достаточные условия оптимальности вырожденного допустимого базисного решения.

### 3.1.2. Основные теоремы

Рассмотрим вспомогательную задачу линейного программирования

$$\min_{y_i} \left\{ \sum_{i \in \mathcal{I}_0} \Delta_i y_i : \sum_{i \in \mathcal{I}_0} y_i v_i = d, \quad y_i \geq 0, \quad i \in \mathcal{I}_0 \right\}, \quad (1.6)$$

где  $d \in \mathcal{L}(V)$  — любой вектор, для которого совместны ограничения в (1.6).

**ЗАМЕЧАНИЕ 3.1.** Задача (1.6) имеет матрицу условий размером  $(m - k) \times (n - k)$ , так как в базисе, определяемом столбцами матрицы  $V$  (далее — в базисе  $V$ ), ее условия-равенства эквивалентны системе  $m - k$  скалярных уравнений. Ниже мы не будем пока переходить к базису  $V$ , чтобы не вводить новых координат.

**ТЕОРЕМА 3.1 (КРИТЕРИЙ ОПТИМАЛЬНОСТИ).** *Допустимое базисное решение задачи (1.1) оптимально тогда и только тогда, когда конечно значение задачи (1.6). Доказательство теоремы 3.1 приведено в Приложении.*

Алгоритм уменьшения целевой функции в задаче (1.1) при неоптимальном допустимом базисном решении базируется на следующей теореме.

**ТЕОРЕМА 3.2 (КРИТЕРИЙ НЕОПТИМАЛЬНОСТИ).** *Текущее допустимое базисное решение задачи (1.1) неоптимально тогда и только тогда, когда найдется множество индексов  $\mathcal{S} \subseteq \mathcal{I}_0$ ,  $|\mathcal{S}| \leq m - k + 1$  такое, что выполняются следующие условия:*

- 1) векторы  $a_i$ ,  $i \in \mathcal{S}$  линейно независимы;
- 2) существуют числа  $\lambda_i > 0$ ,  $i \in \mathcal{S}$ , такие, что

$$\sum_{i \in \mathcal{S}} \lambda_i v_i = 0, \quad \sum_{i \in \mathcal{S}} \Delta_i \lambda_i < 0; \quad \lambda_i > 0, \quad i \in \mathcal{S}. \quad (1.7)$$

*Доказательство.* Пусть  $\mathcal{J}_B \subseteq \mathcal{I}_0$  — множество, состоящее из  $m - k$  индексов, соответствующих векторам текущего базиса в задаче (1.6). По теореме 3.1 текущее допустимое базисное решение задачи (1.1) неоптимально тогда и только тогда, когда целевая функция задачи (1.6) не ограничена на множестве своих допустимых решений. Согласно теории линейного программирования [31], это эквивалентно тому, что на некотором шаге симплексного метода решения задачи (1.6) находится небазисный вектор  $v_p$ ,  $p \in \mathcal{I}_0 \setminus \mathcal{J}_B$  такой, что коэффициенты  $\alpha_j$  его разложения по базису, определяемые однозначно из уравнений

$$v_p = \sum_{j \in \mathcal{J}_B} \alpha_j v_j, \quad (1.8)$$

удовлетворяют неравенствам

$$\Delta_p - \sum_{j \in \mathcal{J}_B} \Delta_j \alpha_j < 0; \quad \alpha_j \leq 0, \quad j \in \mathcal{J}_B. \quad (1.9)$$

Рассмотрим множество

$$\mathcal{S} = \{j : j \in \mathcal{J}_B, \alpha_j < 0\} \cup \{p\},$$

составленное из индексов ненулевых компонент  $\alpha_j$  и индекса  $p$ . Примем

$$\lambda_i = -t\alpha_i, \quad i \in \mathcal{S}, \quad i \neq p, \quad \lambda_p = t, \quad (1.10)$$

где  $t$  — произвольное положительное число. Тогда соотношения (1.8) и (1.9) записываются в виде (1.7). Теорема доказана.

Таким образом, при неоптимальности текущего допустимого базисного решения задачи (1.1) множество  $\mathcal{S}$ , указанное в теореме 3.2, находится в процессе решения задачи (1.6). Покажем, как уменьшить при этом целевую функцию задачи (1.1). В соответствии с равенством в (1.7) и разложением (1.2) имеем

$$\sum_{j \in \mathcal{S}} \lambda_j v_j = 0 \Leftrightarrow \sum_{j \in \mathcal{S}} \lambda_j a_j \in \mathcal{L}(U) \Leftrightarrow \sum_{j \in \mathcal{S}} \lambda_j a_j = \sum_{i \in \mathcal{I}_+} \mu_i a_i, \quad (1.11)$$

где коэффициенты  $\mu_i$  определяются однозначно, так как система уравнений в (1.11) относительно  $\mu_i$  представляет собой  $k$  линейно независимых уравнений с  $k$  неизвестными.

**ЗАМЕЧАНИЕ 3.2.** Линейная комбинация  $\sum_{j \in \mathcal{S}} \lambda_j a_j$  является аналогом вектора, вводимого в базис в обычном симплекс-методе.

Введем множество индексов

$$\mathcal{J} = \{j : j \in \mathcal{I}_+, \mu_j > 0\} \quad (1.12)$$

и, если оно не пусто, положим

$$\theta = \min_j \left\{ \frac{x_j}{\mu_j} : j \in \mathcal{J} \right\} > 0. \quad (1.13)$$

Множество индексов, на котором достигается минимум в (1.13), обозначим

$$\mathcal{R} = \left\{ r : \frac{x_r}{\mu_r} = \theta \right\}.$$

**ТЕОРЕМА 3.3 (МЕТОД УМЕНЬШЕНИЯ ЦЕЛЕВОЙ ФУНКЦИИ).** Пусть текущее допустимое базисное решение  $x$  неоптимально. Тогда

1) если  $\mathcal{J} \neq \emptyset$ , то множество индексов

$$\tilde{\mathcal{I}}_+ = \mathcal{I}_+ \cup \mathcal{S} \setminus \mathcal{R} \quad (1.14)$$



соответствует строгому базису задачи (1.1), отвечающему меньшему значению целевой функции

$$\sum_{i \in \tilde{\mathcal{I}}_+} c_i \tilde{x}_i = \sum_{i \in \mathcal{I}_+} c_i x_i + \theta \sum_{j \in \mathcal{S}} \Delta_j \lambda_j, \quad (1.15)$$

где числа  $\lambda_j$ ,  $j \in \mathcal{S}$  определены по формуле (1.10), а новые переменные, соответствующие строгому базису, определяемому множеством  $\tilde{\mathcal{I}}_+$ , вычисляются по формулам

$$\tilde{x}_i = \begin{cases} x_i - \theta \mu_i, & i \in \mathcal{I}_+ \setminus \mathcal{R}, \\ \lambda_i \theta, & i \in \mathcal{S}. \end{cases} \quad (1.16)$$

2) если  $\mathcal{J} = \emptyset$ , то значение задачи (1.1) равно  $-\infty$ .

Доказательство теоремы 3.3 приведено в Приложении.

**ЗАМЕЧАНИЕ 3.3.** Как следует из (1.11) и (1.13), величина  $\theta \lambda_i$  при любых  $i$  не зависит от множителя  $t$  в (1.10). Это означает, что его выбор не влияет и на значение целевой функции в (1.15).

**ЗАМЕЧАНИЕ 3.4.** Число столбцов в строгом базисе, определяемом множеством  $\tilde{\mathcal{I}}_+$ , согласно (1.14), равно

$$\tilde{k} = |\tilde{\mathcal{I}}_+| = k + |\mathcal{S}| - |\mathcal{R}|.$$

При этом, согласно (1.7),  $|\mathcal{S}| \geq 2$ , т.е. в базис всегда вводится не менее двух векторов. Кроме того,  $|\mathcal{R}| \leq k$ ,  $|\mathcal{S}| \geq 2$ . Поэтому  $\tilde{k} \geq 2$ , т.е. степень вырождения  $(m-1)$  после решения вспомогательной задачи не может реализоваться. Рассмотрим два часто встречающихся частных случая.

а) Пусть множество  $\mathcal{R}$  содержит только один индекс, т.е. из базиса выводится один вектор. Тогда

$$\tilde{k} \geq k + |\mathcal{S}| - 1 \geq k + 1,$$

т.е. при  $|\mathcal{R}| = 1$  степень вырождения уменьшается.

б) Пусть выполняется условие "а" и кроме того  $|\mathcal{S}| = m - k + 1$ , (т.е. в разложении (1.8) нет нулевых коэффициентов). Тогда  $\tilde{k} = m$  и текущий базис становится невырожденным.

### 3.1.3. Описание алгоритма

Алгоритм метода, обоснованного теоремами 3.1 – 3.3, зависит от выбора матрицы  $V$ , дополняющей матрицу строгого базиса  $U$  до базисной матрицы  $B$ . Приведем этот алгоритм сначала для случая, когда  $B$  найдена в результате обычных симплексных итераций, приведших к текущему вырожденному решению. Другой способ выбора базисной матрицы описан ниже в замечании 3.6.

Пусть  $g_i \in \mathbb{R}^k$ ,  $h_i \in \mathbb{R}^{m-k}$  — векторы коэффициентов разложения вектора  $a_i$ , не входящего в строгий базис  $U$ , соответственно по столбцам матриц  $U$  и  $V$ , т.е.  $u_i = Ug_i$ ,  $v_i = Vh_i$ , или

$$\begin{pmatrix} g_i \\ h_i \end{pmatrix} = B^{-1}a_i, \quad i \in \mathcal{I}_0. \quad (1.17)$$

Тогда из соотношений (1.11) следует

$$\mu = \sum_{j \in \mathcal{S}} \lambda_j g_j, \quad (1.18)$$

а задача (1.6) может быть записана в эквивалентной форме

$$\min_{y_i} \left\{ \sum_{i \in \mathcal{I}_0} \Delta_i y_i : \sum_{i \in \mathcal{I}_0} y_i h_i = f, \quad y_i \geq 0 \quad \forall i \in \mathcal{I}_0 \right\}, \quad (1.19)$$

где  $f = B^{-1}d$ . Задача (1.19) содержит  $m - k$  ограничений, причем столбцы  $h_i$ , соответствующие столбцам матрицы  $V$ , образуют, очевидно, единичную матрицу. Это позволяет сразу указать начальное допустимое базисное решение задачи (1.19), а в качестве компонент вектора  $f$  выбрать произвольные положительные числа. В соответствии с изложенным выше алгоритм борьбы с вырожденностью может быть записан в следующем виде.

1. Из условия (1.17) определяются векторы  $g_i$  и  $h_i$ ,  $i \in \mathcal{I}_0$ .
2. Строится вектор  $f$ , компоненты которого — положительные случайные числа.
3. Решается задача (1.19), в которой начальный базис — единичная матрица.

4. Если задача (1.19) имеет конечное значение, то текущий базис  $B$  задачи (1.1) оптимален. Иначе — переход на шаг 5.

5. Из условия (1.10) при  $t = 1$  определяются коэффициенты  $\lambda_j$ ,  $j \in \mathcal{S}$ .

6. По формуле (1.18) находится вектор  $\mu$ .

7. По формуле (1.13) определяется параметр  $\theta$ .

8. В соответствии с условием теоремы 3.3 строится множество  $\tilde{\mathcal{I}}_+$  индексов, соответствующих новому строгому базису задачи (1.1).

**ЗАМЕЧАНИЕ 3.5.** При случайном выборе компонент вектора  $f$  задача (1.19) с вероятностью 1 не является вырожденной. В случае появления вырождения достаточно лишь заново задать компоненты вектора  $f$  в соответствии с шагом 2 алгоритма и пересчитать значения базисных переменных, после чего продолжить решение.

**ЗАМЕЧАНИЕ 3.6.** Возможны и другие способы выбора базисной матрицы  $B$ . Например, можно принять

$$U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}, \quad B = \begin{pmatrix} U_1 & O \\ U_2 & I_{m-k} \end{pmatrix},$$

где  $U_1$  — невырожденная  $k \times k$  — матрица,  $I_{m-k}$  — единичная матрица порядка  $m - k$ . Тогда

$$B^{-1} = \begin{pmatrix} U_1^{-1} & O \\ -U_2 U_1^{-1} & I_{m-k} \end{pmatrix}, \quad g_i = U_1^{-1} u_i, \quad h_i = v_i - U_2 U_1^{-1} u_i, \quad i \in \mathcal{I}_0,$$

т.е. в данном случае достаточно обращать матрицу размерности  $k \times k$ . Поэтому такой выбор базисной матрицы  $B$  может оказаться полезным при  $k \ll m$ .

### 3.1.4. Эквивалентный критерий оптимальности и до- полнения к алгоритму

Теорема 3.3 является обобщением обычного симплекс-метода. При этом роль относительной оценки  $\Delta_i < 0$  играет величина (см. теорему 3.3)

$$D(\lambda) = \sum_{i \in \mathcal{S}} \Delta_i \lambda_i,$$

которую естественно называть *обобщенной относительной оценкой*. Рассмотрим минимальную из величин  $D(\lambda)$  при условии нормировки на вектор  $\lambda$ , заданный в (1.10) с точностью до множителя:

$$D^* = \min_S \left\{ D(\lambda) : \sum_{i \in S} \lambda_i v_i = 0, \quad \sum_{i \in S} \lambda_i = 1, \quad \lambda_i \geq 0, \quad i \in S \right\}.$$

Последняя задача оптимизации фактически является задачей линейного программирования, которая в обычном виде имеет вид

$$D^* = \min_{\lambda} \left\{ \sum_{i \in \mathcal{I}_0} \Delta_i \lambda_i : \sum_{i \in \mathcal{I}_0} \lambda_i v_i = 0, \quad \sum_{i \in \mathcal{I}_0} \lambda_i = 1, \quad \lambda_i \geq 0, \quad i \in \mathcal{I}_0 \right\}. \quad (1.20)$$

Величина  $D^*$  является аналогом минимальной относительной оценки  $\Delta_{\min} = \min_i \Delta_i$  в обычном симплекс-методе и для невырожденного допустимого базисного решения, очевидно, совпадает с ней. Приведем необходимый для дальнейшего другой критерий оптимальности задачи (1.1).

**ТЕОРЕМА 3.4.** 1. Текущее допустимое базисное решение оптимально тогда и только тогда, когда  $D^* \geq 0$  или условия в (1.20) несовместны. В последнем случае текущее допустимое базисное решение является единственным допустимым решением задачи (1.1). 2. Если  $D^* < 0$ , то множество  $S$ , введенное в теореме 3.2, находится из решения задачи (1.20) и соответствует любому текущему базису этой задачи, для которого  $D(\lambda) < 0$ . При этом целевая функция задачи (1.1) уменьшается по формуле (1.15). Доказательство легко проводится от противного с использованием теорем 3.2 и 3.3.

Непосредственно из теоремы (3.2) вытекает следующая лемма.

**ЛЕММА 3.1** (СВЯЗЬ МЕЖДУ ЗАДАЧАМИ (1.20) И (1.6)). Если задача (1.6) не имеет конечного значения, то множество  $\mathcal{J}_B \cup \{p\}$  определяет допустимый базис задачи (1.6), соответствующий вектору  $\lambda$ , ненулевые компоненты которого определяются из соотношения (1.10) при  $t = (1 - \sum_{i \in \mathcal{J}_B} \alpha_i)^{-1}$ . При этом  $D(\lambda) < 0$ .

На практике при решении задачи (1.1) обычным симплекс-методом достаточное условие оптимальности (1.5) заменяется условием  $\varepsilon$ -оптимальности

$$\Delta_{\min} \geq -\varepsilon,$$

где  $\varepsilon$  — заданное положительное число. Такая замена может быть обоснована следующей леммой.

**ЛЕММА 3.2.** *Если оптимальное решение задачи (1.1) удовлетворяет условию*

$$\sum_{i=1}^n x_i^* \leq M, \quad (1.21)$$

где  $M > 0$  — заданное число, то для любого неоптимального допустимого базисного решения  $x$  справедливо неравенство

$$c^T x^* \geq c^T x + M\Delta_{\min}. \quad (1.22)$$

Доказательство леммы 3.2 приведено в Приложении. Аналогично доказывается, что в вырожденном случае имеет место более точная оценка

$$c^T x^* \geq c^T x + MD^*. \quad (1.23)$$

Поэтому вместо эквивалентного критерия оптимальности  $D^* \geq 0$  на практике может быть использован критерий  $\varepsilon$ -оптимальности

$$D^* \geq -\varepsilon. \quad (1.24)$$

Однако непосредственная проверка этого критерия может быть затруднена в силу вырожденности задачи (1.20). Поэтому на практике целесообразно использовать следующий результат.

**ЛЕММА 3.3.** *Значение  $D^*$  задачи (1.20) удовлетворяет неравенству*

$$D^* \geq D(\lambda) + \hat{\Delta}_{\min},$$

где  $\hat{\Delta}_{\min}$  — минимальная относительная оценка для текущего базиса в задаче (1.20).

Доказательство леммы 3.3 следует из формулы (1.22), примененной к задаче (1.20) при  $M = 1$ . Другое доказательство этого утверждения приведено в [46].

С использованием приведенных лемм шаг 5 описанного выше алгоритма может быть модифицирован следующим образом.

5.1. Согласно лемме 3.1 строится допустимое базисное решение  $\lambda$  задачи (1.20), компоненты которого вычисляются по формуле

$$\lambda_i = \frac{1}{1 - \sum_{i \in \mathcal{J}_B} \alpha_i} \begin{cases} -\alpha_i, & i \in \mathcal{J}_B, \\ 1, & i = p, \\ 0 & \text{иначе.} \end{cases}$$

5.2. Ищется величина  $D(\lambda)$  (согласно лемме 3.1,  $D(\lambda) < 0$ ).

5.3. Для текущего базиса задачи (1.20), определяемого множеством индексов  $\mathcal{J}_B \cup \{p\}$ , ищется минимальная относительная оценка  $\hat{\Delta}_{\min}$ .

5.4. Если выполняется условие  $D(\lambda) + \hat{\Delta}_{\min} \geq -\varepsilon$ , то согласно лемме 3.3 текущий базис задачи (1.1) является  $\varepsilon$ -оптимальным, и вычисления завершаются. В противном случае выполняется один из двух следующих шагов.

5.5. Если  $\hat{\Delta}_{\min} \geq 0$ , то решение задачи (1.20), построенное на шаге 5.1, является оптимальным. В этом случае числа  $\lambda_i$  определяются из соотношения (1.10) при  $t = (1 - \sum_{i \in \mathcal{J}_B} \alpha_i)^{-1}$ . Переход к шагу 6.

5.6. Если  $\hat{\Delta}_{\min} < 0$ , то решение, построенное на шаге 5.1, неоптимально. В этом случае можно попытаться уменьшить найденное значение  $D(\lambda)$ , решая задачу (1.20) начиная с базиса, определяемого множеством индексов  $\mathcal{J}_B \cup \{p\}$ . Критерий окончания — выполнение достаточных условий оптимальности для задачи (1.20) или появление при ее решении вырожденных итераций. В результате определяется вектор  $\tilde{\lambda}$ , для которого

$$D(\tilde{\lambda}) \leq D(\lambda) < 0.$$

Тогда в качестве чисел  $\lambda_i$ ,  $i \in \mathcal{S}$ , выбираются ненулевые компоненты вектора  $\tilde{\lambda}$ .

**ЗАМЕЧАНИЕ 3.7** (ОЦЕНИВАНИЕ ЦЕЛЕВОЙ ФУНКЦИИ С ПОЛОЖИТЕЛЬНЫМИ КОЭФФИЦИЕНТАМИ). На практике неравенство (1.24) часто является лишь косвенным признаком того, что решение близко к оптимальному, так как не всегда известно, что справедливо условие (1.21). Однако в случае  $c_i > 0$ ,  $1 \leq i \leq n$  имеет место непосредственная оценка значения задачи (1.1):

$$c^T x \geq c^T x^* \geq \frac{c^T x}{1 - D_c^*}, \quad (1.25)$$

где  $x$  — текущее допустимое базисное решение,

$$D_c^* = \min_{\lambda_i} \left\{ \sum_{i \in \mathcal{I}_0} \Delta_i \lambda_i : \sum_{i \in \mathcal{I}_0} \lambda_i h_i = 0, \quad \sum_{i \in \mathcal{I}_0} c_i \lambda_i = 1, \quad \lambda_i \geq 0, \quad i \in \mathcal{I}_0 \right\}.$$

Оценка (1.25) следует из неравенства

$$c^T x^* \geq c^T x + c^T x^* D_c^*,$$

аналогичного неравенству (1.23). В невырожденном случае оценка (1.25) переходит в оценку

$$c^T x^* \geq \frac{c^T x}{1 - \min_{1 \leq i \leq n} \frac{\Delta_i}{c_i}},$$

полученную ранее в [44] другим способом.

### 3.1.5. Практические результаты

Рассмотрим использование предложенного алгоритма на примере решения описанной в [100] задачи составления расписания авиаперевозок. Пусть  $m$  — число заказов на авиаперевозки или *заданий*,  $n$  — число возможных способов доставки грузов, которые условимся называть *режимами* работы. Будем предполагать, что в каждом режиме выполняется один или несколько заказов по доставке. Обозначим через  $c_j$  суммарные затраты на проведение  $j$ -го режима. Введем матрицу  $A$  с элементами

$$a_{ij} = \begin{cases} 1, & \text{если в } j\text{-м режиме работы выполняется } i\text{-е задание,} \\ 0 & \text{в противном случае.} \end{cases}$$

Обозначим также через  $x_j$  число выполнения заказов при проведении работы  $j$ -го режима. Тогда задача минимизации расходов при выполнении заказов будет иметь вид

$$\min_x \left\{ c^T x : \sum_{i=1}^n x_i a_i = e, \quad x_i \in \mathbb{N}, \quad 1 \leq i \leq n \right\}. \quad (1.26)$$

где  $e = (1, \dots, 1)^T \in \mathbb{R}^m$ , и первое условие означает, что каждое задание будет выполнено ровно один раз.

Задача (1.26) представляет собой задачу целочисленного программирования и может быть решена при помощи соответствующего метода. При этом на предварительном этапе требуется решить задачу линейного программирования, полученную из задачи (1.26) заменой условия  $x_i \in \mathbb{N}$  на более слабое  $x_i \geq 0$  при  $1 \leq i \leq n$ . При решении такой задачи с использованием стандартного симплекс-метода часто наблюдаются большие последовательности вырожденных итераций. Приведем результаты применения к задаче (1.26) описанного выше алгоритма, а также алгоритма, описанного в [100].

При численных расчетах решалась задача для  $n = 3135$  и  $m = 80$ . Коэффициенты матрицы  $A$  выбирались случайно из множества  $\{0, 1\}$ . Коэффициенты  $c_j$  также задавались случайно в интервале от 1000 до 20000. Результаты расчетов приведены в таблицах. Использование стандартного алгоритма приводит к последовательностям вырожденных итераций, число которых достигает 50. Целевая функция при этом практически остается постоянной. Всего было сделано 1324 итерации, и затем были выполнены достаточные условия оптимальности. Использование нового алгоритма и алгоритма Вулфа позволило сократить число итераций при решении более чем вдвое. При использовании алгоритма Вулфа на каждой вырожденной итерации решалась вспомогательная задача той же размерности, что и исходная, при использовании нового алгоритма размерность вспомогательной задачи зависела от числа ненулевых базисных переменных. Это позволило сократить время расчетов по сравнению с алгоритмом Вулфа примерно на 20%. Было проведено большое число экспериментов, при которых решались задачи различной размерности с различными условиями. В большом количестве



случаев новый алгоритм оказывался более эффективным и позволял сократить время вычислений порой вдвое, особенно при малых  $m - k$ . Однако в ряде случаях (обычно при малых  $k$ ) предпочтительнее оказывался алгоритм Вулфа. Это не позволяет сделать однозначный вывод о преимуществе нового метода борьбы с вырождением перед методом Вулфа, но говорит о его высокой эффективности и целесообразности проведения дальнейших исследований в этой области.

## ПРИЛОЖЕНИЕ

*Доказательство теоремы 3.1.* Критерием Куна—Таккера оптимальности текущего допустимого базисного решения является существование такого вектора  $\pi \in \mathbb{R}^m$ , что выполняются условия (1.3) и

$$\pi^T a_i \leq c_i, \quad i \in \mathcal{I}_0. \quad (1.27)$$

Из разложения (1.2) вектора  $a_i$  получаем  $\pi^T a_i = \pi_u^T u_i + \pi_v^T v_i$ , где  $\pi_u$  и  $\pi_v$  — ортогональные проекции вектора  $\pi$  на  $\mathcal{L}(U)$  и  $\mathcal{L}(V)$  соответственно. Тогда соотношение (1.27) переписывается в виде

$$\pi_v^T v_i \leq c_i - \pi_u^T u_i, \quad i \in \mathcal{I}_0, \quad (1.28)$$

а  $\pi_u$  однозначно определяется из условия

$$\pi_u^T U = c_u^T, \quad (1.29)$$

которое эквивалентно условию (1.3). Согласно теории двойственности [31, 68], неравенства (1.28) совместны относительно  $\pi_v$  тогда и только тогда, когда конечно значение задачи

$$\min_{y_i} \left\{ \sum_{i \in \mathcal{I}_0} (c_i - \pi_u^T u_i) y_i : \sum_{i \in \mathcal{I}_0} y_i v_i = d, \quad y_i \geq 0, \quad i \in \mathcal{I}_0 \right\}. \quad (1.30)$$

Целевая функция этой задачи с учетом (1.2) равна

$$\sum_{i \in \mathcal{I}_0} (\Delta_i y_i + \pi_0^T v_i y_i) = \sum_{i \in \mathcal{I}_0} \Delta_i y_i + \pi_0^T d,$$

откуда и следует утверждение теоремы.

*Замечание к доказательству.* Фактически нами доказано, что в формулировке теоремы 3.1 вместо задачи (1.6) может быть рассмотрена задача (1.30), отличающаяся от нее только тем, что в целевой функции коэффициенты  $\Delta_i(\pi_0)$  заменены на  $(c_i - \pi_u^T u_i)$ . Это означает, что мы можем не искать некоторого решения  $\pi_0$  системы (1.3), а найти однозначно вектор  $\pi_u$  из системы (1.29) по формуле  $\pi_u^T = c_u^T (U^T U)^{-1} U^T$ .

*Доказательство теоремы 3.3.* Возьмем вектор  $\tilde{x}$ , компоненты которого  $\tilde{x}_i$  при  $i \in \tilde{\mathcal{I}}_+$  вычисляются по формуле (1.16), а в противном случае равны нулю. Покажем, что  $\tilde{x}$  является допустимым решением задачи (1.1). Из (1.13) и условия  $\lambda_i > 0$  (см. (1.7)) следует, что  $\tilde{x}_i > 0$  для  $i \in \tilde{\mathcal{I}}_+$ . Кроме того,

$$\begin{aligned} \sum_{i \in \tilde{\mathcal{I}}_+} \tilde{x}_i a_i &= \sum_{i \in \tilde{\mathcal{I}}_+ \setminus \mathcal{R}} x_i a_i + \theta \left( \sum_{i \in \mathcal{S}} \lambda_i a_i - \sum_{i \in \tilde{\mathcal{I}}_+ \setminus \mathcal{R}} \mu_i a_i \right) = \\ &= \sum_{i \in \tilde{\mathcal{I}}_+} x_i a_i + \theta \left( \sum_{i \in \mathcal{S}} \lambda_i a_i - \sum_{i \in \tilde{\mathcal{I}}_+} \mu_i a_i \right) + \left( \sum_{i \in \mathcal{R}} \mu_i a_i - \theta \sum_{i \in \mathcal{R}} x_i a_i \right) = b, \end{aligned}$$

так как оба слагаемых в круглых скобках равны нулю в силу (1.11) и (1.13). Допустимость вектора  $\tilde{x}$  доказана. Далее, для любого допустимого решения  $\tilde{x}$  справедливо соотношение [59]

$$\sum_{i=1}^n c_i \tilde{x}_i = \sum_{i \in \mathcal{I}_+} c_i x_i + \sum_{i \in \mathcal{I}_0} \Delta_i \tilde{x}_i. \quad (1.31)$$

Подставляя в него формулу (1.16) и учитывая, что  $\Delta_i = 0$  при  $i \in \mathcal{I}_+ \setminus \mathcal{R}$ , получаем формулу (1.15). Эта формула справедлива при любом  $\theta$ , для которого вектор  $\tilde{x}$  имеет неотрицательные компоненты. Поэтому если множество (1.12) пусто, то вектор  $\tilde{x}$  допустим при любых  $\theta$ , и целевая функция не ограничена на множестве допустимых решений. Для доказательства теоремы осталось показать, что при  $\mathcal{J} \neq \emptyset$  и выборе  $\theta$  из формулы (1.13) столбцы  $a_i$ ,  $i \in \tilde{\mathcal{I}}_+$  линейно независимы, т.е. образуют строгий базис. Для этого составим нетривиальную линейную комбинацию этих векторов и приравняем ее к нулю:

$$\sum_{i \in \tilde{\mathcal{I}}_+ \setminus \mathcal{R}} y_i a_i + \sum_{i \in \mathcal{S}} y_i a_i = 0. \quad (1.32)$$

Отметим, что среди чисел  $y_i$ ,  $i \in \mathcal{S}$  есть ненулевые, так как в противном случае векторы  $a_i$ ,  $i \in \tilde{\mathcal{I}}_+ \setminus \mathcal{R}$  оказываются линейно зависимыми, что противоречит линейной независимости векторов  $a_i$ ,  $i \in \mathcal{I}_+$ . Из (1.32) и (1.2) следует

$$\sum_{i \in \mathcal{S}} y_i a_i \in \mathcal{L}(U) \Leftrightarrow \sum_{i \in \mathcal{S}} y_i v_i = 0.$$

Сравнивая это уравнение с уравнением в (1.7), мы можем принять, что  $y_i = \lambda_i$ ,  $i \in \mathcal{S}$  (так как обе эти величины определены с точностью до постоянной). Тогда из (1.11) и (1.32) следует, что

$$\sum_{i \in \mathcal{I}_+ \setminus \mathcal{R}} (y_i + \mu_i) a_i + \sum_{i \in \mathcal{R}} \mu_i a_i = 0,$$

откуда  $\mu_i = 0$ ,  $i \in \mathcal{R}$ .

*Доказательство леммы 3.2.* Пусть  $x$  — некоторое текущее неоптимальное допустимое базисное решение задачи (1.1),  $\tilde{x}$  — произвольное решение, отличное от  $x$ . Учитывая, что в этом случае  $\Delta_{\min} < 0$ , из соотношения (1.31) получим

$$c^T \tilde{x} = c^T x + \sum_{i \in \mathcal{I}_0} \Delta_i \tilde{x}_i \geq c^T x + \Delta_{\min} \sum_{i \in \mathcal{I}_0} \tilde{x}_i \geq c^T x + M \Delta_{\min}.$$

Из произвольности выбора допустимого вектора  $\tilde{x}$  и следует утверждение леммы.

## Использование стандартного алгоритма.

Номер итерации	Размер строгого базиса	Значение целевой функции	Время расчетов, с
1	80	2965,5	
624-631	69	1677,7	3,61
674-735	41	1674,5	29,48
898-927	52	1673,5	30,16
1043-1069	36	1673,1	13,15
1130-1153	32	1672,2	11,57
1243-1279	27	1671,6	18,29
1323	25	1670,7	0,59

*Примечание.* Выполнено итераций: всего – 1324, вырожденных – 607.

Общее время расчетов – 13 минут 36,96 секунд.

## Использование алгоритма Вульфа.

Номер итерации	Размер строгого базиса	Размерность подзадачи	Число итераций в подзадаче	Значение целевой функции	Время расчетов, с
1	80			2868,3	
554	77	$80 \times 3135$	1	1681,2	0,38
560	65	$80 \times 3135$	13	1672,3	5,49
564	54	$80 \times 3135$	21	1671,3	8,51
566	41	$80 \times 3135$	36	1670,8	17,08
568	26	$80 \times 3135$	170	1670,3	72,89
569	23	$80 \times 3135$	150	1670,2	97,49

*Примечание.* Выполнено итераций: всего – 569, вырожденных – 36.

Общее время расчетов – 9 минут 12,61 секунд.

Использование нового алгоритма.

Номер итерации	Размер строгого базиса	Размерность подзадачи	Число итераций в подзадаче	Значение целевой функции	Время расчетов, с
1	80			2887,3	
554	78	$2 \times 3057$	2	1675,1	1,10
560	65	$15 \times 3070$	7	1672,1	5,22
564	41	$39 \times 3094$	82	1671,1	24,34
565	32	$48 \times 3103$	240	1670,8	59,59
566	28	$52 \times 3107$	100	1670,6	34,88

*Примечание.* Выполнено итераций: всего – 566, вырожденных – 22.

Общее время расчетов – 7 минут 16,87 секунд.

## 3.2. Обобщенная задача линейного программирования

### 3.2.1. Виды обобщенных задач и соотношения между ними

Рассмотрим теперь обобщенную задачу линейного программирования, в которой сами векторы условий могут принадлежать некоторым ограниченным замкнутым множествам  $\mathcal{A}_i$ , а каждый коэффициент в целевой функции может зависеть от вектора  $a_i$ .

Таким образом, векторы  $a_i$  являются параметрами оптимизации:

$$W = \min_{x_i, a_i} \left( \sum_{i=1}^n c_i(a_i)x_i : \sum_{i=1}^n x_i a_i = b, \quad x \geq 0, \quad a_i \in \mathcal{A}_i \quad \forall i = 1, \dots, n \right). \quad (2.1)$$

Мы будем рассматривать только случай линейной зависимости коэффициентов от

вектора  $a$  :

$$c_i(a) = c_{i1} + c_{i2}^T a.$$

В [31] задача (2.1) рассматривается лишь для случая выпуклых множеств  $\mathcal{A}_i$  при  $c_{i2} = 0 \forall i$  и называется обобщенной задачей Вулфа. Для задач коррекции, рассмотренных в главах 2, 6, каждое из этих множеств — граница выпуклого множества (например,  $\mathcal{A}_i$  — эллипсоиды) и, кроме того,  $c_{i2} = 0$ ,  $c_{i1} > 0$ . При этом [31, 44] естественно искать оптимальный базис из достаточного условия оптимальности, которое в отличие от (1.5) можно записать в виде:

$$\Delta_{\min} \doteq \min_{i=1, \dots, n} \Delta_i = 0, \quad (2.2)$$

где вычисление целевой функции в отличие от (1.4) представляет собой уже некоторую оптимальную подзадачу:

$$\Delta_i = \min_{a \in \mathcal{A}_i} (c_i(a) - \pi^T a). \quad (2.3)$$

**ЗАМЕЧАНИЕ 3.8.** В отличие от условия (1.5) выражение (2.2) представляет собой равенство, так как точки базиса входят в множество оптимизации и для базисных точек  $\Delta_i = 0$ .

Если  $\Delta_{\min} < 0$ , то вектор, реализующий решение подзадачи (2.2), вводится в базис и целевая функция не увеличивается, если допустимое решение — не вырождено. Этот вектор выбирается из бесконечного множества, в связи с чем предлагаемая модификация симплекс-метода носит название метода генерации столбцов. Если решение подзадачи (2.3) находится достаточно просто, например, аналитически, то метод генерации столбцов так же эффективен, как и соответствующая задача обычного линейного программирования таких же размеров.

Для случая, когда множества  $\mathcal{A}_i$  — выпуклые многогранники, базис, удовлетворяющий условию (2.3) всегда существует [31]. Однако в общем случае для задачи (2.1) не всегда можно найти базис, отвечающий условию (2.3) [6]. Поэтому (1.5) приходится заменять на  $\varepsilon$ -оптимальное условие [44, 18]

$$\Delta > -\varepsilon, \quad (2.4)$$

которому удовлетворяет некоторое множество базисов при любом достаточно малом  $\varepsilon > 0$  [18]. На практике это может привести к плохой обусловленности матриц  $\varepsilon$ -оптимального базиса при малых  $\varepsilon$  [45, 6], а также, как и в (1.1), — к многократным итерациям без уменьшения целевой функции. При этом для (2.1) неизвестны методы предотвращения заикливания. Задачу (2.1) можно рассматривать как нахождение минимума среди всех решений множества обычных задач линейного программирования вида (1.1), каждая из которых соответствует определенному выбору векторов  $a_i$  из множеств  $\mathcal{A}_i$ . Для решения задачи (2.1) рассмотрим аналогичную задачу, но с возможностью выбора нескольких векторов  $a_i$  из этих множеств, так что соответствующие векторам  $a_i$  переменные  $x_i(a_i)$  входят в целевую функцию с коэффициентом  $c_i(a_i)$ . При этом указанную задачу можно строго записать следующим образом [9]. Нужно найти конечные множества точек  $\mathcal{F}_i \subset \mathcal{A}_i$  и переменные  $x_i(a_i) \quad \forall a \in \mathcal{F}_i$  из условия

$$w = \inf_{\mathcal{F}_i, x_i(a)} \left( \sum_{i=1}^n \sum_{a \in \mathcal{F}_i} c_i(a) x_i(a) : \sum_{i=1}^n \sum_{a \in \mathcal{F}_i} x_i(a) a = b, \right. \\ \left. x_i(a) \geq 0, \quad \mathcal{F}_i \subset \mathcal{A}_i \quad \forall i = 1, \dots, n \right). \quad (2.5)$$

**ЗАМЕЧАНИЕ 3.9.** Задачу (2.5) можно записать в эквивалентном и для многих более привычном виде, где параметрами оптимизации служат числа  $n_i$  и векторы  $a_{ij} \in \mathcal{A}_i$  (а не множества  $\mathcal{F}_i$ ) [18]:

$$w = \inf_{n_i, a_i, x_{ij}(a_i)} \left( \sum_{i=1}^n \sum_{j=1}^{n_i} c_i(a_{ij}) x_{ij}(a_{ij}) : \sum_{i=1}^n \sum_{j=1}^{n_i} x_{ij} a_{ij} = b, \right. \\ \left. x_{ij} \geq 0, \quad a_{ij} \in \mathcal{A}_i \quad \forall i = 1, \dots, n \right). \quad (2.6)$$

Мы будем использовать также еще более простую запись задачи (2.5), которая употребляется в [57]:

$$w = \inf_{a_i, x_{ij} a_{ij}} \left( \sum_{i=1}^n \sum_j c_i(a_{ij}) x_{ij}(a_{ij}) : \sum_{i=1}^n \sum_j x_{ij} a_{ij} = b, \right. \\ \left. x_{ij} \geq 0, \quad a_{ij} \in \mathcal{A}_i \quad \forall i = 1, \dots, n \right). \quad (2.7)$$

Задача (2.5) имеет самостоятельный интерес. Например, к задаче такого вида сводится оптимальная задача выбора состава астроизмерений [23, 18]. Назовем ее также обобщенной задачей линейного программирования и ниже дадим



теорию и алгоритм решений этой задачи. Совокупность переменных оптимизации в (2.5)

$$\{\mathcal{F}_i, \quad x_i(a) \quad \forall a \in \mathcal{F}_i, \quad i = 1, \dots, n\}, \quad (2.8)$$

удовлетворяющих ограничениям этой задачи, будем называть допустимым решением; допустимое решение, дающее оптимум, — оптимальным решением. Аналогичные определения используем и для (2.1). Допустимое решение последней является, очевидно, допустимым и для задачи (2.5) (случай  $\mathcal{F}_i = a_i$  в (2.5)). Поэтому  $w \leq W$ . Введем понятия, аналогичные используемым в линейном программировании. Будем называть допустимое решение (2.8) базисным, если векторы

$$a_{i_1}, \dots, a_{i_k} \quad (i_j \in \{1, \dots, n\} \forall j = 1, \dots, k), \quad (2.9)$$

соответствующие ненулевым компонентам  $x_j(a_j)$ ,  $j = i_1, \dots, i_k$ , линейно независимы. Очевидно  $k \leq m$ . Векторы (2.9) будем называть строгим базисом допустимого базисного решения. Если  $k < m$ , то допустимое базисное решение назовем вырожденным, если  $k = m$  — невырожденной. Рассмотрим задачу (2.5), фиксируя в ее условиях множества  $\mathcal{F}_i$  и полагая их равными множествам некоторой допустимого решения (2.8). Тогда эта задача становится обычной задачей линейного программирования и, следовательно, имеет оптимальное допустимое базисное решение  $\{x_1, \dots, x_m, 0, \dots, 0\}$ , где  $x_i \geq 0$ ,  $i = 1, \dots, m$ . Этому решению соответствует допустимое базисное решение задачи (2.5), дающая не большее значение целевой функции, чем исходная допустимая точка (2.8). Отсюда получаем следующий вывод.

**ТЕОРЕМА 3.5.** *Если задача (2.5) имеет решение, то существует оптимальное допустимое базисное решение задачи. При этом в (2.5) операция  $\inf$  заменяется операцией  $\min$ , а множества  $\mathcal{F}_i$  можно считать состоящими из  $m$  точек.*

Рассмотрим условия существования единственного решения у задач (2.1) и (2.5). Прежде всего отметим, что в обобщенной задаче Вулфа [31] множества в (2.1) — выпуклые и решение задачи (2.1) легко получается из решения задачи (2.5) путем замены в (2.8) множеств  $\mathcal{F}_i$  и соответствующих переменных  $x_i(a)$  их вы-

пуклой комбинацией, при этом  $w = W$ . Однако уже в упомянутой задаче выбора астроизмерений, где множества  $\mathcal{A}_i$  — эллипсы в трехмерном пространстве, не содержащие в своей плоскости начало координат, может реализоваться случай  $w < W$ , где  $W$  соответствует решению задачи (2.5) при условии, что каждое множество  $\mathcal{F}_i$  состоит из одной точки. В задаче линейной идеальной коррекции [44, 18] каждое множество  $\mathcal{A}_i$  есть поверхность выпуклой области, симметричной относительно начала координат в  $\mathbb{R}^m$ . В этих и других случаях множества  $\mathcal{A}_i$  обладают свойством, которое мы называем нуль-выпуклостью [18]. А именно, множество  $\mathcal{A}$  — нуль-выпуклое, если для любых точек  $a_1, a_2 \in \mathcal{A}$  и числа  $0 < \lambda < 1$  существует такое число  $\kappa \geq 1$ , что

$$\kappa[\lambda a_1 + (1 - \lambda)a_2] \in \mathcal{A}.$$

Если это условие верно только для  $\kappa > 1$ , то множество мы называем строго нуль-выпуклым. Имеет место следующий критерий нуль-выпуклости [18], аналогичный критерию выпуклости множества [67].

**ЛЕММА 3.4.** *Множество  $\mathcal{A}$  нуль-выпукло тогда и только тогда, когда для любых точек  $a_i \in \mathcal{A}$ ,  $i = 1, \dots, n$ , и чисел  $\lambda_i \geq 0$  таких, что  $\lambda_1 + \dots + \lambda_n = 1$ , существует число  $\kappa \geq 1$  (в случае строгой нуль-выпуклости  $\kappa > 1$ ), что  $\kappa(\lambda_1 a_1 + \dots + \lambda_n a_n) \in \mathcal{A}$ .*

Рассмотрим задачи (2.1) и (2.5), считая, что множества  $\mathcal{A}_i$  для них одни и те же и являются нуль-выпуклыми. Пусть (2.8) — допустимое решение задачи (2.5). Определим компоненты допустимого решения задачи (2.1) из условия

$$x_i^* = \frac{1}{\kappa_i} \sum_{a \in \mathcal{F}_i} x_i(a), \quad a_i^* = \kappa_i \sum_{a \in \mathcal{F}_i} \lambda_i(a) a, \quad \lambda_i(a) = \frac{x_i(a)}{\sum_{a \in \mathcal{F}_i} x_i(a)}, \quad (2.10)$$

где  $\kappa_i \geq 1$  выбрано из условия  $a_i^* \in \mathcal{A}_i$ , что возможно в силу леммы 3.4. Из (2.10)  $x_i^* \leq \sum_{a \in \mathcal{F}_i} x_i(a)$ , причем равенство имеет место только при  $\kappa_i = 1$ . Отсюда нетрудно получить следующий результат.

**ТЕОРЕМА 3.6.**

1. Если для всех  $i$  множества  $\mathcal{A}_i$  — строго нуль-выпуклые и  $c_i(a_i) > 0$ ,

то задачи (2.1) и (2.5) эквивалентны, т.е. решение одной из них является решением другой.

2. Если множества  $\mathcal{A}_i$  — нуль-выпуклые и  $c_i \geq 0$ , то решение задачи (2.1) получается из решения (2.5) преобразованием (2.10). При этом  $w = W$ .
3. Если множества  $\mathcal{A}_i$  — выпуклые, то также то решение задачи (2.1) получается из решения (2.5) преобразованием (2.10) и  $w = W$ .

### 3.2.2. Критерий оптимальности для обобщенной задачи линейного программирования

Рассмотрим вспомогательную задачу обобщенного линейного программирования размерности  $m - k$ :

$$\min_{\mathcal{F}_i, y_i(a)} \left\{ \sum_{i=1}^n \sum_{a \in \mathcal{F}_i} (c_i(a) - \pi_0^T a) y_i(a) : \sum_{i=1}^n \sum_{a \in \mathcal{F}_i} y_i(a) v = d, \mathcal{F}_i \in \mathcal{A}_i, y_i(a) \geq 0 \right\}, \quad (2.11)$$

где  $d \in \mathcal{L}(V)$  — любой вектор, для которого совместны ограничения в (2.11),  $\mathcal{F}_i$  — произвольные конечные множества, вектор  $v$  определяется из условия вида (1.2);  $\pi_0$  — произвольное решение системы (1.3), в которой  $B$  — матрица строгого базиса (2.9) допустимого базисного решения.

**ТЕОРЕМА 3.7 (АНАЛОГ ТЕОРЕМЫ 3.1).** *Допустимое базисное решение задачи (2.5) оптимально тогда и только, когда конечно значение задачи (2.11).*

*Доказательство.* Необходимость. Пусть строгий базис (2.9) соответствует оптимальному базисному решению задачи (2.5). Рассмотрим соответствующую (2.5) обычную задачу линейного программирования, множество векторов условий которой состоит из векторов (2.9) и некоторых конечных множеств  $\mathcal{F}_i \in \mathcal{A}_i$ . Она записывается в виде

$$\tilde{w} = \min_{x_j \geq 0, x_i(a) \geq 0} \left( \sum_{j=1}^k c_{ij} x_j + \sum_{i=1}^n \sum_{a \in \mathcal{F}_i} c_i(a) x_i(a) : \sum_{j=1}^k x_j a_{ij} + \sum_{i=1}^n \sum_{a \in \mathcal{F}_i} x_i(a) a = b \right). \quad (2.12)$$

Очевидно  $\check{w} \geq w$ , и отсюда следует, что строгий базис (2.9) оптимален в (2.12), так как ему соответствует значение целевой функции, равное  $w$ . Теперь в силу теоремы 3.1 конечно решение задачи вида (2.11), соответствующей задаче (2.12). Ввиду произвольного выбора множеств  $\mathcal{F}_i$  получаем необходимые условия теоремы 3.7.

Достаточность. Пусть конечно значение задачи (2.11). Тогда конечно значение задачи вида (2.11) при любом выборе множеств  $\mathcal{F}_i$ . Поэтому для любого допустимого базисного решения (2.8) имеем

$$\sum_{j=1}^k c_{i_j} x_{i_j}(a_{i_j}) \leq \sum_{i=1}^n \sum_{a \in \mathcal{F}_i} c_i(a) x_i(a).$$

Достаточность и теорема доказаны.

**ЗАМЕЧАНИЕ 3.10.** Имеется существенное различие между теоремами 3.1 и ее аналога 3.7. Для обычной задачи линейного программирования (1.1) на каждой итерации совокупность векторов, вводимых в базис, может быть найдена в процессе решения задачи линейного программирования (1.6) размерности  $m-k$ , которая выбором вектора  $d$  условий может быть сделана невырожденной. Для аналога задачи (1.6), соответствующего задаче (2.5), этой невырожденности добиться, вообще говоря, не удастся ввиду континуальности множеств  $\mathcal{F}_i$ . Для (2.5) также имеют место аналоги теорем 3.2 — 3.4, которые формулируются очевидным образом. В остальном алгоритм решения задачи (2.5) после нахождения вводимых в строгий базис векторов (из решения задачи (2.11) или аналога задачи (1.6)) описан в теореме 3.3. На каждом шаге алгоритма целевая функция уменьшается, и за цикливания быть не может. Возникает вопрос о сходимости алгоритма генерации столбцов.

### 3.2.3. О сходимости алгоритма генерации столбцов.

На каждом шаге метода генерации столбцов целевая функция уменьшается на  $-\theta \Delta_{\min}$ . Отсюда получаем следующее утверждение.

**ЛЕММА 3.5.** Если  $\theta \Delta_{\min}$  не стремится к нулю с увеличением числа итераций, то алгоритм сходится за конечное число шагов к оптимальному решению.

На практике случай, указанный в лемме 3.5 обычно не имеет места. Рассмотрим другой случай.

**ЛЕММА 3.6.** Пусть задача (2.5) имеет конечное решение. Тогда, если стремиться к нулю  $\Delta_{\min}$ , то алгоритм сходится к оптимальному решению по функционалу.

Утверждение леммы (3.6) следует из неравенства 1.22:

$$c^T x^* \geq c^T x + M \Delta_{\min},$$

которое справедливо и для обобщенной задачи. Здесь постоянная  $M$  определяется из условия  $\sum_{i=1}^n x_i^* \leq M$ , а  $\Delta_{\min}$  находится по формуле (2.2)

Теоретически еще возможен случай, когда величина  $\theta$  стремится к нулю с бесконечным увеличением числа итераций, а  $\Delta_{\min}$  при этом к нулю не стремится. В этом случае алгоритм сходится к некоторому вырожденному допустимому базисному решению. Если мы возьмем это решение за исходное, то можно опять уменьшать целевую функцию на конечную величину при помощи теоремы 3.7. Таким образом, симплексный алгоритм решения задачи 2.1 обоснован.

### **3.2.4. Об эквивалентном критерии оптимальности и дополнениях к алгоритму**

Все написанное в пункте 3.1.4 относительно эквивалентного критерия оптимальности, распространяется на обобщенную задачу линейного программирования. При этом имеют место также оценки сверху для минимума в задаче с положительными коэффициентами в целевой линейной функции. Мы не выписываем здесь эквивалентный критерий (см. его в [6]), так как он очевидно записывается из 2.11. Оценки снизу имеют такой же вид, как и приведенные в пункте 3.1.4 выражения для вырожденной задачи линейного программирования.

## Глава 4

# Теория и алгоритмы решения задач $L$ - и $MV$ -оптимального планирования эксперимента и их применение

В данной главе для решения указанных задач используется вариационный подход, при котором метод наименьших квадратов заменяется в соответствии с теоремой Гаусса - Маркова оптимальным линейным несмещенным алгоритмом. Это позволило установить математическую аналогию между задачами  $L$ - и  $MV$ -оптимального планирования эксперимента и линейной идеальной коррекции параметров системы. На этой основе получены критерий оптимальности и разработаны эффективные алгоритмы решения этих задач. Работоспособность этих алгоритмов продемонстрирована при решении задач оптимального распределения измерений в случае полиномиальной регрессии, для которого найдено аналитическое решение  $MV$ -оптимальной задачи.

### 4.1. Задачи $L$ - и $MV$ -оптимального планирования и исторический комментарий

#### 4.1.1. Постановка задачи

Пусть  $\theta$  — вектор неизвестных параметров размерности  $m$ , и при  $i = 1, \dots, n$  проводится  $r_i$  измерений векторной функции  $H_i^T \theta$  размерности  $n_i$ , где  $H_i$  — заданные матрицы размера  $m \times n_i$ . Предполагается, что ошибки измерений

есть некоррелированные между собой случайные величины с нулевыми математическими ожиданиями и единичными дисперсиями. Тогда, если векторы  $y_i$  есть средние арифметические указанных выше  $r_i$  измеренных значений, то осредненные уравнения измерений могут быть записаны в виде

$$y_i = H_i^T \theta + \varepsilon_i, \quad i = 1, \dots, n, \quad (1.1)$$

где осредненные ошибки измерений  $\varepsilon_i$  есть некоррелированные между собой векторы, математические ожидания которых равны нулю, а их ковариационные матрицы есть

$$D\varepsilon_i = \frac{1}{r_i} I_{n_i},$$

где  $I_{n_i}$  — единичная матрица порядка  $n_i$ . Общий случай, когда

$$D\varepsilon_i = V_i, \quad i = 1, \dots, n$$

( $V_i$  — положительно определенные матрицы) сводится к рассматриваемому случаю умножением на  $V_i^{-\frac{1}{2}}$  неосредненных уравнений измерений.

Пусть задано общее число измерений:

$$\sum_{i=1}^n r_i = N.$$

Вместо чисел  $r_i$  будем использовать вектор  $p$  с координатами

$$p_i = r_i/N, \quad i = 1, \dots, n.$$

Далее мы будем, как это обычно делают, пренебрегать целочисленностью чисел  $r_i$  и считать, что  $p$  — непрерывный план эксперимента, т.е. любой вектор из симплекса

$$\Sigma_n = \{p \in \mathbb{R}^n : p \geq 0, \quad e^T p = 1\},$$

где  $e^T = (1, \dots, 1)$ .

Пусть  $b_j \in \mathbb{R}^m$  — заданные векторы, а

$$l_j = b_j^T \theta, \quad j = 1, \dots, s$$

— контролируемые параметры (обычно  $s \leq m$ ). При заданном плане  $p$  рассмотрим линейные оценки каждого из этих параметров:

$$\hat{l}_j = \sum_{i \in \mathcal{I}_p} \Phi_{ij}^\top y_i, \quad \mathcal{I}_p \doteq \{i : p_i > 0\}, \quad (1.2)$$

где векторы  $\Phi_{ij} \in \mathbb{R}^{n_i}$  определяют линейный оценитель параметра  $l_j$ . Этот оценитель дает несмещенные оценки параметров  $b_j^\top \theta$  тогда и только тогда, когда он удовлетворяет так называемым *условиям несмещенности* ([96], с.195; см. также ниже замечание 4.3):

$$\sum_{i \in \mathcal{I}_p} H_i \Phi_{ij} = b_j, \quad j = 1, \dots, s, \quad (1.3)$$

т.е.  $b_j$  есть линейная комбинация тех матриц  $H_i$ , для которых  $p_i > 0$ . Далее будем обозначать  $\mathcal{P}_n$  множество тех  $p$  из  $\Sigma_n$ , для которых выполняются условия несмещенности (1.3) (для некоторых  $\Phi_{ij}$ ).

Для каждого  $p \in \mathcal{P}_n$  рассмотрим наилучшие линейные несмещенные оценки, для которых оценитель определяется из условия минимальных дисперсий при линейном несмещенном оценивании:

$$ND\hat{l}_j = \min_{\Phi_{ij}} \left\{ \sum_{i \in \mathcal{I}_p} \frac{\|\Phi_{ij}\|^2}{p_i} : \sum_{i \in \mathcal{I}_p} H_i \Phi_{ij} = b_j \right\}, \quad j = 1, \dots, s. \quad (1.4)$$

Задача оптимального планирования эксперимента состоит в нахождении плана  $p \in \mathcal{P}_n$ , доставляющего нижнюю грань заданному критерию оптимальности  $L(p)$  [2 — 6]:

$$L = \inf\{L(p) : p \in \mathcal{P}_n\}. \quad (1.5)$$

Мы будем рассматривать два критерия оптимальности, предполагая далее, что оценки (1.2) являются наилучшими нелинейными несмещенными. Первый критерий

$$L(p)^2 = N \sum_{j=1}^s D\hat{l}_j \quad (1.6)$$



есть критерий *L*-оптимальности [33]. Вторым критерий

$$L(p)^2 = N \max\{D\hat{l}_1, \dots, D\hat{l}_s\} \quad (1.7)$$

будем называть *MV<sub>s</sub>-критерием* (при  $s = m$  и  $l_j = \theta_j$ ,  $j = 1, \dots, m$  критерий (1.7) называют просто *MV-критерием* [33]). В зависимости от критерия (1.6) или (1.7) задачу (1.5) будем соответственно называть *L-задачей* или *MV<sub>s</sub>-задачей*.

При решении задачи (1.5) будем предполагать, что

$$\text{rank}(H_1^T, \dots, H_n^T) = m \quad (1.8)$$

(т.е. ранг составной матрицы максимален), что равносильно допущению о возможности линейного несмещенного оценивания всех компонент вектора  $\theta$  в случае использования всех  $n$  групп измерений.

**ЗАМЕЧАНИЕ 4.1.** Задача (1.5) может быть рассмотрена для более общей по сравнению с (1.1) модели измерений

$$y(t) = H(t)^T \theta + \varepsilon(t), \quad t \in \mathcal{T},$$

где  $\mathcal{T}$  — заданное множество возможных значений некоторой векторной переменной  $t$ . Излагаемая ниже теория и алгоритмы пригодны для указанного обобщения. Однако при изложении теории мы ограничимся случаем (1.1) конечного множества  $\mathcal{T}$  и рассмотрим бесконечное множество в случае полиномиальной регрессии (см. замечание 4.13).

**ЗАМЕЧАНИЕ 4.2.** Нетрудно показать, что суммирование по  $i \in \mathcal{I}_p$  в соотношениях (1.2) — (1.4) можно заменить суммированием по  $i = 1, \dots, n$ , если доопределить слагаемые в (1.4) при  $p_i = 0$  следующим образом:

$$\frac{\|\Phi_{ij}\|^2}{p_i} = \begin{cases} 0, & \Phi_{ij} = 0 \text{ и } p_i = 0, \\ +\infty, & \Phi_{ij} \neq 0 \text{ и } p_i = 0. \end{cases}$$

В дальнейшем будем считать указанные замены при суммировании и оптимизации произведенными.

ЗАМЕЧАНИЕ 4.3. Согласно теореме Гаусса-Маркова ([96], с.201, [33], [95]), при выполнении условия (1.3) наилучшая нелинейная несмещенная оценка совпадает с оценкой взвешенного метода наименьших квадратов  $\hat{l}_j = b_j^\top \hat{\theta}$ , где  $\hat{\theta}$  — любое решение нормальных уравнений

$$M(p)\hat{\theta} = \sum_{i=1}^n p_i H_i y_i, \quad (1.9)$$

а

$$M(p) = \sum_{i=1}^n p_i H_i H_i^\top$$

— нормированная на число  $N$  информационная матрица. Эти уравнения всегда совместны, но решение  $\hat{\theta}$  будет единственно только в случае невырожденной матрицы  $M(p)$  [96]. План  $p$  в этом случае называется *невырожденным*. В противном случае план  $p$  является *вырожденным*, решение уравнений (1.9) неоднозначно. Однако условие (1.3) эквивалентно тому, что оценка  $b_j^\top \hat{\theta}$  каждой функции  $b_j^\top \theta$ ,  $j = 1, \dots, s$ , одна и та же для всех решений уравнения (1.9). При этом параметры  $b_j^\top \theta$  называются *оцениваемыми* [96]. Таким образом, условия несмещенности и оцениваемости эквивалентны. Далее, условия несмещенности (1.3) можно трактовать как принадлежность векторов  $b_j$  линейному многообразию, порожденному матрицами  $\sqrt{p_1}H_1, \dots, \sqrt{p_n}H_n$  или, эквивалентно, столбцами информационной матрицы [96]. Это означает, что найдутся векторы  $\lambda_j \in \mathbb{R}^m$  такие, что выполняется условие оцениваемости параметров  $b_j^\top \theta$ :

$$M(p)\lambda_j = b_j, \quad j = 1, \dots, s. \quad (1.10)$$

В постановку задачи (1.5) входят несколько целочисленных параметров, определяющих ее сложность:  $m$ ,  $s$ ,  $n$ ,  $n_i$ ,  $i = 1, \dots, n$ . Простейший случай  $s = 1$  оценивания одного параметра при  $n_i = 1$ ,  $i = 1, \dots, n$  рассматривался впервые в работе [89], где задача фактически сведена к задаче линейного программирования с матрицей условий размерности  $m \times n$ . Для этого случая в работах [2, 26] обнаружено совпадение в описании этой задачи и задачи о наихудшей корреляции [43]. При этом может быть найдено решение, содержащее не более  $m$

положительных компонент вектора  $p$ . В работах [35, 45] была рассмотрена задача, обратная к  $MV_s$ -задаче: минимизировать суммарное число измерений при условии ограничения сверху на дисперсии контролируемых параметров. В [45] при допущении, что оптимальный план невырожден, даны достаточные условия оптимальности  $MV_s$ -задачи (или обратной к ней) и предложен некоторый алгоритм, состоящий из серии решений обобщенной задачи линейного программирования. Однако при  $s > 2$  этот алгоритм существенно усложняется. Кроме того, в [45] показано, что существует решение, содержащее не более  $m_0 = ms - \frac{1}{2}s(s-1)$  положительных компонент вектора  $p$ . Этот результат для случая  $n_i = 1, i = 1, \dots, n$  ранее был доказан в [35]. Для  $L$ -задачи (а также для некоторых других задач планирования эксперимента) подобный результат также известен [34].

В классической теории планирования эксперимента [76, 95] обычно рассматривается случай  $n_i = 1, i = 1, \dots, n$ , и предлагаются различные градиентные алгоритмы решения  $L$ -задачи, которые, реализация которых сопряжена с определенными трудностями [33]. Нам неизвестны простые алгоритмы решения  $MV_s$ -задачи, а также легко проверяемые для нее условия оптимальности. Ниже предлагается единый подход к решению  $L$ - и  $MV_s$ -оптимальных задач, основанный на вариационном представлении дисперсий  $D\hat{l}_j$ . Он позволяет свести  $L$ -задачу к обобщенной задаче линейного программирования, а  $MV_s$ -задачу — к аналогичной задаче с векторным параметром в правых частях условий, сформулировать близкие по структуре условия оптимальности для этих задач и получить эффективные алгоритмы их решения. Настоящие исследования в большой степени стимулировала работа [45].

## 4.2. Сведение $L$ -задачи к задаче оптимальной линейной импульсной коррекции и алгоритм ее решения

В соответствии с (1.4) и замечанием 4.2 целевую функцию (1.6) можно преобразовать так:

$$\begin{aligned} L(p)^2 &= \min_{\Phi_{ij}} \left\{ \sum_{j=1}^s \sum_{i=1}^n \frac{\|\Phi_{ij}\|^2}{p_i} : \sum_{i=1}^n H_i \Phi_{ij} = b_j, \quad j = 1, \dots, s, \right\} = \\ &= \min_{u_i} \left\{ \sum_{i=1}^n \frac{\|u_i\|^2}{p_i} : \sum_{i=1}^n B_i u_i = b \right\}. \end{aligned}$$

Здесь вектор  $u_i$  новых переменных и его норма определяются из условия

$$u_i = (\Phi_{i1}^T, \dots, \Phi_{is}^T)^T \in \mathbb{R}^{n_i s}, \quad \|u_i\| = \left( \sum_{j=1}^s \|\Phi_{ij}\|^2 \right)^{\frac{1}{2}}, \quad (2.1)$$

где выражение  $\frac{\|u_i\|^2}{p_i}$  понимается как в замечании 4.2 и использованы обозначения:

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix}, \quad B_i = \begin{pmatrix} H_i & 0_{m \times n_i} & \dots & 0_{m \times n_i} \\ 0_{m \times n_i} & H_i & \dots & 0_{m \times n_i} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{m \times n_i} & \dots & 0_{m \times n_i} & H_i \end{pmatrix},$$

т.е.  $b \in \mathbb{R}^M$ ,  $B_i$  — матрица размера  $M \times n_i s$ , где  $M = ms$ ,  $0_{m \times n_i}$  — нулевая матрица размерности  $m \times n_i$ .

**ТЕОРЕМА 4.1.** *Оптимальные значения вектора  $p$ , линейного оценителя (2.1) и значение  $L$ -задачи находятся из соотношений*

$$p_i^* = \frac{\|u_i^*\|}{L}, \quad i = 1, \dots, n, \quad (2.2)$$

где  $u_i^*$  — решение задачи

$$L = \min_{u_i} \left\{ \sum_{i=1}^n \|u_i\| : \sum_{i=1}^n B_i u_i = b \right\}. \quad (2.3)$$

*Доказательство.* Подставляя по-новому записанную целевую функцию в (1.5) и переставляя затем порядок оптимизации по  $p$  и  $u_i$ , найдем аналитически экстремум по вектору  $p$  при помощи множителей Лагранжа. Для этого сначала заметим, что согласно леммам 4.2, 4.3 (см. Приложение) минимум выпуклой полунепрерывной снизу функции  $\sum_{i=1}^n \frac{\|u_i\|^2}{p_i}$  по переменной  $p$  на множестве  $\mathcal{P}_n$  можно заменить на минимум этой функции на симплексе  $\Sigma_n$ . При фиксированных переменных  $u_i$  этот минимум достигается при  $p_i = \|u_i\|(\sum_{i=1}^n \|u_i\|)^{-1}$ , откуда и получаем указанный результат.

**ЗАМЕЧАНИЕ 4.4.** При  $s = 1$  задача (2.3) является задачей линейного программирования [89, 2, 44, 18]. При этом симплекс-методом находится оптимальное решение, для которого отличны от нуля не более  $m$  переменных  $p_i$ . Отметим, что приведенное нами доказательство теоремы 4.1 строго учитывает принадлежность вектора  $p$  множеству  $\mathcal{P}_n$ , в то время как в основополагающей работе [89], где рассмотрен случай  $s = 1$ , и известных нам последующих работах (в том числе и в нашей работе [2]) минимум по вектору  $p$  отыскивался на множестве  $\Sigma_n$ , т.е. неявно предполагалась оцениваемость параметра  $l_1$  на этом множестве. Однако указанная математическая неточность в указанных работах не отразилась на правильности результата.

Задача (2.3) есть задача оптимальной линейной импульсной коррекции [44, 18, 82], где  $u_i$  и  $B_i$  — соответственно корректирующий импульс и матрица его влияния в  $i$ -й коррекции, вектор  $b$  — требуемое суммарное влияние коррекции, а целевая функция характеризует суммарные затраты на коррекцию. Задача (2.3) является частным случаем так называемой проблемы моментов [67].

Заменой переменных

$$u_i = x_i \gamma_i, \quad x_i = \|u_i\|, \quad \gamma_i \in \mathbb{R}^{n_i s}, \quad \|\gamma_i\| = 1$$

задача (2.3) сводится к следующей задаче [44, 18]:

$$L = \min_{x_i, A_i} \left\{ \sum_{i=1}^n x_i : \sum_{i=1}^n x_i A_i = b, \quad x_i \geq 0, \quad A_i \in \mathcal{A}_i, \quad i = 1, \dots, n \right\}, \quad (2.4)$$

где

$$\mathcal{A}_i = \left\{ A \in \mathbb{R}^M : A = B_i \gamma, \quad \|\gamma\| \leq 1 \right\}, \quad M = ms. \quad (2.5)$$

В (2.5) формально должно стоять условие  $\|\gamma\| = 1$ . Однако расширение множества изменения векторов условий  $A_i$  не меняет значения задачи (2.4). Это следует из того, что каждому оптимальному  $x_i > 0$  всегда соответствует оптимальный вектор  $A_i$ , принадлежащий относительной границе множества (2.5) в подпространстве

$$\left\{ A \in \mathbb{R}^M : A = B_i \gamma, \quad \gamma \in \mathbb{R}^{n_i s} \right\},$$

так как замена  $A_i$  на  $gA_i$ , где  $0 < g < 1$ , приводит к увеличению целевой функции. Относительная граница множества  $\mathcal{A}_i$  есть эллипсоид размерности  $M$ , который является линейным образом в  $\mathbb{R}^M$  единичной сферы  $\|\gamma\| = 1$  из  $\mathbb{R}^{n_i s}$ . Задача (2.4) называется *обобщенной задачей линейного программирования* и отличается от обычной задачи линейного программирования тем, что каждый вектор условий  $A_i$  не задан, а может быть произвольно выбран из заданного выпуклого множества (см. раздел 3.2 и [31, 57]).

#### 4.2.1. Необходимое и достаточное условие оптимальности $L$ -задачи

Для задачи (2.4) могут быть сформулированы необходимые и достаточные условия Куна—Таккера. Назовем *допустимым решением задачи (2.4)* положительные коэффициенты  $x_1, \dots, x_k$  и соответствующие векторы  $A_1, \dots, A_k$  (нумерация переменных задачи (2.4) условна) такие, что удовлетворяются условия  $\sum_{i=1}^k x_i A_i = b$ . Если эти векторы линейно независимы (при этом, очевидно,  $k \leq M$ ), то будем говорить о *допустимом базисном решении*.

Задача (2.4) эквивалентна следующей обобщенной задаче линейного программирования [18, 9]:

$$L = \min_{\mathcal{F}_i, x_i(A)} \left\{ \sum_{i=1}^n \sum_{A \in \mathcal{F}_i} x_i(A) : \sum_{i=1}^n \sum_{A \in \mathcal{F}_i} x_i(A) A = b \right\}, \quad (2.6)$$

где  $\mathcal{F}_i$  — конечные подмножества множества  $\mathcal{A}_i$ . Отличие задачи (2.4) от задачи (2.6) состоит в том, что в последней допускается наличие нескольких векторов из одного и того же множества  $\mathcal{A}_i$  с ненулевыми коэффициентами  $x_i(A)$ . Однако оптимальные решения этих задач совпадают [18, 9]. Для задачи (2.6), а следовательно, и для задачи (2.4) справедлив аналог теоремы Куна—Таккера ([18], с.275).

**ТЕОРЕМА 4.2 (КУНА—ТАККЕРА).** *Допустимое решение задачи (2.4) является оптимальным тогда и только тогда, когда существует вектор  $\pi$  размерности  $M$  такой, что*

$$\pi^T A_i = 1, \quad \text{если } x_i > 0, \quad (2.7)$$

$$\pi^T A \leq 1 \quad \forall A \in \bigcup_{i=1}^n \mathcal{A}_i. \quad (2.8)$$

**ЗАМЕЧАНИЕ 4.5.** Идея доказательства теоремы 4.2 состоит в следующем [18]. Конструируется обычная задача линейного программирования, векторы условий которой есть векторы оптимального допустимого базисного решения решения задачи (2.4) и любой вектор  $A$  из  $\bigcup_{i=1}^n \mathcal{A}_i$ . Если применить к этой задаче теорему Куна—Таккера и воспользоваться произвольностью вектора  $A$ , то получим теорему 4.2. Эта теорема также следует из результатов в [70].

**ЗАМЕЧАНИЕ 4.6.** Из (2.7) нетрудно получить, что если существует оптимальное решение задачи (2.4), то существует и оптимальное допустимое базисное решение, удовлетворяющее (2.7), (2.8).

Представим вектор  $\pi$  в виде составного вектора

$$\pi^T = (\pi_1^T, \dots, \pi_s^T), \quad (2.9)$$

где  $\pi_j \in \mathbb{R}^m$ . Условия (2.7), (2.8) тогда могут быть записаны в виде

$$\beta \doteq \max\{\beta_1, \dots, \beta_n\} = 1, \quad (2.10)$$

где величины  $\beta_i$  вычисляются аналитически:

$$\begin{aligned} \beta_i &= \max\{\pi^T A : A \in \mathcal{A}_i\} = \\ &= \max\{\pi^T B_i \gamma : \|\gamma\| = 1\} = \|B_i^T \pi\| = \left( \sum_{j=1}^s \|H_i^T \pi_j\|^2 \right)^{\frac{1}{2}}. \end{aligned} \quad (2.11)$$

Максимум в (2.11) достигается при

$$A_i = B_i \gamma_i, \quad \gamma_i = \frac{1}{\beta_i} B_i^T \pi \Leftrightarrow \gamma_{ij} = \frac{1}{\beta_i} H_i^T \pi_j, \quad j = 1, \dots, s, \quad (2.12)$$

где  $\gamma_{ij} \in \mathbb{R}^{n_i}$  — блоки составного вектора  $\gamma_i^T = (\gamma_{i1}^T, \dots, \gamma_{is}^T) \in \mathbb{R}^{n_i s}$ .

Теорема Куна—Таккера, дополненная условиями оцениваемости (1.10) или (1.3) (эквивалентными условиям-равенствам в (2.4)), дают следующие необходимые и достаточные условия оптимальности для  $L$ -задачи, обсуждаемые ниже.

**ТЕОРЕМА 4.3 (НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ ОПТИМАЛЬНОСТИ ДЛЯ  $L$ -ЗАДАЧИ).** *План  $p^*$  является  $L$ -оптимальным тогда и только тогда, когда найдутся векторы  $\lambda_j \in \mathbb{R}^m$ ,  $j = 1, \dots, s$ , такие, что выполняются условия:*

- а)  $M(p^*)\lambda_j = b_j$ ,  $j = 1, \dots, s$  (т.е. выполняется условие оцениваемости (1.10));
- б)  $\max_{1 \leq i \leq n} \sum_{j=1}^s \|H_i^T \lambda_j\|^2$  достигается при всех  $i$  таких, что  $p_i^* > 0$ , и этот максимум равен  $L^2$ , где  $L$  есть оптимальное значение  $L$ -задачи.

*Доказательство.* Согласно теореме 4.1, условия оптимальности  $L$ -задачи эквивалентны этим условиям для задачи (2.4). Для векторов  $A_i$ ,  $i = 1, \dots, k$ ,



удовлетворяющих условию (2.7), величины  $\beta_i$  в (2.11) равны единице. Поэтому из (2.12) следует, что  $A_i = B_i B_i^T \pi$ ,  $i = 1, \dots, k$ , и условия-равенства задачи (2.4) (которые являются условиями несмещенности (1.3)) с учетом выражения (2.2) на оптимальном решении примут вид:

$$\sum_{i=1}^n L p_i^* B_i B_i^T \pi = b \Leftrightarrow L \sum_{i=1}^n p_i^* H_i H_i^T \pi_j = b_j, \quad j = 1, \dots, s. \quad (2.13)$$

Вводя векторы

$$\lambda_j = \pi_j L, \quad j = 1, \dots, s,$$

и переписывая полученные необходимые и достаточные условия оптимальности (2.10), (2.13), получим доказательство теоремы 4.3.

**ЗАМЕЧАНИЕ 4.7.** Необходимые и достаточные условия оптимальности для  $L$ -задачи, насколько нам известно, ранее были сформулированы в похожем на теорему 4.3 виде лишь при допущении невырожденности информационной матрицы  $M(p^*)$  [33]. В частности, при условии, что  $\text{rank}(b_1, \dots, b_s) = m$ , информационная матрица невырождена [33] (что следует из (1.10)) и это позволяет иногда находить решение при помощи теоремы 4.3.

## 4.2.2. Получение оптимального плана с минимальным числом положительных компонент

Пусть каким-либо образом найдены фигурирующие в теореме 4.3 векторы  $\lambda_j$ , соответствующие оптимальному плану  $p^*$ , у которого компоненты  $p_i^*$  положительны при  $i = 1, \dots, \tilde{n}$  ( $\tilde{n} \leq n$ , нумерация по  $i$  условна). Рассмотрим уравнения условия а) теоремы 4.3 как систему линейных уравнений

$$\sum_{i=1}^{\tilde{n}} p_i^* P_{ij} = b_j, \quad j = 1, \dots, s \quad (2.14)$$

относительно неизвестных значений  $p_i^*$ . Матрица  $(P_{ij})$  имеет ранг [45]

$$m_0 = M - \frac{s(s-1)}{2}.$$

Поэтому существует базисное решение системы (2.14), содержащее не более  $m_0$  положительных компонент  $p_i$ . Это решение находится стандартным в теории линейного программирования приемом — введением искусственных переменных и решением задачи линейного программирования с матрицей условий размера  $m_0 \times \tilde{n}$  [31, 57]. Получаемый в итоге план  $p_0$  с не более чем  $m_0$  положительными компонентами удовлетворяет, очевидно, условию б) теоремы 4.3. Поэтому он является оптимальным для  $L$ -задачи.

### 4.2.3. Алгоритм решения задачи (2.4)

Введем некоторые понятия (см. раздел 3.2 и [18, 9]), аналогичные используемым в линейном программировании. Невырожденная матрица

$$B = (A_1, \dots, A_k, A_{k+1}, \dots, A_M), \quad (2.15)$$

включающая все векторы  $A_i, i = 1, \dots, k$ , которым в (2.4) соответствуют коэффициенты  $x_i > 0$ , и любые векторы  $A_{k+1}, \dots, A_M \in \bigcup_{i=1}^n \mathcal{A}_i$ , называется *базисной*. Если  $k = M$ , то допустимое базисное решение является *невырожденным* и соответствует, согласно (2.2), невырожденному плану  $p$ , в противном случае — *вырожденным*. Вектор  $\pi$  в (2.7) определяется однозначно лишь для невырожденного допустимого базисного решения. Поэтому оптимальное допустимое базисное решение может отыскиваться на основе достаточного условия оптимальности, которое будет необходимым, если проверяемый план является невырожденным [44, 18, 31, 57]. Опишем этот способ, называемый *методом генерации столбцов* для задачи типа (2.4) [57]. Пусть  $B$  — некоторая базисная матрица (2.15), а вектор  $\pi$  определяется из условия

$$\pi^T B = e^T. \quad (2.16)$$

Тогда, если выполняется условие (2.10), то базисный вектор

$$x_B = B^{-1}b$$

является оптимальным. В противном случае в базис по правилам симплекс-метода вводится тот вектор  $A_i$ , определяемый из (2.12), на котором достигается максимум в (2.10). При этом целевая функция для невырожденного допустимого базисного решения ( $k = M$  в (2.7)) уменьшается, а для вырожденного — не увеличивается. В общем случае оптимальное решение достигается за бесконечное число итераций. При этом, вообще говоря, может не существовать базиса, для которого удовлетворяется достаточное условие оптимальности (2.10) [18]. Однако всегда существует множество базисов, для каждого из которых  $\beta = 1 + \varepsilon$ , где  $\varepsilon$  — произвольное малое число. Если  $L_\varepsilon$  — значение целевой функции в (2.10), соответствующее одному из этих базисов, то имеет место оценка значения в (2.4) [44]:

$$\frac{L_\varepsilon}{1 + \varepsilon} \leq L \leq L_\varepsilon.$$

При этом с требуемой точностью находится почти оптимальное решение.

Таким образом, если величина  $1 - \beta$  стремится к нулю при увеличении числа итераций, то алгоритм сходится в том смысле, что мы можем найти сколь угодно близкое к величине  $L$  значение целевой функции. Большой опыт вычислений показывает, что при разумном задании точности  $\varepsilon$  (не более  $10^{-10}$ ) указанный метод генерирования столбцов всегда позволяет получить почти оптимальное решение. При этом алгоритм так же эффективен, как и обычный симплекс-метод для задачи линейного программирования с матрицей условий размера  $M \times n$ . Теоретически возможный случай, когда  $1 - \beta$  не стремится к нулю при увеличении числа итераций, возможен лишь тогда, когда предельное допустимое базисное решение является вырожденным. На практике нам такой случай не встречался, но неосуществимость его не доказана. В этом случае может быть использован предложенный в [9] способ строгого уменьшения целевой функции, соответствующей вырожденному допустимому базисному решению.

**ЗАМЕЧАНИЕ 4.8.** При нахождении оптимального решения задачи (2.4) мы также находим и вектор  $\pi$  из (2.16) и затем векторы  $\lambda_j$ , как это указано при доказательстве теоремы 4.3. Далее, как указано выше, можно получать оптималь-

ный план, содержащий не более  $m_0$  положительных компонент, находя базисное решение системы (2.14) при  $\tilde{n} \leq M$ .

### 4.3. Сведение $MV_s$ -задачи к параметрической задаче оптимальной линейной импульсной коррекции и алгоритм ее решения

Целевую функцию (1.7) можно записать в виде

$$L(p)^2 = N \max_{\mu \in \Sigma_s} \sum_{j=1}^s \mu_j D(\hat{l}_j), \quad p \in \mathcal{P}_n, \quad (3.1)$$

где искусственно вводимый вектор переменных  $\mu^T = (\mu_1, \dots, \mu_s)$  принадлежит симплексу

$$\Sigma_s = \{\mu \in \mathbb{R}^s, \quad \mu \geq 0 : \quad e^T \mu = 1\}.$$

Таким образом, задача (1.5) становится минимаксной. В Приложении доказаны следующие результаты:

**ТЕОРЕМА 4.4.** *Если параметры  $b_j^T \theta, j = 1, \dots, s$  являются оцениваемыми в смысле замечания 4.3, то в минимаксной задаче (1.5), (3.1) операции максимума и инфимума можно поменять местами.*

**ТЕОРЕМА 4.5** (НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ ОПТИМАЛЬНОСТИ ДЛЯ  $MV_s$ -ЗАДАЧИ). *План  $p^* \in \Sigma_n$  является  $MV_s$ -оптимальным тогда и только тогда, когда найдутся вектор  $\mu^* \in \Sigma_s$  и векторы  $\lambda_j \in \mathbb{R}^m, j = 1, \dots, s$ , удовлетворяющие условию а) теоремы 4.3 и условиям:*

- б)  $\max_{1 \leq i \leq n} \sum_{j=1}^s \mu_j^* \|H_i^T \lambda_j\|^2$  достигается при всех  $i$  таких, что  $p_i^* > 0$ ;
- в)  $\max_{1 \leq j \leq s} b_j^T \lambda_j$  достигается при всех  $j$  таких, что  $\mu_j^* > 0$ .

*При этом значения максимумов в б) и в) совпадают и равны  $L^2$ , где  $L^2$  — значение  $MV_s$ -задачи.*

ЗАМЕЧАНИЕ 4.9. В работе [45] для случая, когда план  $p^*$  — невырожденный, доказана достаточность условий а), б), в) теоремы 4.5 для оптимальности этого плана.

Используя утверждение теоремы 4.4 и записывая вариационное представление дисперсий (1.4), получим:

$$L^2 = \max_{\mu \in \Sigma_s} \inf_{p \in \mathcal{P}_n} \min_{\Phi_{ij}} \left\{ \sum_{j=1}^s \sum_{i=1}^n \mu_j \frac{\|\Phi_{ij}\|^2}{p_i} : \sum_{i=1}^n H_i \Phi_{ij} = b_j, \text{ если } \mu_j > 0 \right\}.$$

Условие  $\mu_j > 0$  вводится здесь, чтобы избежать неопределенностей  $0 \cdot \infty$  (см. замечание 4.2). Переставляя здесь операции минимизации по  $p_i$  и  $\Phi_{ij}$ , меняя порядок суммирования по  $i$  и  $j$  и оптимизируя затем по  $p_i$  как и при доказательстве теоремы 4.1, находим

$$L = \max\{L(\mu) : \mu \in \Sigma_s\}, \quad (3.2)$$

где  $L(\mu)$  есть значение подзадачи

$$L(\mu) = \min_{\Phi_{ij}} \left\{ \sum_{i=1}^n \left( \sum_{\mu_j > 0} \mu_j \|\Phi_{ij}\|^2 \right)^{\frac{1}{2}} : \sum_{i=1}^n H_i \Phi_{ij} = b_j, \text{ если } \mu_j > 0 \right\}. \quad (3.3)$$

Если при этом  $\mu_j^*$  — оптимальное значение в (3.2),  $\Phi_{ij}^*$  — оптимальные значения в (3.3), то

$$p_i^* = \frac{\left( \sum_{j=1}^s \mu_j^* \|\Phi_{ij}^*\|^2 \right)^{1/2}}{L}, \quad i = 1, \dots, n.$$

Будем рассматривать пока только векторы  $\mu$  с положительными компонентами ( $\mu > 0$ ). Потом мы перейдем к общему случаю (см. замечание 4.11). При  $\mu > 0$  можно ввести новые переменные, аналогично (2.1):

$$u_i = (\sqrt{\mu_1} \Phi_{i1}^T, \dots, \sqrt{\mu_s} \Phi_{is}^T)^T \in \mathbb{R}^{n_i s}$$

Тогда подзадача (3.3) запишется в виде

$$L(\mu) = \min_u \left\{ \sum_{i=1}^n \|u_i\| : \sum_{i=1}^n B_i u_i = b(\mu) \right\}, \quad (3.4)$$

где  $b(\mu)^T = (\sqrt{\mu_1}b_1^T, \dots, \sqrt{\mu_s}b_s^T)^T \in \mathbb{R}^M$ .

Задача (3.2) с учетом (3.4) по терминологии [18] есть проектная задача оптимальной линейной импульсной коррекции, в которой оптимальный корректируемый вектор  $b(\mu^*)$  отыскивается из условия, что минимальные затраты на коррекцию при каждом  $\mu$  будут максимальными среди всех  $\mu \in \Sigma_s$ .

**ЗАМЕЧАНИЕ 4.10.** Подзадача (3.4) может быть записана в виде обобщенной задачи линейного программирования (2.4) с заменой  $L$  на  $L(\mu)$  и  $b$  на  $b(\mu)$ . Поэтому она при каждом фиксированном  $\mu$  эффективно решается методом генерации столбцов, который был рассмотрен в предыдущем разделе.

### 4.3.1. Редукция задачи (3.2) к задаче многомерной максимизации и алгоритм ее решения

Используя идеи, отмеченные в замечании 4.5, нетрудно показать, что двойственная задача к подзадаче (3.4) имеет вид:

$$L(\mu) = \max\{b^T(\mu)\pi : \pi \in \Pi\}, \quad (3.5)$$

$$\begin{aligned} \Pi &\doteq \left\{ \pi \in \mathbb{R}^M : \|B_i^T \pi\| \leq 1, \quad i = 1, \dots, n \right\} = \\ &= \left\{ \pi \in \mathbb{R}^M : \pi^T A \leq 1 \quad \forall A \in \bigcup_{i=1}^n \mathcal{A}_i \right\}. \end{aligned}$$

Другой способ получения соотношений (3.5) основан на работе [70].

**ЗАМЕЧАНИЕ 4.11.** Задача (3.5) есть обобщенная задача линейного программирования, значение которой, как известно [31, 57], непрерывно зависит от вектора  $b(\mu)$  и, следовательно, от вектора  $\mu$ . Поэтому соотношение двойственности (3.5) имеет место для всех  $\mu \in \Sigma_s$ . Если при этом некоторые компоненты вектора  $\mu$  равны нулю, то соответствующие уравнения в условиях задачи (3.4) не учитываются. Необходимо только следить, чтобы соответствующие векторы  $b_j$  удовлетворяли условию оцениваемости (1.3).

Соотношение (3.5) позволяет записать задачу (3.2) в виде задачи максимизации по  $\pi$  и  $\mu$ :

$$L = \max_{\pi, \mu} \{b^T(\mu)\pi : \pi \in \Pi, \mu \in \Sigma_s\}. \quad (3.6)$$

Используя симметричность эллипсоидов  $\mathcal{A}_i$ , нетрудно показать, что максимум по  $\mu$  при фиксированном  $\pi$  достигается при

$$\mu_j = \frac{(b_j^T \pi_j)^2}{\sum_{j=1}^s (b_j^T \pi_j)^2}, \quad j = 1, \dots, s, \quad (3.7)$$

где  $\pi_j \in \mathbb{R}^m$  определены в (2.9). Этот максимум равен

$$\max \{b^T(\mu)\pi : \mu \in \Sigma_s\} = \left( \sum_{j=1}^s (b_j^T \pi_j)^2 \right)^{\frac{1}{2}}.$$

Теперь в соответствии с (3.7) можно предложить конструктивный способ покоординатного увеличения целевой функции в (3.6). Возьмем некоторое значение  $\mu^0 \in \mathbb{R}^s, \mu^0 > 0$  и определим последовательности  $\mu^k, \pi^k$  следующим образом. Вектор  $\mu^{k+1}$  находится по  $\pi^k$  из формул (3.7). Вектор  $\pi^k$  определяется по  $\mu^k$  из формулы (2.16), где оптимальная базисная матрица  $B$  находится, согласно замечанию 4.10, путем решения задачи (3.4) методом генерации столбцов. Согласно (3.6), получаем

$$b(\mu^k)^T \pi^k \leq b(\mu^{k+1})^T \pi^k \leq b(\mu^{k+1})^T \pi^{k+1}, \quad k = 0, 1, 2, \dots \quad (3.8)$$

**ЗАМЕЧАНИЕ 4.12.** Процедура (3.8) позволяет монотонно увеличивать целевую функцию в (3.6). Однако, если некоторые компоненты вектора  $\mu$  обращаются в нуль, то как показывает опыт, это может не привести к нахождению глобального максимума в (3.6). В работе [29], использующей результаты настоящей статьи,  $MV$ -задача сводится к обобщенной задаче линейного программирования, что позволяет эффективно находить глобальный оптимум. В то же время отметим, что если в процедуре (3.8) существует предел  $\tilde{\mu} > 0$  последовательности  $\mu^k$ , то он равен оптимальному значению  $\mu^*$  в задаче (3.5). Действительно, если

$\tilde{\mu} > 0$ , то все векторы  $\lambda_j$  определяются однозначно через векторы  $\pi_j$  согласно теореме 4.5. При этом нетрудно проверить выполнение всех условий теоремы 4.5, используя соотношения (3.7).

#### 4.4. Нахождение аналитического решения для случая для случая полиномиальной регрессии.

Случай полиномиальной регрессии характеризуется следующими условиями. На отрезке  $[-1, 1]$  задаются векторы  $H_m(t) = (1, t, \dots, t^{m-1})^T$ , которые являются исходными для  $L$ - и  $MV$ -задач.

**ЗАМЕЧАНИЕ 4.13.** Здесь мы рассматриваем в соответствии с замечанием 4.1 оптимальные задачи планирования с бесконечным множеством возможных моментов измерений. При этом возникают обобщения задач (2.4), которые нужно решать с той лишь разницей по сравнению с дискретным случаем, что величина  $\beta$  определяется по формуле (ср. с (2.10))

$$\beta = \max_{t \in [-1, 1]} \left( \sum_{j=1}^s \|H_m(t)^T \pi_j\|^2 \right)^{\frac{1}{2}}.$$

При вычислении величины  $\beta$  находится максимум полинома степени  $2m - 2$ . Если не учитывать трудоемкость операций по нахождению этого максимума, то решение обобщения задачи (2.3) для бесконечного множества возможных моментов измерений по трудоемкости сопоставимо с решением обычной задачи линейного программирования с матрицей условий размера  $M \times (M + 1)$  [18].

В случае  $s = m$  и  $l_j = \theta_j$ ,  $j = 1, \dots, m$   $L$ -задача называется  $A$ -задачей. Используя теоремы 4.3 и 4.5 и проводя те же рассуждения, что и при исследовании  $D$ -оптимальной задачи в случае полиномиальной регрессии ([33], с.114), нетрудно получить, что существуют оптимальные планы для  $A$ - и  $MV$ -задач, которые сосредоточены в  $m$  точках, симметричных относительно точки  $t = 0$ , которая



входит в оптимальный план при нечетном  $m$ . Точки  $t = -1$  и  $t = 1$  всегда входят в оптимальный план, точка  $t = 0$  — только при нечетном  $m$ , а весовая функция  $p(t)$  является четной. Эти выводы для  $A$ -задачи получены в [95], где найдены  $A$ -оптимальные планы для полиномиальной регрессии при  $m \leq 11$ . Такие же планы получены нами путем численного решения задачи (2.4), которая работает достаточно эффективно при  $m \leq 11$ .

Перейдем к решению  $MV$ -задачи для полиномиальной регрессии. Предварительно выпишем из [95] (с.230 — 238) решение  $C$ -оптимальной задачи планирования эксперимента, в которой критерий оптимальности — дисперсия  $D\hat{\theta}_j$  скалярного параметра  $\theta_j$ . Для параметров  $\theta_m, \theta_{m-2}, \dots$  минимальные дисперсии равны коэффициентам при  $t^{m-1}, t^{m-2}, \dots$  полинома Чебышева степени  $m-1$ , а оптимальные точки  $t_i$ , в которых весовая функция  $p_j(t)$ , соответствующая каждому параметру  $\theta_j, j = m, m-2, \dots$ , не равна нулю, есть точки максимума полинома Чебышева:

$$t_i = \cos\left(\frac{m-i}{m-1}\pi\right), \quad i = 1, \dots, m.$$

При этом, согласно теореме 4.1,

$$p_j(t_i) = \frac{|\Phi_{ij}|}{\sum_{i=1}^m |\Phi_{ij}|} \quad i = 1, \dots, m,$$

где  $\Phi_{ij}$  определяются из системы уравнений

$$\sum_{i=1}^m \Phi_{ij} H_m(t_i) = e_j.$$

Здесь  $e_j$  —  $j$ -й единичный вектор. Для параметров  $\theta_{m-1}, \theta_{m-3}, \dots$  формулы оптимального распределения имеют тот же вид, что и для параметров  $\theta_m, \theta_{m-2}, \dots$ , но с заменой в вышеприведенных формулах  $m$  на  $m-1$ .

Теперь решим  $MV$ -задачу для  $m \leq 11$ . При  $m \leq 4$  строго наибольшей среди  $C$ -оптимальных дисперсий  $D\hat{\theta}_1, \dots, D\hat{\theta}_m$  является  $D\hat{\theta}_m$  и, кроме того, на  $C$ -оптимальном плане для параметра  $\theta_m$  справедливы неравенства  $D\hat{\theta}_j < D\hat{\theta}_m \quad \forall j < m$ . Это означает, что  $C$ -оптимальный план для параметра  $\theta_m$

является  $MV$ -оптимальным при всех  $m \leq 4$ . Аналогично показывается, что при  $6 \leq m \leq 11$   $MV$ -оптимальным является  $C$ -оптимальный план для параметра  $\theta_{m-2}$ . При  $m = 5$  имеет место равенство  $D\hat{\theta}_3 = D\hat{\theta}_5$  для  $C$ -оптимальных дисперсий. В этом случае решение  $MV$ -задачи можно найти, подбирая веса  $p(t_i) = p_i^*$  в точках  $t_i$  максимума полинома Чебышева и параметры  $\mu_3$  и  $\mu_5$  из условия  $\mu_3 + \mu_5 = 1$ , так чтобы удовлетворялись необходимые и достаточные условия теоремы 4.5. При этом оказывается, что  $\mu_3 = \frac{8}{15}$ .

Приведем найденные указанным способом оптимальный спектр распределения измерений  $p(t)$  при  $t \geq 0$  (функция  $p(t)$  — четная) и значения  $L$  задачи (3.2).

$m$	Весовая функция $p(t)$ при $t \geq 0$			
2	$p(1) = \frac{1}{2};$			
3	$p(0) = \frac{1}{2},$	$p(1) = \frac{1}{4};$		
4	$p(\frac{1}{2}) = \frac{1}{3},$	$p(1) = \frac{1}{6};$		
5	$p(0) = \frac{5/2}{4+\sqrt{15}},$	$p(\frac{1}{\sqrt{2}}) = \frac{1}{2} - \frac{2}{4+\sqrt{15}},$	$p(1) = \frac{3/4}{4+\sqrt{15}};$	
6	$p(0, 309) = 0, 265,$	$p(0, 809) = 0, 175,$	$p(1) = 0, 060;$	
7	$p(0) = 0, 222,$	$p(0, 5) = 0, 194,$	$p(\frac{\sqrt{3}}{2}) = 0, 139,$	$p(1) = 0, 055;$
8	$p(0, 222) = 0, 179,$	$p(0, 623) = 0, 152,$	$p(0, 901) = 0, 117,$	$p(1) = 0, 051;$
9	$p(0) = 0, 156,$	$p(0, 383) = 0, 147,$	$p(\frac{1}{\sqrt{2}}) = 0, 124,$	$p(0, 924) = 0, 103,$ $p(1) = 0, 047;$
10	$p(0, 174) = 0, 134,$	$p(0, 5) = 0, 123,$	$p(0, 766) = 0, 106,$	$p(0, 940) = 0, 092,$ $p(1) = 0, 043;$
11	$p(0) = 0, 120,$	$p(0, 309) = 0, 116,$	$p(0, 588) = 0, 106,$	$p(0, 809) = 0, 093,$ $p(0, 951) = 0, 083,$
		$p(1) = 0, 040.$		

$$L = \begin{cases} 2^{m-2}, & 2 \leq m \leq 4, \\ \frac{16}{\sqrt{15}} + 4, & m = 5, \\ (m-1)2^{m-4}, & 6 \leq m \leq 11. \end{cases}$$

Указанное выше решение было получено также численно при помощи изложенного в предыдущем разделе алгоритма, который оказался достаточно трудоемким по объему вычислений, так как медленно шел процесс увеличения целевой функции при изменении параметра  $\mu$  по формуле (3.7).

## ПРИЛОЖЕНИЕ

*Доказательство теоремы 4.4.* Согласно [96], если параметр  $b_j\theta$  является оцениваемым, т.е. если выполняется условие (1.10), то функционал (3.1) равен

$$ND(\hat{l}_j) = b_j^T M(p)^+ b_j, \quad (4.1)$$

где  $M(p)^+$  — псевдообратная к  $M(p)$  матрица Мура—Пенроуза (вместо матрицы  $M(p)^+$  может быть использована любая г-обратная матрица  $M(p)^-$ , определяемая из условия  $M(p)M(p)^-M(p) = M(p)$  [96]). Согласно предположению теоремы 4.4, множество  $\mathcal{P}_n$ , определенное после формулы (1.3), непусто. Кроме того, по приведенной ниже лемме 4.1 оно выпукло. Из (3.1), (1.5), (4.1) получаем:

$$L^2 = \inf_{p \in \Sigma_n} \max_{1 \leq j \leq s} ND(\hat{l}_j) = \inf_{p \in \mathcal{P}_n} \max_{\mu \in \Sigma_s} \mathcal{L}(\mu, p), \quad (4.2)$$

где лагранжиан

$$\mathcal{L}(\mu, p) = \sum_{j=1}^s \mu_j b_j^T M(p)^+ b_j : \quad \Sigma_s \times \mathcal{P}_n \longrightarrow \mathbb{R} \quad (4.3)$$

при фиксированном  $p \in \mathcal{P}_n$  линеен по переменной  $\mu$  на симплексе  $\Sigma_s$ . Кроме того, по лемме 4.1 он является выпуклым по переменной  $p$  на выпуклом множестве  $\mathcal{P}_n$ . По несимметричной теореме о минимаксе [63] запишем (4.2) в виде

$$L^2 = \sup_{\mu \in \Sigma_s} \inf_{p \in \mathcal{P}_n} \mathcal{L}(\mu, p) = \inf_{p \in \mathcal{P}_n} \max_{\mu \in \Sigma_s} \mathcal{L}(\mu, p). \quad (4.4)$$

В силу леммы 4.1 функционал в (4.4) полунепрерывен снизу по переменной  $p$  на всем симплексе  $\Sigma_n$  (и равен  $+\infty$  вне множества  $\mathcal{P}_n$ ). По теореме Вейерштрасса записанная слева в (4.2) исходная  $MV_s$  задача, а вместе с ней и эквивалентная ей задача

$$L^2 = \min_{p \in \mathcal{P}_n} \max_{\mu \in \Sigma_s} \mathcal{L}(\mu, p), \quad (4.5)$$

имеют решение  $p^* \in \mathcal{P}_n$ . Далее, нижняя грань лагранжианов (4.3) по переменной  $p \in \mathcal{P}_n$  есть конечная вогнутая полунепрерывная сверху на симплексе  $\Sigma_s$  функция, которая по теореме Вейерштрасса достигает на нем своего максимума в некоторой точке  $\mu^* \in \Sigma_s$ :

$$L^2 = \max_{\mu \in \Sigma_s} \inf_{p \in \mathcal{P}_n} \mathcal{L}(\mu, p). \quad (4.6)$$

Теорема 4.4 доказана.

*Доказательство теоремы 4.5.* Согласно соотношению (3.6), необходимые и достаточные условия оптимальности  $MV_s$ -задачи складываются из таких условий для задач (3.2) и (3.3). Необходимые и достаточные условия последней задачи могут быть записаны, согласно замечанию 4.10, в виде условий а) и б) теоремы 4.5. При этом для  $\mu_j > 0$  получаем  $\lambda_j = \pi_j^* L / \sqrt{\mu_j^*}$ , а при  $\mu_j = 0$  вектор  $\lambda_j$  есть произвольное решение системы (1.10).

Равенства (4.4) — (4.6) эквивалентны существованию седловой точки  $(\mu^*, p^*)$  у лагранжиана (4.3) ([83], с.172). Иными словами,

$$\mathcal{L}(\mu^*, p^*) = \max_{\mu \in \Sigma_s} \mathcal{L}(\mu, p^*), \quad (4.7)$$

$$\mathcal{L}(\mu^*, p^*) = \min_{p \in \mathcal{P}_n} \mathcal{L}(\mu^*, p). \quad (4.8)$$

Условие (4.8) может быть использовано для доказательства условий а) и б) теоремы 4.5, которые, однако уже получены нами другим путем. Для лагранжиана (4.3) условие (4.7) равносильно тому, что

$$\max_{1 \leq j \leq s} b_j^T M^+(p^*) b_j \quad (4.9)$$

достигается при всех  $j$  таких, что  $\mu_j^* > 0$ , так как в противном случае функцию  $\mathcal{L}(\mu, p^*)$  можно было бы увеличить, уменьшив до нуля то значение  $\mu_j > 0$ , для которого максимум в (4.9) не достигается, и увеличив то значение  $\mu_j$ , для которого этот максимум достигается.

Ввиду тождества

$$b_j^T M(p^*)^+ b_j = \lambda_j^T M(p^*) M(p^*)^+ M(p^*) \lambda_j = \lambda_j^T M(p^*) \lambda_j = \lambda_j^T b_j$$

условие (4.9) можно записать в виде условия в) теоремы 4.5. Таким образом, значение максимума в условии в), равное максимуму в (4.9), совпадает с величиной  $\mathcal{L}(\mu^*, p^*) = L^2$ , как это нетрудно получить из (4.7).

**ЛЕММА 4.1.** Пусть  $K$  - неотрицательно определенная матрица (обозначается  $K \geq 0$ ). Тогда:

- а) функция  $y^T K^+ y$  выпукла и полунепрерывна снизу по вектору  $y$  и матрице  $K$  в совокупности;
- б) справедливо тождество

$$\frac{y^T K^+ y}{2} = \sup_x \left\{ x^T y - \frac{x^T K x}{2} \right\}, \quad (4.10)$$

причем верхняя грань в (4.10) конечна тогда и только тогда, когда вектор  $y$  принадлежит образу оператора  $K$ , т.е.  $Kx = y$  для некоторого  $x$ ; на любом таком  $x$  достигается максимум в (4.10).

*Доказательство.* Утверждение а) следует из линейности по рассматриваемым переменным целевой функции в (4.10) [68]. Утверждение б) получается прямыми вычислениями ([68], с.125).

**ЛЕММА 4.2.** *Если функция  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$  выпукла и  $\bar{\Sigma}$  — замыкание выпуклого множества  $\Sigma \subset \mathbb{R}^n$ , то*

$$\inf_{x \in \Sigma} f(x) = \inf_{x \in \bar{\Sigma}} f(x). \quad (4.11)$$

*Доказательство.* Пусть  $\bar{x} \in \bar{\Sigma}$  и  $x_0 \in \text{ri } \Sigma$  ( $\text{ri } \Sigma$  — относительная внутренность множества  $\Sigma$ ). Тогда  $\alpha x_0 + (1 - \alpha)\bar{x} \in \text{ri } \Sigma$  при любом  $\alpha \in (0, 1)$  [68]. В силу выпуклости функции одной переменной  $f(\alpha x_0 + (1 - \alpha)\bar{x})$  на отрезке  $[0, 1]$  получаем, что

$$\overline{\lim}_{\alpha \rightarrow +0} f(\alpha x_0 + (1 - \alpha)\bar{x}) \leq f(\bar{x})$$

и поэтому правая часть (4.11) не меньше ее левой части. Обратное неравенство очевидно.

**ЛЕММА 4.3.** *Замыкание выпуклого множества  $\mathcal{P}_n$ , введенного после формулы (1.3), совпадает со всем симплексом  $\Sigma_n$ .*

*Доказательство.* Относительная внутренность симплекса  $\Sigma_n$  есть  $\text{ri } \Sigma_n = \{p \in \Sigma_n : p_i > 0 \ \forall \ i = 1, \dots, n\}$ . Докажем включение  $\text{ri } \Sigma_n \subset \mathcal{P}_n$ . Для этого достаточно проверить, что все матрицы  $M(p)$  при  $p \in \text{ri } \Sigma_n$  невырождены. Предположим противное:  $M(\bar{p})x = 0$  для некоторых векторов  $\bar{p} \in \text{ri } \Sigma_n$ ,  $x \neq 0$ . Тогда получим

$$0 = x^T M(\bar{p})x = \sum_{i=1}^n \bar{p}_i x^T M_i x, \quad M_i \doteq H_i H_i^T \geq 0,$$

хотя в то же время

$$x^T M_i x \geq 0, \quad \bar{p}_i > 0, \quad i = 1, \dots, n.$$

Отсюда следует, что  $x^T M_i x = 0$ , т.е.  $M_i x = 0 \ \forall \ i = 1, \dots, n$ . Тогда при любом  $p \in \Sigma_n$  получим

$$M(p)x = \sum_{i=1}^n p_i M_i x = 0,$$

т.е. все информационные матрицы  $M(p)$  вырождены. Это противоречит условию (1.8).

Таким образом,

$$\Sigma_n = \overline{\text{ri } \Sigma_n} \subset \overline{\mathcal{P}_n} \subset \overline{\Sigma_n} = \Sigma_n,$$

что и требовалось доказать.

## 4.5. Оптимальное планирование лазерных наблюдений спутников ЛАГЕОС-1,2

### 4.5.1. Введение

Одной из важнейших как с научной, так и с прикладной точек зрения, является задача определения параметров вращения Земли по лазерным наблюдениям ее искусственных спутников. Число используемых наблюдений обычно ограничено вследствие ряда причин, возникающих при использовании большого числа наблюдений: плохой вычислительной эффективностью алгоритма оценивания; потерей точности при решении обычно используемого метода наименьших квадратов; бесполезностью привлечения дополнительных наблюдений в связи с заметным влиянием даже малой корреляции между наблюдениями, не проявляющимся при умеренном числе наблюдений [84, 85]. В связи с этим может быть поставлена задача оптимального по некоторому критерию распределения моментов наблюдений вдоль заданного интервала. Целесообразность постановки такой задачи обусловлена следующими обстоятельствами.

Во-первых, за последние годы существенно увеличилось число спутников, наблюдаемых лазерными дальномерами, в результате чего над многими наблюдательными станциями одновременно можно наблюдать несколько спутников. Вследствие этого возникает проблема назначения приоритетов наблюдаемых объектов. Весьма важным критерием при выборе таких приоритетов может оказаться принадлежность пролета или его части оптимальному плану наблюдений.

Во-вторых, в ближайшие годы планируется трансформация сети лазерных станций наблюдений, в результате которой часть станций будет перемещена на новые места. Одним из показателей, которыми следует руководствоваться при выборе новых пунктов размещения станций может являться соответствие этих пунктов оптимальному плану.

В-третьих, оптимальный план может использоваться как дополнительная информация при обработке уже имеющегося наблюдательного материала, например, при назначении весов наблюдениям.

Решение оптимальной задачи существенно зависит от допущений относительно статистических характеристик ошибок наблюдений. Классическим допущением является некоррелированность всех измерений. Тогда при линейной модели наблюдений оптимальная задача рассматривается в *теории планирования эксперимента* [33, 95]. *Гарантирующий подход* развивается в ряде работ (см., например, [84, 85, 43, 18, 49]). При этом подходе предполагается известным лишь некоторое множество возможных значений статистических характеристик ошибок наблюдений (например, математических ожиданий и ковариаций ошибок наблюдений) и оптимизируется состав измерений при наиболее неблагоприятном по данному критерию значениям этих характеристик. В некоторых случаях оптимальные моменты для этих двух подходов близки или полностью совпадают (см. ниже исторический комментарий).

Здесь мы рассматриваем решение задачи в рамках теории планирования эксперимента, считая возможную корреляцию между измерениями несущественной при используемом умеренном числе наблюдений. При этом будем использовать новые алгоритмы планирования эксперимента [21] для достоверного определения координат полюса Земли  $x_p$  и  $y_p$  по результатам независимых наблюдений вдоль орбиты спутника. Рассмотрим модель оценивания.

**Модель оценивания и ее линеаризация.** Современная теория вращения Земли позволяет прогнозировать значения координат полюса на интервалах времени порядка 1 года с точностью порядка 0.01 секунды дуги, а неравномерность вращения Земли, характеризующуюся величиной UT1-UTC (см. [24]), — с



точностью порядка 0.001 секунды. Следовательно, значения поправок к прогнозируемым величинам не превосходит 10 процентов этих величин. Таким образом, в данной задаче вполне оправдана линеаризация исходных уравнений наблюдений. В [24] дается подробный вывод линеаризованных уравнений наблюдений относительно поправок к координатам полюса, к величине, характеризующей неравномерность вращения Земли UT1-UTC, к координатам спутника и к составляющим вектора его скорости. Следует отметить, что имеются в виду поправки к значениям уточняемых величин в момент времени, соответствующий середине обрабатываемой дуги орбиты спутника (для спутника ЛАГЕОС, как правило, длина дуги составляет 5 суток). Таким образом, для каждой дуги орбиты в момент наблюдения  $t$  уравнения наблюдений могут быть записаны в линеаризованном относительно поправок виде:

$$\varrho_o(t) - \varrho_c(t, \Theta^0) = \nabla(\varrho_c(t, \Theta^0))(\Theta - \Theta^0), \quad (5.1)$$

где  $\varrho_o(t)$  и  $\varrho_c(t, \Theta^0)$  — соответственно наблюдаемое и расчетное значения дальности до спутника,  $\Theta$  — вектор неизвестных параметров размерности  $m$  (в матричных операциях все вводимые векторы считаются столбцами),  $\Theta^0$  — некоторое номинальное (или расчетное) значение этого вектора,  $\nabla(\varrho_c(t, \Theta))$  — градиент (вектор-строка) от дальности по вектору  $\Theta$ . В настоящей статье мы рассматриваем случай когда вектор  $\Theta$  состоит из координат и скоростей спутника в некоторый начальный момент времени и координат полюса  $x_p$  и  $y_p$  на заданную эпоху, т.е. полагаем  $m = 8$ , а значения ряда других параметров, определяющих траекторию спутника (таких как гравитационная постоянная Земли, коэффициенты разложения потенциала Земли и т.п.) принимаем известными достаточно точно, чтобы не учитывать их изменение в линеаризованной модели (5.1).

В следующих трех подразделах дана математическая постановка и изложена методика решения задачи оптимизации, которая в последнем разделе апробирована на задаче определения оптимальных планов при уточнении координат полюса по лазерным наблюдениям спутника ЛАГЕОС.

## 4.5.2. Математическая постановка задачи

Через  $\theta$  обозначим вектор с компонентами  $(\Theta_j - \Theta_j^0)/a_j$ , где  $a_j$  — нормировочные множители, значения которых выбираются в соответствии с физическим смыслом задачи (см. ниже замечание 1). Пусть заданы  $n$  моментов времени (или точек дуги орбиты), в каждом из которых могут повторяться наблюдения с заданного измерительного пункта (на практике эти моменты интерпретируются как нормальные места, в окрестности которых должны проводиться наблюдения). Тогда линеаризованная модель наблюдений (5.1) может быть записана в виде

$$y_{ik} = H_i^T \theta + \varepsilon_{ik}, \quad k = 1, \dots, r_i, \quad i = 1, \dots, n,$$

где  $r_i$  — число повторений наблюдений в каждой точке дуги,  $y_{ik}$  — наблюдаемые относительно номинальной траектории значения,  $H_i$  — заданные векторы,  $\varepsilon_{ik}$  — ошибки наблюдений. Будем полагать, что эти ошибки есть некоррелированные между собой при всех  $i$  и  $k$  случайные величины, математические ожидания которых равны нулю, а дисперсия ошибки  $\varepsilon_{ik}$  равна  $\sigma_i^2$  при любом  $k$ . Более общий случай, когда возможна корреляция между величинами  $\varepsilon_{ik}$  при заданном  $i$  и различных  $k$ , легко сводится к рассматриваемому [21]. Отметим также, что может быть без всяких осложнений рассмотрена более общая модель наблюдений, в которой  $y_{ik}$  — векторы произвольного размера (например, размера два, если одновременно измеряются дальность и радиальная скорость спутника) [21].

**Замечание 1.** Нам представляется естественным следующий выбор нормировочных множителей  $a_j$ , введенных выше. Положим равным их дисперсиям  $D\hat{\Theta}_j$ , вычисленным при условии, что все  $r_i$ ,  $i = 1, \dots, n$ , равны друг другу, т.е. суммарное число наблюдений равномерно распределено по нормальным местам. При таком выборе постоянных  $a_j$  мы можем видеть изменение точности каждого контролируемого параметра при решении задачи (1.5).

### 4.5.3. Проверка алгоритмов на примере определения координат полюса Земли

Изложенные в двух предыдущих разделах алгоритмы реализованы в соответствующей компьютерной программе. В качестве исследуемого наблюдательного материала использовались наблюдения спутника ЛАГЕОС, произведенные на семи пятисуточных дугах. Использование нормирующих множителей, описанных в разделе 4.5.3, обеспечивает повышенную наглядность результатов, так как в этом случае при равномерном распределении моментов наблюдений дисперсии всех параметров равны 1.

В процессе вычислений было выявлено, что решать задачу (1.5) целесообразней всего на интервалах равным 1 суткам. Оказалось, что планы  $L$ - и  $MV$ -задач как правило совпадают, а полученные значения дисперсий отличаются не более, чем на 10%. Поскольку решение  $MV$ -задач требует существенно большего времени вычислений, целесообразно пользоваться  $L$ -критерием. Используя решение задачи (1.5), можно уменьшить дисперсии оценок координат полюса приблизительно в 2 раза.

В качестве исследуемого наблюдательного материала использовались нормальные точки спутника ЛАГЕОС на интервале времени 08.02.1994 - 15.03.1994, распределение которых по станциям представлено таблицей 1.

Таблица 1

Распределение по станциям обрабатываемых наблюдений.

Номер	Название	Число наблюдений
7090	Yaragade	440
7835	Grasse	403
7530	Bar Giyyora	27
7105	Washington	283
7109	Quincy	227
7097	Easter Island	6
7840	Herstmonseux	328
7810	Zimmervald	82
7545	Cagliari	159
8834	Wetzell	430
7839	Graz	276
7110	Monument Peak	341
1884	Riga	137
7403	Arequipa	20
7837	Shanghai	31
7838	Simosato	8
7210	Maui	316
7939	Matera	34
7882	Cabo san lucas	21
7811	Borovec	42
7080	Fort Davis	103

Оказалось, что варьируя число контролируемых параметров  $s$ , можно обеспечить улучшение оценок не только контролируемых параметров, но и остальных. Исследовались следующие варианты значений числа контролируемых параметров:  $x_p$  или  $y_p$  ( $s = 1$ );  $x_p$  и  $y_p$  ( $s = 2$ ); 3 составляющих вектора скорости спутника ( $s = 3$ ), координаты полюса и составляющие вектора скорости спутника ( $s = 5$ ); координаты и составляющие вектора скорости спутника ( $s = 6$ ), все перечисленные параметры ( $s = 8$ ). Результаты вычислений для одной из дуг, соответствующие перечисленным вариантам, приведены в таблице 2.

Таблица 2

Значения среднеквадратических ошибок определяемых параметров при различном числе контролируемых параметров.

$s$	$x$	$y$	$z$	$V_x$	$V_y$	$V_z$	$x_p$	$y_p$
1	22.0	28.9	24.9	7.0	27.0	19.5	0.4	7.0
1	17.7	13.5	13.6	13.1	14.7	16.1	12.1	0.6
2	1.4	1.2	1.2	0.8	1.3	1.2	0.5	0.7
3	1.9	0.5	0.7	0.5	0.5	0.5	1.1	1.8
5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
6	0.6	0.5	0.6	0.6	0.6	0.6	1.1	0.8

В случае  $s = 1$  достигалось наименьшее значение дисперсии этого параметра (одной из координат полюса), однако дисперсия другой координаты, как правило, была такой же как в случае  $s = 2$ , а дисперсии координат спутника достигали 10. В случае  $s = 2$  дисперсии координат полюса принимали значения в диапазоне 0.5-0.7, а дисперсии координат спутника превосходили 1. При  $s = 3$  и  $s = 6$  дисперсии координат полюса превосходили 1, что делало это решение неприемлемым. В случае  $s = 5$  все дисперсии оказывались меньшими 1, причем дисперсии координат полюса в этом случае мало отличались от значений, имевших место при  $s = 2$ . При  $s = 8$  время вычислений неприемливо велико. Таким образом, установив  $s = 5$ , можно получить оптимальный план измерений за минимальное время вычислений.

Анализ результатов вычислений позволяет сделать вывод о том, что пролеты, содержащие базисные точки для случая  $s = 1$  содержат базисные точки,

соответствующие случаю  $s = 2$ , в то же время пролеты, содержащие базисные точки, соответствующие случаю  $s = 2$ , так же содержат точки базиса, соответствующего  $s = 5$ . Типичный пример указанного рода представлен таблицами 3 — 5. Действительно, точки 1, 64, 110, 113, 120 являются общими для всех трех вариантов, пары точек 11 и 10, 30 и 31, 32 и 33 принадлежат соответственно одним и тем же пролетам. Точка 43 (см. табл. 5) дополняет множество пролетов, соответствующее варианту  $s = 2$  до множества пролетов, соответствующего  $s = 5$ .

Таблица 3

Оптимальный базис на интервале 49405.5-49406.5 MJD для случая 1  
контролируемого параметра  $x_p$

$n$	$t$	$dt$	Станция	$p$
1	49405.51674311	0.00000000	7835	0.02689
11	49405.52986577	0.01312266	7835	0.06237
30	49405.61704183	0.10029872	7090	0.27737
32	49405.66137965	0.14463654	7835	0.15113
64	49405.71788675	0.20114364	7110	0.06113
110	49405.82591678	0.30917366	1884	0.17099
113	49405.82727205	0.31052894	7545	0.07512
120	49406.10088920	0.58414609	7403	0.17499

Наибольшие веса получили наблюдения следующих станций  
7835, 7090, 7810, 7105, 7403, 8834, 7838, 7840, 7545, 7109, 7210.

Таким образом, можно сделать вывод о том, что для определения координат полюса на некотором интервале времени следует использовать не все возможные наблюдения, а только те которые соответствуют оптимальному базису. Имея заранее информацию о структуре оптимального базиса можно освободить некоторые станции от наблюдений исследуемого спутника (в данном случае ЛАГЕОСа), назначив наивысший приоритет этому спутнику для станций соответствующих оптимальному базису.

Таблица 4

Оптимальный базис на интервале 49405.5-49406.5 MJD для случая 2  
контролируемых параметров  $x_p$  и  $y_p$

$n$	$t$	$dt$	Станция	$p$
1	49405.51674311	0.00000000	7835	0.15539
10	49405.52916438	0.01242127	7835	0.02485
12	49405.53382421	0.01708110	7105	0.09437
31	49405.61756961	0.10082650	7090	0.21264
32	49405.66137965	0.14463654	7835	0.14254
47	49405.69302788	0.17628477	7105	0.06136
64	49405.71788675	0.20114364	7110	0.04105
110	49405.82591678	0.30917366	1884	0.04324
113	49405.82727205	0.31052894	7545	0.09282
114	49406.09216235	0.57541924	7403	0.05022
120	49406.10088920	0.58414609	7403	0.08152

Таблица 5

Оптимальный базис на интервале 49405.5-49406.5 MJD для случая 5  
контролируемых параметров  $V_x, V_y, V_z, x_p, y_p$

$n$	$t$	$dt$	Станция	$p$
1	49405.51674311	0.00000000	7835	0.16526
11	49405.52986577	0.01312266	7835	0.07056
12	49405.53382421	0.01708110	7105	0.06875
31	49405.61756961	0.10082650	7090	0.14255
33	49405.66257872	0.14583561	7835	0.17556
43	49405.68807653	0.17133342	7110	0.00263
47	49405.69302788	0.17628477	7105	0.05739
64	49405.71788675	0.20114364	7110	0.00735
110	49405.82591678	0.30917366	1884	0.04678
113	49405.82727205	0.31052894	7545	0.09658
114	49406.09216235	0.57541924	7403	0.08866
120	49406.10088920	0.58414609	7403	0.07794

#### 4.5.4. Заключение

Подводя итог изложенному выше, можно заключить, что использование предложенной оптимизации позволяет уменьшить дисперсию определяемых значений координат полюса на 30-40%. Применяв предложенную методику, координирующий центр получает возможность оптимальным образом планировать работу сети станций.

Кроме того, предлагаемая методика может оказаться весьма эффективной при планировании и обработке иных космических экспериментов, таких как оценивание геодинимических параметров путем совместной обработки наблюдений нескольких спутников, обработка и согласование наблюдений одного спутника, выполненных на различной технике, обработка альтиметрических измерений и т. д.



## Глава 5

# О возможности решения проблемы моментов методами линейного программирования и применение к задачам робастного и минимаксного оценивания

Рассматриваются условия, при которых задача минимизации нормы при линейных ограничениях (проблема моментов) может быть эффективно решена методами линейного программирования. В качестве приложения рассмотрены задача минимаксного оценивания при наличии немоделируемых возмущений, задача робастного оценивания при неопределенности закона распределения ошибок, ряд задач минимаксного оценивания при неточно заданной ковариационной матрице ошибок измерений. Предлагается единый подход, позволяющий свести их к задачам обобщенного линейного программирования и эффективно решать их обычным симплекс-методом с некоторой модификацией при проверке условия оптимизации.

## 5.1. Введение.

Различные задачи робастного и оптимального оценивания параметров системы сводятся к так называемой проблеме моментов [39, 67] следующего вида

$$p(x^*) = \min_x \{p(x) : Ax \geq b\}. \quad (1.1)$$

Здесь  $x \in \mathbb{R}^n$ ,  $b \in \mathbb{R}^m$ ,  $A$  —  $(m \times n)$ -матрица,  $p(x)$  — некоторая норма (или полунорма, т.е.  $p(x)$  может равняться нулю при  $x \neq 0$ ) вектора  $x$ ,  $m \leq n$ . Здесь и далее верхним индексом "\*" помечены оптимальные значения переменных. При этом запись (1.1) включает в себя и случай, когда ограничения (или часть из них) имеют вид равенств

$$Ax = b,$$

так как эти равенства эквивалентны неравенствам  $Ax \leq b$ ,  $Ax \geq b$ . Для задачи (1.1) можно найти точное решение лишь в некоторых частных случаях. Мы покажем, что задачи определенного класса могут быть эффективно решены методами линейного программирования. Такими являются, в частности, рассмотренные ниже задачи получения робастных нелинейных оценок и ряд задач минимаксного оценивания при неточно заданных характеристиках ошибок измерений и возмущений в действующих на систему силах.

Мы рассмотрим три случая.

1. **Минимизация опорной функции.** Для некоторых задач функция  $p(x)$  может быть представлена как опорная функция некоторого симметричного относительно нуля выпуклого компакта  $\Gamma$  (см., например, задачу оценивания с немоделируемыми возмущениями из раздела 5.2.2), т.е.

$$p(x) = \max \{x^T \gamma : \gamma \in \Gamma\} = x^T \gamma(x), \quad (1.2)$$

и явно или достаточно просто находится функция  $\gamma(x)$ , реализующая оптимум в (1.2).

2. **Минимизация максимума квадратичной функции.** Для задач робастного оценивания уже не удастся найти соотношение вида (1.2). Однако имеет

место аналогичная более сложная зависимость (см. формулу (3.3) с квадратичной целевой функцией от вектора  $\gamma$ ), позволяющая применить с некоторыми изменениями теорию, развитую в разделе 5.2 и найти эффективный алгоритм решения задачи робастного оценивания (см. раздел 5.3).

**3. Случай известной двойственной нормы функции  $p(x)$ .** Известно [38], что в (1.2) можно принять

$$\Gamma = \Gamma^* \doteq \{\gamma : p^*(\gamma) \leq 1\}, \quad (1.3)$$

где

$$p^*(\gamma) \doteq \max\{\gamma^T \alpha : p(\alpha) = 1\} \quad (1.4)$$

– норма, двойственная к  $p(x)$ . Если  $p(x)$  – полунорма, то в случае существования максимума в (1.4)  $p^*(\gamma)$  также будем называть двойственной нормой. Однако явный вид множества  $\Gamma^*$  не всегда известен. Для ряда задач (см., например, задачи минимаксного оценивания при неточно заданной ковариационной матрице ошибок измерений из раздела 5.4.2) можно найти множество  $\Gamma = \Gamma^*$ . (что позволяет существенно улучшить алгоритм при  $m \ll n$ ).

## **5.2. Решение проблемы моментов в случае минимизации опорной функции и применение к задаче минимаксного оценивания при наличии немоделируемых возмущений.**

### **5.2.1. Обоснование алгоритма.**

Результаты этого раздела получены путем использования и обобщения идей из [28]. Это позволяет нам решать более общие, чем в [28] задачи.

ЛЕММА 5.1. Пусть имеет место соотношение (1.2). Тогда задача (1.1) сводится к обобщенной задаче линейного программирования

$$p(x^*) = \max_{\lambda, \mu_j, \gamma^j} \left\{ b^T \lambda : \sum_{i=1}^m \lambda_i \begin{pmatrix} A^i \\ 0 \end{pmatrix} + \sum_j \mu_j \begin{pmatrix} -\gamma^j \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \lambda^T = (\lambda_1, \dots, \lambda_m) \geq 0, \mu_j \geq 0 \right\}, \quad (2.1)$$

в которой  $A^i$  – столбцы матрицы  $A^T$ , а векторы условий  $\gamma^j \in \mathbb{R}^n$  могут выбираться произвольно из условия  $\gamma^j \in \Gamma$  (при этом число слагаемых по  $j$  в (2.1) можно считать не превосходящим  $n + 1$ ).

Доказательство леммы 5.1 приведено в разделе 5.5.

Задачу (2.1) можно формально рассматривать как обычную задачу линейного программирования с бесконечным числом векторов условий. Для нее может быть обоснован метод генерации столбцов [31, 57, 58], состоящий в применении симплексного метода для решения этой задачи. При этом проверка условия оптимальности и поиск вектора, вводимого в базис, приводит уже к перебору бесконечного числа векторов из  $\Gamma$ , т.е. к вспомогательной оптимальной задаче. Если эта вспомогательная задача решается эффективно (например, аналитически), то симплексный алгоритм так же эффективен для задачи (2.1), как и для обычной задачи линейного программирования. Мы покажем, что вспомогательная задача имеет вид (1.2), т.е. имеет по предположению простое решение. Отметим, что для задачи (2.1) не удастся, как правило, найти решение за конечное число шагов [18]. Однако мы имеем на каждом шаге оценку близости к оптимальному решению (см. ниже теорему 5.1).

Обоснуем симплексный алгоритм подробно. Для его описания введем понятия, аналогичные используемым в обычной задаче линейного программирования (см. раздел 3.2).

Решение обобщенной задачи линейного программирования (2.1) достигается как и в обычной задаче линейного программирования на допустимом базисном решении [31, 57, 58, 18, 9]. Под ним мы будем подразумевать переменные  $\lambda_i$  ( $i = 1, \dots, k$ ),  $\mu_j$ , ( $j = 1, \dots, s + 1 - k$ ) и соответствующие переменным  $\mu_j$  векторы  $\gamma^j$ , которые удовлетворяют следующим условиям: выполнены ограничения

этой задачи и не вырождена базисная матрица  $B$ , столбцами которой являются соответствующие векторы ограничений (нумерация здесь и далее условна). Если все переменные  $\lambda_i, \mu_j$  положительны, то это решение – *невырожденное*, в противном случае – *вырожденное*.

Обозначим через  $c_B^T = (b_1, \dots, b_k, 0, \dots, 0)$  – вектор коэффициентов целевой функции, соответствующий текущему базисному решению задачи (2.1). Определим, следуя [31, 59]  $(n + 1)$ -вектор  $\Pi$  множителей Лагранжа, т.е.  $n$ -вектор  $\pi$  и скаляр  $\rho$  из системы

$$\begin{aligned} c_B^T - \Pi^T B &= 0 \Leftrightarrow b_i - \pi^T A^i = 0, \quad i = 1, \dots, k, \\ \pi^T \gamma^j - \rho &= 0, \quad j = 1, \dots, s + 1 - k. \end{aligned} \quad (2.2)$$

**ЗАМЕЧАНИЕ 5.1.** Можно считать, что векторы  $A^i$  удовлетворяют условиям

$$b_i - \pi^T A^i \leq 0, \quad i = k + 1, \dots, m. \quad (2.3)$$

Действительно, если зафиксировать векторы  $\gamma^j$ , то задача (2.1) становится обычной задачей линейного программирования. Решив ее симплекс-методом за конечное число шагов, мы удовлетворим достаточным условиям оптимальности, которые и есть условия (2.3) [31, 59].

Если (2.3) выполнено, то достаточные условия оптимальности для задачи (2.1) имеют вид [57, 58]:

$$-\min_{\gamma \in \Gamma} (-\gamma^T, 1) \Pi = 0 \Leftrightarrow \max_{\gamma \in \Gamma} \pi^T \gamma = \rho. \quad (2.4)$$

Если это условие не выполняется, а вектор  $\gamma(\pi)$  реализует максимум в (1.2) при  $x = \pi$ , то при вводе в базис по симплексным правилам вектора  $(\gamma(\pi)^T, 1)$  целевая функция задачи (2.1) не уменьшается (в случае невырожденного допустимого решения – увеличивается). При этом оценку близости текущего решения к оптимальному дает следующая теорема.

**ТЕОРЕМА 5.1.** *Предположим, что алгоритм генерации столбцов для задачи (2.1) модернизирован в соответствии с замечанием 5.1, т.е. для  $\pi$  из*

(2.2) выполняются условия (2.3). Тогда:

1. Для оптимального решения  $p(x^*)$  задачи (1.1) имеет место оценка:

$$\rho \leq p(x^*) \leq p(\pi). \quad (2.5)$$

2. Если  $p(\pi) = \rho$ , то оптимальное решение задачи (1.1) есть  $x^* = \pi$ , а ее значение равно  $p(x^*) = \rho$ .

Доказательство теоремы 5.1 дано в разделе 5.5.

Таким образом, для проверки условия оптимальности и ввода в базис вектора условий нужно уметь решать подзадачу нахождения вектора  $\gamma(\pi)$ , которая по сделанному выше предположению о функции  $p(x)$  решается эффективно (например, аналитически).

**ЗАМЕЧАНИЕ 5.2.** Отметим некоторые существенные отличия решения обобщенной и обычной задач линейного программирования на примере задачи (2.1).

1. В отличие от обычной задачи линейного программирования достаточное условие оптимальности (2.4) может вообще не выполняться ни на одном базисе [18]. Условие (2.5) позволяет найти почти оптимальный базис с заданной точностью.
2. На каждом шаге симплекс-метода целевая функция в (2.1) увеличивается на величину  $\theta(p(\pi) - \rho)$ , где

$$\theta = \min_{\alpha_i > 0} \frac{\mu_i}{\alpha_i},$$

$\alpha_i$  – коэффициенты разложения вводимого в базис вектора по базису. При этом, если

$$p(\pi) - \rho \rightarrow 0 \text{ при } n \rightarrow \infty \quad (2.6)$$

(а это и происходит во всех встречавшихся нам задачах), то алгоритм сходится согласно неравенству (2.5). Теоретически возможен случай, когда условие

(2.6) не выполняется. Тогда из изложенного в этом замечании нетрудно вывести, что должно выполняться условие (см. также [18])

$$\theta \rightarrow 0 \text{ при } n \rightarrow \infty.$$

Это условие означает, что происходит стремление к вырожденному базисному решению. В [9] (см. раздел 3.2) предложен алгоритм, позволяющий избежать вырожденных итераций (т.е. итераций, на которых целевая функция не изменяется) и изменить значение целевой функции путем ввода в базис сразу нескольких векторов.

3. Если не учитывать трудности нахождения вектора  $\gamma(\pi)$ , то по трудоемкости вычислений задача (2.1) эквивалентна обычной задаче линейного программирования с  $n + 1$  условиями-равенствами и  $n + 2$  переменными.

**ЗАМЕЧАНИЕ 5.3.** Ряд задач оптимального оценивания, некоторые из которых рассмотрены ниже, сводятся к задаче вида

$$p(x^*) = \min_x \{p(x) : Ax = b\}. \quad (2.7)$$

к решению которой применимы приведенные выше результаты со следующими упрощениями. В задаче (2.1) на переменные  $\lambda_i$  нет ограничений на знак. Поэтому эти переменные можно всегда считать базисными [59]. Это означает, во-первых, что фактически размерность базисной матрицы можно считать равной  $n + 1 - m$ . К тому же выводу можно прийти, исключив переменные  $\lambda_i$  из уравнений-равенств в (2.1). Во-вторых, модернизация алгоритма, указанная в условии теоремы 5.1, не требуется, так как всегда выполняются условия (2.3).

### **5.2.2. Решение минимаксной задачи с немоделируемыми возмущениями.**

Пусть движение системы описывается дифференциальными уравнениями

$$\dot{q}(t) = R(t)q(t) + F(t)u(t), \quad t \in [0, T], \quad (2.8)$$

где  $q(t)$  – текущий  $m$ – вектор состояния,  $u(t)$  –  $r$ – вектор возмущений, значение которого мало, но неизвестно. Неизвестную функцию  $u(t)$  будем называть *немоделируемым возмущением*. Мы будем полагать известными лишь ограничения

$$p(u(t)) \leq U(t),$$

где  $p(\cdot)$  означает некоторую норму, а  $U(t)$  – заданная функция.

Приведем два вида функции  $p(\cdot)$ , соответствующих неучтенной составляющей  $u(t)$  при коррекции траектории космического аппарата. Вид этой функции зависит от способа коррекции (см., например, [18]). В межпланетных перелетах обычно проводится коррекция с использованием одного двигателя после соответствующей ориентации космического аппарата в пространстве. Этому способу соответствует модельная функция затрат

$$p_i(x) = \|x\| \quad (\text{евклидова норма}). \quad (2.9)$$

При коррекции искусственных спутников Земли обычно коррекция проводится при помощи шести двигателей, направленных по положительным и отрицательным направлениям трех осей координат при сохранении постоянной ориентации космического аппарата. При этом модельная функция затрат равна

$$p_i(x) = \|x\|_1 \doteq \sum_{i=1}^{n_i} |x_i|. \quad (2.10)$$

Пусть уравнения измерений имеют вид

$$z_i = h(t_i)^T q(t_i) + \varepsilon_i, \quad i = 1, \dots, n, \quad t_1 \leq t_2 \leq \dots \leq t_n, \quad (2.11)$$

где  $h(t_i)$  – известные векторы, а  $\varepsilon_i$  – ошибки измерений. Пусть далее нужно оценить параметр

$$l = a^T q(\tau)$$

( $a$  – известный вектор,  $\tau$  – заданный момент времени) при помощи линейного



несмещенного алгоритма [39, 96]

$$\hat{l} = \sum_{i=1}^n x_i z_i, \quad (2.12)$$

где условия несмещенности имеют вид:  $\hat{l} = l$ , если  $\varepsilon_i = 0 \forall i$ . Решение уравнения (2.8) в форме Коши есть

$$q(t) = Q(t)[\theta + s(t)], \quad \theta \doteq q(0), \quad s(t) \doteq \int_0^t \Psi(t)u(t) dt, \quad \Psi(t) \doteq Q^{-1}(t)F(t),$$

$Q(t)$  – фундаментальная матрица  $m \times m$ , определяемая из системы дифференциальных уравнений в вариациях:

$$\dot{Q}(t) = R(t)Q(t), \quad Q(0) = I_m$$

(здесь  $I_m$  – единичная матрица  $m \times m$ ). С учетом этого решения модель измерений (2.11) записывается в виде схемы линейной регрессии

$$z_i = H(t_i)^T \theta + \gamma_i, \quad i = 1, \dots, n, \quad (2.13)$$

где  $\gamma_i = \varepsilon_i + h(t_i)s(t_i)$  – суммарная ошибка измерений и модели,  $H(t) = Q(t)^T h(t)$ . При этом условия несмещенности алгоритма оценивания имеют вид

$$\sum_{i=1}^n x_i H_i = b, \quad b \doteq Q(\tau)^T a. \quad (2.14)$$

Отсюда можно получить выражение для ошибки оценивания [18]:

$$\hat{l} - l = \sum_{i=1}^n x_i \varepsilon_i + \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \Psi(t)^T Y_j(t) u(t) dt + \int_t^T \Psi(t)^T u(t) dt, \quad (2.15)$$

где

$$Y_j(t) \doteq \sum_{i=j}^n x_i H(t_i) - bI(\tau - t),$$

$$I(y) \doteq \begin{cases} 0, & y \leq 0, \\ 1, & y > 0 \end{cases}$$

– функция Хевисайда.

В минимаксной задаче с немоделируемыми возмущениями предполагается, что ошибки измерений  $\varepsilon_i$  ограничены по модулю значениями  $\sigma_i$ , а оптимальный вектор  $x^*$  находится из условия [18], [28], [48]

$$p(x^*) = \min_x \max_{u(t), \varepsilon_i} \left\{ \hat{l} - l : |\varepsilon_i| \leq \sigma_i, p(u(t)) \leq U(t) \right\}, \quad (2.16)$$

где ограничения на вектор  $x = (x_1, \dots, x_n)^T$  есть условия несмещенности (2.14).

Соотношение (2.15) есть линейная функция вектора  $x$ . Поэтому задача (2.16) имеет вид задачи (2.7), где вектор  $\gamma$  выражается через немоделируемое возмущение и ошибки измерения из формулы (2.13).

Отсюда следует, что

$$\begin{aligned} \max_{u(t), \varepsilon_i} \left\{ \hat{l} - l : |\varepsilon_i| \leq \sigma_i, p(u(t)) \leq U(t) \right\} = & \sum_{i=1}^n \sigma(t_i) |x_i| + \\ & + \sum_{j=1}^n \int_{t_{j-1}}^{t_j} U(t) p^*(\Psi(t)^T Y_j(t)) dt + \int_{t_n}^T U(t) p^*(\Psi(t)^T b) dt, \end{aligned}$$

где символом  $p^*(\cdot)$  обозначена двойственная норма. Эта норма в случае (2.9) также является евклидовой, т.е.  $\|x\|^* = \|x\|$ . В случае (2.10) двойственная норма равна максимальному модулю  $|x_r|$  компонент вектора  $x$ . В обоих указанных случаях явно находятся  $\varepsilon_i$  и  $u(t)$ , т.е. вектор  $\gamma$ , доставляющий максимум в подзадаче (1.2) находится явными вычислениями. Таким образом, в этих случаях полностью выполняются условия теоремы 5.1 и может быть применен симплексный алгоритм решения задачи (2.1), столь же эффективный как и решение обычной задачи линейного программирования той же размерности. Примеры решения этой задачи для случая (2.10) приведены в [28], [48].

## 5.3. О решении задачи робастного оценивания.

### 5.3.1. Задача робастного оценивания.

Рассмотрим линейную модель измерений

$$z_i = H_i^T \theta + \xi_i, \quad i = 1, \dots, n,$$

где  $z_i$  и  $\xi_i$  – результаты и ошибки измерений соответственно,  $H_i \in \mathbb{R}^s$  – известные векторы,  $\theta \in \mathbb{R}^s$  – неизвестный вектор, относительно которого известны только ограничения

$$A\theta \geq b, \tag{3.1}$$

где  $b$  – заданный  $m$ -вектор,  $A$  –  $m \times s$  матрица. Если, например, ограничения на компоненты  $\theta_j$  вектора  $\theta$  имеют естественный вид

$$|\theta_j| \leq c_j, \quad j = 1, \dots, s,$$

то это соответствует следующим параметрам в (3.1)

$$m = 2s, \quad A = \begin{pmatrix} I \\ -I \end{pmatrix}, \quad b^T = (-c_1, \dots, -c_s, -c_1, \dots, -c_s),$$

где  $I$  – единичная матрица размера  $s \times s$ . Предположим, что ошибки измерений есть независимые между собой случайные величины, функция распределения каждой из которых представляется в виде выпуклой комбинации двух функций распределения:

$$F_\xi(x) = (1 - \varepsilon)\Phi(x) + \varepsilon G(x),$$

где  $\varepsilon$  – некоторое положительное число, не большее единицы,  $\Phi(x)$  – функция стандартного нормального распределения, вид функции  $G(x)$  распределения вероятности неизвестен. Здесь  $\varepsilon$  можно интерпретировать как вероятность того, что ошибка измерений распределена не по нормальному закону. Для такой модели в предположении, что  $n$  достаточно велико и может быть применен закон

больших чисел, предлагается устойчивый к возможно большим ошибкам измерений (робастный) метод оценивания (так называемый  $M$ -метод) [42, 32, 69, 66], впервые обоснованный Хьюбером [91] для рассматриваемого нами случая. В этом методе оценка  $\hat{\theta}$  вектора  $\theta$  определяется следующим образом:

$$\hat{\theta} = \arg \min_{\theta} \{F(\theta) : A\theta \geq b\}, \quad (3.2)$$

где

$$F(\theta) = \sum_{i=1}^n f(x_i(\theta)), \quad x_i(\theta) = z_i - H_i^T \theta,$$

функция  $f(x)$  имеет вид

$$f(x) \doteq \begin{cases} x^2/2 & , \quad |x| \leq C, \\ |x|C - C^2/2 & , \quad |x| > C, \end{cases}$$

а величина  $C$  определяется как функция  $\varepsilon$  из соотношения

$$\frac{1}{1-\varepsilon} = \int_{-C}^C \varphi(x) dx + \frac{2}{C} \varphi(C),$$

$$\varphi(x) \doteq \frac{1}{\sqrt{2\pi}} \exp(-x^2/2).$$

### 5.3.2. Решение задачи робастного оценивания с помощью обобщенного линейного программирования.

Функция  $F(\theta)$  представляется, как нетрудно проверить, в виде:

$$F(\theta) = \max_{\gamma} \left\{ x(\theta)^T \gamma - \frac{1}{2} \gamma^T \gamma : |\gamma_i| \leq C, i = 1, \dots, n \right\}, \quad (3.3)$$

где  $x(\theta)$  – вектор с компонентами  $x_i(\theta)$ , а максимум в (3.3) достигается при  $\gamma_i(\theta) \doteq (\gamma_1(\theta), \dots, \gamma_n(\theta))^T$  :

$$\gamma_i(\theta) = \begin{cases} x_i(\theta), & |x_i(\theta)| \leq C, \\ C \operatorname{sign}(x_i(\theta)), & |x_i(\theta)| > C. \end{cases} \quad (3.4)$$

Подставляя (3.3) в (3.2) и используя теорему Неймана–Кнезера [27], переставим операции минимизации по  $\theta$  и максимизации по  $\gamma$ . Тогда  $\hat{\theta}$  определяется путем решения задачи

$$F(\hat{\theta}) = \sup_{\gamma} \inf_{\theta} \left\{ -(H\gamma)^T \theta + z^T \gamma - \frac{1}{2} \gamma^T \gamma : \right. \\ \left. A\theta \geq b, |\gamma_i| \leq C, i = 1, \dots, n \right\}, \quad (3.5)$$

где  $z$  – вектор с компонентами  $z_i$ ,  $H$  – матрица размером  $s \times n$ , столбцами которой являются  $H_i$ . Заменяя в (3.5) минимизацию по  $\theta$  двойственной задачей максимизации, получим:

$$\max_{\gamma, \lambda} \left\{ b^T \lambda + z^T \gamma - \frac{1}{2} \gamma^T \gamma : A^T \lambda + H\gamma = 0, \lambda \geq 0, |\gamma_i| \leq C \quad i = 1, \dots, n \right\}, \quad (3.6)$$

где  $\lambda$  –  $m$ -мерный вектор.

Чтобы избавиться от нелинейного слагаемого  $\gamma^T \gamma$  в целевой функции и перевести его в ограничения, введем новую скалярную переменную  $\eta$ , удовлетворяющую условию  $C_0 \geq \eta \geq \frac{1}{2} \gamma^T \gamma$ . Очевидно, можно принять  $C_0 = nC^2/2$ . Тогда задача (3.6) переписется в эквивалентном виде:

$$\max_{\gamma, \lambda, \eta} \{ b^T \lambda + z^T \gamma - \eta : A^T \lambda + H\gamma = 0, \lambda \geq 0 \},$$

где переменные  $\gamma$  и  $\eta$  принадлежат ограниченному выпуклому множеству

$$\mathcal{M} = \left\{ \gamma, \eta : C_0 \geq \eta \geq \frac{1}{2} \gamma^T \gamma, |\gamma_i| \leq C, i = 1, \dots, n \right\} \in \mathbb{R}^{n+1}. \quad (3.7)$$

Учитывая это и используя далее те же рассуждения, что и при доказательстве леммы 5.1, получаем следующий вывод.

**ЛЕММА 5.2 (АНАЛОГ ЛЕММЫ 5.1).** *Задача (3.2) сводится к обобщенной задаче линейного программирования с  $s + 1$  ограничениями- равенствами:*

$$F(\hat{\theta}) = \max_{\lambda, \mu_j, \gamma^j, \eta_j} \left\{ b^T \lambda + \sum_j (z^T \gamma^j - \eta_j) \mu_j \right\} \quad (3.8)$$

при условиях

$$\sum_{i=1}^m \lambda_i \begin{pmatrix} A^i \\ 0 \end{pmatrix} + \sum_j \mu_j \begin{pmatrix} H\gamma^j \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \lambda \geq 0, \quad \mu_j \geq 0, \quad \begin{pmatrix} \gamma^j \\ \eta_j \end{pmatrix} \in \mathcal{M}.$$

Здесь  $A^i$  – столбцы матрицы  $A^T$ , а векторы условий  $(\gamma^j, \eta_j)^T$  могут выбираться произвольно из множества  $\mathcal{M}$  (при этом число слагаемых по  $j$  в (3.8) можно считать не превосходящим  $n + 1$ ).

Отметим, что, в отличие от задачи (2.1), в задаче (3.8) не только векторы условий могут выбираться из определенного множества, но и коэффициенты целевой функции при  $\mu_j$  линейно зависят от выбора этих векторов условий.

Обозначим через  $c_B^T = (b_1, \dots, b_k, z^T \gamma^1 - \eta_1, \dots, z^T \gamma^{s+1-k} - \eta_{s+1-k})$  вектор коэффициентов целевой функции, соответствующий текущему базисному решению задачи (3.8) (здесь нумерация условна,  $0 \leq k \leq s$ ). Определим  $(s + 1)$ -вектор  $\Pi$  множителей Лагранжа, т.е.  $m$ -вектор  $\pi$  и скаляр  $\rho$  из системы вида (2.2):

$$c_B^T - \Pi^T B = 0 \Leftrightarrow \begin{cases} b_i - \pi^T A^i = 0, & i = 1, \dots, k, \\ (z - H^T \pi)^T \gamma^j - \eta_j - \rho = 0, & j = 1, \dots, s - k + 1. \end{cases} \quad (3.9)$$

**ЗАМЕЧАНИЕ 5.4.** Для задачи (3.8) также справедливо замечание 5.1 из раздела 5.2.1. Это означает, что можно путем указанной в замечании 5.1 модернизации симплексного алгоритма добиться на каждой итерации выполнения условий (2.3).

Если (2.3) выполнено, то достаточные условия оптимальности для задачи (3.8) с учетом (3.3) имеют вид [57, 58]:

$$\begin{aligned} 0 &= \min_{\gamma, \eta} \left\{ -z^T \gamma + \eta + \pi^T H \gamma : \gamma \in \Gamma, \eta \geq \frac{1}{2} \gamma^T \gamma \right\} = \\ &= \rho - \max_{\gamma \in \Gamma} \left\{ (z^T - \pi^T H) \gamma - \frac{1}{2} \gamma^T \gamma \right\} = \rho - F(\pi) \end{aligned} \quad (3.10)$$

Если (3.10) не выполняется, а вектор  $\gamma(\pi)$ , вычисленный по формуле (3.4), реализует максимум в (3.3) при  $\theta = \pi$ , то при вводе в базис по симплексным правилам вектора  $((H\gamma(\pi))^T, 1)$  целевая функция задачи (3.8) не уменьшается (в случае невырожденного допустимого решения – увеличивается). При этом справедлив результат, аналогичный теореме 5.1.

**ТЕОРЕМА 5.2 (АНАЛОГ ТЕОРЕМЫ 5.1).** *Предположим, что алгоритм генерации столбцов для задачи (3.8) модернизирован в соответствии с замечанием 5.4, т.е. для  $\pi$  из (3.9) выполняются условия (2.3). Тогда:*

1. Для оптимального решения  $F(\hat{\theta})$  задачи (3.2) имеет место оценка:

$$\rho \leq F(\hat{\theta}) \leq F(\pi). \quad (3.11)$$

2. Если  $F(\pi) = \rho$ , то оптимальное решение задачи (3.2) есть  $\hat{\theta} = \pi$ , а ее значение равно  $F(\hat{\theta}) = \rho$ .

Доказательство теоремы 5.2 дано в разделе 5.5.

Таким образом, проверка условия оптимальности осуществляется аналитически. Если не учитывать затраты на аналитические расчеты вектора  $\gamma(\pi)$ , то по трудоемкости вычислений задача (3.8) эквивалентна обычной задаче линейного программирования с  $s + 1$  условиями-равенствами и  $s + 2$  переменными. При этом условие (3.11) позволяет найти почти оптимальный базис с заданной точностью. Отметим также, что пункты 1 и 2 замечания 5.2 также имеют место для задачи (3.8).

Можно предположить, что предлагаемый симплексный алгоритм (ниже в таблице – СА) будет тем эффективнее по сравнению с итеративным методом (ниже – ИМ), использующем на каждой итерации метод наименьших квадратов с некоторыми весами, чем больше величины  $s$ ,  $\varepsilon$  и  $m$ . Это предположение оправдывается как показывает расчеты. Для полиномиальной регрессии при моделировании неизвестного распределения равномерным с большой дисперсией даже при отсутствии ограничений на оцениваемые коэффициенты полинома ( $m = 0$ ) при  $s \geq 5$  предлагаемый алгоритм эффективнее итерационного метода примерно в 2 раза при различных значениях  $\varepsilon \geq 0.01$ . Некоторые результаты приведены в таблице для  $s = 7$  при различных значениях  $\varepsilon$  и разных реализациях результатов измерений. Неизвестное распределение  $G(x)$  моделируется как равномерное.

Таблица 1.

Оценки вектора  $(10, 70, -40, 30, -10, 4, 9)$ , полученные на выборках, состоящих из 250 элементов, при различных параметрах шума.

$$\varepsilon = 0,05 \quad R[-10, 10]$$

метод	оценка							число итераций
ИМ	10,27	70,09	-40,44	30,06	-9.81	4.35	8.12	14
СА	9,91	70,31	-40,02	29,99	-10.21	4.23	9.13	9

$$\varepsilon = 0,01 \quad R[-100, 100]$$

метод	оценка							число итераций
ИМ	10,36	70,50	-40,38	29,63	-9.89	3.98	8.62	24
СА	10,32	69,40	-40,13	30,01	-10.22	4.15	9.37	12

#### 5.4. Решение проблемы моментов в случае известной двойственной нормы и применение к задаче минимизации гарантированной дисперсии.

Предлагаемый в разделе 5.2 алгоритм оперирует с матрицами порядка  $n + 1$  и при больших  $n$  оказывается трудоемким. В задачах минимаксного оценивания обычно  $m \ll n$ . В следующем частном, но достаточно распространенном случае, мы предлагаем симплексный алгоритм, оперирующий с матрицами порядка  $m$ .

Этот случай имеет различные приложения к задачам минимаксного оценивания, некоторые из которых рассмотрены ниже.



### 5.4.1. Обоснование алгоритма.

Введением новых переменных  $u \in \mathbb{R}^1, \alpha \in \mathbb{R}^n$  из условий

$$x = u\alpha, \quad p(\alpha) = 1, \quad u \geq 0$$

задача (1.1) сводится к обобщенной задаче линейного программирования

$$p(x^*) = \min_{u \geq 0, a \in \mathcal{A}} \{u : ua \geq b\}; \quad \mathcal{A} \doteq \{a : a = A\alpha, p(\alpha) = 1\}. \quad (4.1)$$

Согласно дополнению к теореме Каратеодори (см. теорему 1.3 Каратеодори-Малоземова из [33]) каждая граничная точка компакта может быть представлена как выпуклая комбинация неизвестных  $n$  точек из выпуклой оболочки компакта. Поэтому вектор  $a$  в (4.1) может быть представлен в виде выпуклой линейной комбинации

$$a = \sum_j \mu_j a_j, \quad a_j \in \text{conv} \mathcal{A}, \quad \sum_j \mu_j = 1, \quad \mu_j \geq 0,$$

где  $\text{conv} \mathcal{A} = \{a : a = A\alpha, p(\alpha) \leq 1\}$  – выпуклая оболочка множества  $\mathcal{A}$ . Вводя переменные  $y_j = u\mu_j$  и  $z_i \geq 0$ , приводим задачу (4.1) к виду

$$p(x^*) = \min_{y_j, z_i} \left\{ \sum_j y_j : \sum_j y_j a_j - \sum_{i=1}^m z_i e_i = b, \quad a_j \in \mathcal{A}, \quad y_j \geq 0, \quad z_i \geq 0 \right\}, \quad (4.2)$$

где  $e_i$  – единичные векторы. В (4.2) формально должно стоять условие  $a_j \in \text{conv} \mathcal{A}$ , которое заменено равносильным условием  $a_j \in \mathcal{A}$ .

Обозначим через  $c_B^T = (1, \dots, 1, 0, \dots, 0)$  – вектор коэффициентов целевой функции, соответствующий текущей базисной матрице  $B$  задачи (4.2) (число единиц равно числу  $k$  положительных переменных  $y_j$  в базисном решении). Определим  $m$ -вектор  $\pi$  множителей Лагранжа из системы [31, 59]

$$c_B^T - \pi^T B = 0 \Leftrightarrow 1 - \pi^T a_i = 0, \quad i = 1, \dots, k, \quad \pi^T e_i = 0, \quad j = k + 1, \dots, m.$$

**ЗАМЕЧАНИЕ 5.5.** Используя те же соображения, что и в замечании 5.1, так модернизируем симплексный алгоритм, что текущее базисное решение будет

удовлетворять условиям

$$\pi^T e^i \geq 0, \quad i = 1, \dots, m. \Leftrightarrow \pi \geq 0, \quad (4.3)$$

Для этого при фиксированных  $a_j$  базисные векторы  $e_i$  нужно выбирать в задаче (4.2) оптимально путем решения обычной задачи линейного программирования.

При условии (4.3) достаточное условие оптимальности имеет вид [57, 58]

$$\min_{a_j \in \mathcal{A}} (1 - a_j^T \pi) \Leftrightarrow 1 - p^*(A^T \pi) = 0. \quad (4.4)$$

Если на текущем шаге  $1 - p^*(A^T \pi) < 0$  и вектор  $\alpha$  доставляет экстремум в (1.4) при  $\gamma = A^T \pi$ , то при введении в базис вектора  $A\alpha$  целевая функция задачи (4.2) не уменьшается (в случае невырожденного допустимого решения – увеличивается).

Достаточное условие (4.4) может, вообще говоря, никогда не выполниться при проведении симплексных итераций [18]. Однако всегда может быть найдено решение, сколь угодно близкое к оптимальному, как показывает следующая теорема.

**ТЕОРЕМА 5.3.** *Предположим, что алгоритм генерации столбцов модернизирован в соответствии с замечанием 5.5, т.е. выполняются условия (4.3). Тогда если*

$$1 - p^*(A^T \pi) > -\varepsilon \quad (\varepsilon > 0),$$

*то*

$$\sum_j y_j \geq p(x^*) > \sum_j y_j / (1 + \varepsilon). \quad (4.5)$$

Доказательство теоремы 5.3 дано в разделе 5.5.

Теорема 5.3 дает достаточные условия почти оптимальности и оценку близости текущего решения к оптимальному. Если не учитывать трудности нахождения двойственной нормы (1.4), то по трудоемкости вычислений задача (4.2) эквивалентна обычной задаче линейного программирования с  $m$  условиями-равенствами и  $m + 1$  переменными.

**ЗАМЕЧАНИЕ 5.6.** Итак, мы приходим к следующему выводу. Задача (1.1) сводится к обобщенной задаче линейного программирования размерности  $m$  (обычно  $m \ll n$ ). При этом на каждой итерации подзадача поиска вводимого в базис вектора сводится к нахождению двойственной нормы  $p^*(\gamma)$  при  $\gamma = A^T \pi$ . Если эта подзадача решается сравнительно легко, то эффективен предлагаемый алгоритм решения задачи (1.1). Покажем, что подзадача 1.4 сводится к задаче вида (1.1), но с одним ограничением-равенством. Действительно, введем новую переменную  $z$  из условий

$$\alpha = z/p(z), \quad \gamma^T z = 1.$$

Тогда задача (1.4) записывается в виде

$$\frac{1}{p^*(\gamma)} = \min \{p(z) : \gamma^T z = 1\}. \quad (4.6)$$

**ЗАМЕЧАНИЕ 5.7 (АНАЛОГ ЗАМЕЧАНИЯ 5.3).** Если рассматривается задача (2.7) вместо задачи (1.1), то в задаче (4.2) переменные  $z_i$  отсутствуют и теорема 5.3 справедлива без каких-либо допущений.

### 5.4.2. Минимизация гарантированной дисперсии.

Рассмотрим ошибку несмещенного линейного алгоритма из пункта 5.2.2 при отсутствии немоделируемых возмущений. Согласно (2.15) она имеет вид

$$\hat{l} - l = \sum_{i=1}^n x_i \varepsilon_i,$$

где компоненты оценителя  $x_i$  удовлетворяют условию несмещенности (2.14). Пусть нам известно лишь множество  $\mathcal{K}$ , которому принадлежит матрица ковариаций ошибок измерений. Примем

$$p(x)^2 = D_{\text{гар}} \doteq \max_{K \in \mathcal{K}} D(\hat{l} - l),$$

т.е.  $p(x)^2$  – гарантированная дисперсия оценки (здесь  $D(\cdot)$  – знак дисперсии). Рассмотрим два случая.

Допустим, что множество  $\mathcal{K}$  определяется условиями

$$|k_{ij}| \leq k \leq 1 \quad (i \neq j),$$

где  $k_{ij}$  — коэффициенты корреляции между ошибками измерений, число  $k$  задано, а дисперсии ошибок известны и для простоты записи приняты далее единицами. Кроме того, следует наложить необходимое для корреляционной матрицы условие неотрицательной определенности. Тогда [18]

$$p(x)^2 = (1 - k) \sum_{i=1}^n x_i^2 + k \left( \sum_{i=1}^n |x_i| \right)^2.$$

При этом функция  $p^*(x)$  в (4.6) находится аналитически [20] и поэтому может быть применена теорема 5.3 для решения задачи (2.7).

Рассмотрим другой важный для практики случай. Пусть вектор ошибок измерений  $\varepsilon$  разбивается на  $s$  некоррелированных между собой подвекторов и допускается произвольная корреляция между ошибками внутри каждого подвектора. Тогда при допущении, что ошибка каждого измерения имеет единичную дисперсию, гарантированная дисперсия оценки есть

$$p(x)^2 = D_{\text{гар}} = \sum_{r=1}^s \left( \sum_i |x_i^r| \right)^2,$$

где  $x_i^r$  — компоненты вектора оценивателя  $x$ , соответствующие подвектору измерений с номером  $r$ . Для этого случая функция  $p^*(\gamma)$  в (4.6) находится аналитически. Действительно, зафиксировав оптимальными все переменные, кроме переменных  $z_i^1$ , получаем, что в задаче (2.7) имеется решение, в котором не равно нулю может быть только одно некоторое переменное  $z_{n(1)}^1$  (так как оптимальные переменные  $z_i^1$ , находятся из решения задачи линейного программирования с одним ограничением-равенством). Поэтому задача (4.6) приводится к задаче

$$\frac{1}{p^*(\gamma)^2} = \min_{n(r), z_{n(r)}} \left\{ \sum_{r=1}^s (z_{n(r)}^r)^2 : \sum_{r=1}^s z_{n(r)}^r \gamma_{n(r)} = 1 \right\}, \quad (4.7)$$

где номера  $n(r)$  и величины  $z_{n(r)}$  подлежат определению, а  $\gamma_{n(r)}$  — соответствующие компоненты вектора  $\gamma = A^T \pi$ . Задача (4.7) имеет следующее решение при

фиксированных  $n(r)$  [96]

$$z_{n(r)} = \frac{\gamma_{n(r)}}{p^*(\gamma)}, \quad p^*(\gamma) = \sum_{r=1}^s \gamma_{n(r)}^2,$$

откуда определяются оптимальные для задачи (4.7) значения

$$n(r) = \arg \max_i \{|\gamma_i^r|\}.$$

Таким образом, теорема 5.3 применима и для этого случая.

## 5.5. Приложение.

*Доказательство леммы 5.1.* Подставим (1.2) в (1.1), переставим операции минимума и максимума (это возможно согласно теореме Неймана–Кнезера [27]) и учтем, что согласно теории двойственности

$$\inf_x \{x^T \gamma : Ax \geq b\} = \max_{\lambda} \{b^T \lambda : A^T \lambda = \gamma, \lambda \geq 0, \lambda \in \mathbb{R}^m\}.$$

Тогда получим

$$p(x^*) = \max_{\lambda, \gamma} \{b^T \lambda : A^T \lambda = \gamma, \gamma \in \Gamma, \lambda \geq 0\}. \quad (5.1)$$

Очевидно, максимизацию в предыдущем соотношении можно проводить лишь по граничным точкам множества  $\Gamma$ . Согласно дополнению к теореме Каратеодори (см. теорему 1.3 Каратеодори–Малоземова из [33]) каждая граничная точка компакта может быть представлена как выпуклая комбинация неизвестных  $n$  точек из выпуклой оболочки компакта. Поэтому, учитывая выпуклость множества  $\Gamma$ , получаем

$$\gamma = \sum_j \mu_j \gamma^j, \quad \gamma^j \in \Gamma, \quad \sum_j \mu_j = 1, \quad \mu_j \geq 0.$$

Тогда задача (5.1) максимизации по  $\lambda$  запишется в виде задачи (2.1). Лемма доказана.

*Доказательство теоремы 5.1.* Пусть  $\lambda^r \doteq (\lambda_1, \dots, \lambda_k, 0, \dots, 0)$  – текущее базисное решение задачи (2.1), соответствующее базису (2.2), а  $\gamma \doteq \sum_{j=1}^{n+1-k} \mu_j \gamma^j$ .

Из условий (2.1) следует, что  $A^T \lambda = \gamma$  и  $\sum_{j=1}^{n+1-k} \mu_j = 1$ . С учетом этого из (2.2) вытекают соотношения

$$b^T \lambda = \pi^T A^T \lambda = \pi^T \gamma = \pi^T \sum_{j=1}^{n+1-k} \mu_j \gamma^j = \rho.$$

Пусть  $\lambda^*$  – решение задачи (2.1). С учетом предыдущих равенств очевидно, что  $b^T \lambda^* \geq b^T \lambda = \rho$ . Кроме того, из (2.2) и (2.3) следует, что  $\pi$  – допустимый вектор для задачи (1.1). Поэтому  $p(\pi) \geq b^T \lambda^*$ , поскольку оптимальные значения задач (1.1) и (2.1) совпадают. Таким образом неравенство (2.5) доказано.

Если  $p(\pi) = \rho$ , то из (2.5) следует, что  $\rho = b^T \lambda^*$ , а  $\pi$  – решение задачи (1.1). Теорема 5.1 доказана.

*Доказательство теоремы 5.2.* Пусть  $\lambda^T \doteq (\lambda_1, \dots, \lambda_k, 0, \dots, 0)$  – текущее базисное решение задачи (3.8), соответствующее базису (3.9) и

$$\gamma \doteq \sum_{j=1}^{n+1-k} \mu_j \gamma^j, \quad \eta \doteq \sum_{j=1}^{n+1-k} \mu_j \eta_j.$$

Тогда из (3.9) следует  $\pi^T A^T \lambda = b^T \lambda$ , а из (3.8) получим  $A^T \lambda = -H \gamma$ . Отсюда, с учетом (3.9), вытекает, что целевая функция в (3.8) равна  $b^T \lambda + z^T \gamma - \eta = \rho$ . Далее доказательство совпадает с рассуждениями в теореме 5.1.

*Доказательство теоремы 5.3.* Для того, чтобы доказать неравенство (4.5), запишем соотношение между оптимальными и текущими значениями целевой функции в задаче (4.3) [59]:

$$p(x^*) = \sum_j y_j + \sum_j (1 - \pi^T a^j) y_j^* + \sum_{i=1}^m (-\pi^T e_i) z_i,$$

где звездочкой помечены оптимальные переменные, а вектор  $\pi$  соответствует текущему базисному решению. Отсюда с учетом формул (4.3), (4.4) получаем неравенство

$$p(x^*) \geq \sum_j y_j + (1 - p^*(A^T \pi)) \sum_j y_j^*,$$

где  $p^*(\cdot)$  означает, как и ранее, двойственную норму. Теперь неравенство (4.5) получается путем использования последнего соотношения, соотношения  $\sum_j y_j^* = p(x^*)$  и условия  $1 - p^*(A^T \pi) > -\varepsilon$ . Теорема доказана.

## Глава 6

# Об оптимальной линейной идеальной коррекции при ограничениях на корректирующие импульсы

Рассматривается задача линейной идеальной коррекции параметров системы при достаточно общих ограничениях на корректирующие импульсы. Показывается, что эта задача сводится к обобщенной задаче линейного программирования, которая в важных частных случаях может быть решена столь же эффективно, как и обычная задача линейного программирования. Приводятся примеры коррекции орбиты искусственного спутника Земли, демонстрирующие эффективность предлагаемого метода. Для них показывается, что если ограничить величины корректирующих импульсов одной константой, то при уменьшении этой константы почти до ее минимально возможного значения не происходит значительного увеличения суммарного расхода топлива, а увеличивается лишь количество импульсов. Это означает возможность проведения непрерывной коррекции с двигателем малой тяги, столь же эффективной как и импульсной коррекции, которая моделирует коррекцию с двигателем большой тяги на малом интервале времени.

## 6.1. Постановка задачи и невозможность ее решения методом, используемым для задачи без ограничений на импульсы

### 6.1.1. Постановка задачи

Рассмотрим задачу обобщенной линейной импульсной коррекции [44, 18], считая ее идеальной, т.е. не имеющей ошибок исполнения импульсов. Пусть  $b \in \mathbb{R}^m$  — корректируемый вектор,  $u_i \in \mathbb{R}^{n_i}$  — корректирующие импульсы,  $B_i$  —  $(m \times n_i)$ – матрицы влияния  $i$ –го импульса на корректируемый вектор,  $i = 1, \dots, n$ . Величина  $n_i$  есть число независимых параметров коррекции.

Рассматривая далее в качестве приложения импульсную коррекцию космического аппарата, заметим, что если может быть реализовано любое направление импульса в трехмерном пространстве, то  $n_i = 3$ . Если задана плоскость или прямая, которым должен принадлежать этот импульс, то  $n_i$  равно двум или единице.

В линейной идеальной коррекции импульсы должны находиться из условия

$$\sum_{i=1}^n B_i u_i = b. \quad (1.1)$$

Кроме этого будем предполагать, что

$$u_i \in \mathcal{U}_i, \quad i = 1, \dots, n, \quad (1.2)$$

где  $\mathcal{U}_i$  — замкнутые выпуклые множества.

Пусть неубывающие положительные функции  $p_i(u_i)$  характеризуют энергетические затраты на исполнение  $i$ –го импульса. Здесь мы будем считать для определенности, что функция  $p_i(x)$  есть некоторая норма вектора  $x = (x_1, \dots, x_{n_i})^T$ . При коррекции траектории космического аппарата вид этой функции зависит от способа коррекции (см., например, [18]). В межпланетных перелетах обычно проводится коррекция с использованием одного двигателя после соответствующей ориентации космического аппарата в пространстве. Этому способу соответствует модельная функция затрат



$$p_i(x) = \|x\| \quad (\text{евклидова норма}). \quad (1.3)$$

При коррекции искусственных спутников Земли обычно коррекция проводится при помощи шести двигателей, направленных по положительным и отрицательным направлениям трех осей координат при сохранении постоянной ориентации космического аппарата. При этом модельная функция затрат равна

$$p_i(x) = \|x\|_1 \doteq \sum_{i=1}^{n_i} |x_i|. \quad (1.4)$$

Отметим, что при  $n_i = 1$  функции (1.3) и (1.4) совпадают.

Оптимальная задача обобщенной линейной импульсной коррекции может быть сформулирована в виде

$$L = \min_{u_i} \left\{ \sum_{i=1}^n p_i(u_i) : \sum_{i=1}^n B_i u_i = b, \quad u_i \in \mathcal{U}_i \right\}. \quad (1.5)$$

В случае, когда все функции затрат имеют вид (1.4) и есть только ограничения сверху на эти затраты в каждый момент времени, задача (1.5) решается методами линейного программирования [101]. Однако, как показано ниже, уже при наличии функций затрат вида (1.3) решение задачи так просто не получается.

В случае, когда нет дополнительных ограничений (1.2) на импульсы, заменой переменных

$$u_i = x_i \gamma_i, \quad x_i = p_i(u_i), \quad \gamma_i \in \mathbb{R}^{n_i}, \quad p_i(\gamma_i) = 1$$

задача (1.5) сводится к следующей задаче [44, 18]:

$$L = \min_{x_i, a_i} \left\{ \sum_{i=1}^n x_i : \sum_{i=1}^n x_i a_i = b, \quad x_i \geq 0, \quad a_i \in \mathcal{A}_i, \quad i = 1, \dots, n \right\}, \quad (1.6)$$

где

$$\mathcal{A}_i = \{a \in \mathbb{R}^m : a = B_i \gamma, \quad p_i(\gamma) \leq 1\}. \quad (1.7)$$

В (1.7) формально должно стоять условие  $p_i(\gamma) = 1$ . Однако расширение множества изменения векторов условий  $a_i$  не меняет значения задачи (1.6). Это

следует из того, что каждому оптимальному  $x_i > 0$  всегда соответствует оптимальный вектор  $a_i$ , принадлежащий границе множества (1.7) (так как замена  $a_i$  на  $ga_i$ , где  $0 < g < 1$ , приводит к увеличению целевой функции). Граница множества  $\mathcal{A}_i$  есть эллипсоид размерности не большей  $m$ , который является линейным образом множества  $p_i(\gamma) = 1$  (единичной окружности в норме  $p_i(\cdot)$ ) из  $\mathbb{R}^{n_i}$  в  $\mathbb{R}^m$ . Задача (1.6) называется *обобщенной задачей линейного программирования* и отличается от обычной задачи линейного программирования тем, что каждый вектор условий  $a_i$  не задан, а может быть произвольно выбран из заданного выпуклого множества (см. раздел 3.2 и [31, 57]).

Задачу (1.6) можно формально рассматривать как обычную задачу линейного программирования с бесконечным числом векторов условий, заполняющих множества  $\mathcal{A}_i$ . Исходя из этого, а также из геометрических соображений, иллюстрируемых ниже на примере, в [44] был предложен симплексный алгоритм решения этой задачи. Фактически, этот алгоритм является применением метода *генерации столбцов* для обобщенной задачи линейного программирования к задаче (1.6), который рассматривается в [31, 57]. Более полное обоснование этого метода для задачи (1.6) рассматривается в [18, 9]. Для описания этого алгоритма введем понятия, аналогичные используемым в обычном линейном программировании (см. раздел 3.2).

Назовем *допустимым решением* задачи (1.6) переменные  $x_i$ ,  $i = 1, \dots, n$ , и соответствующие векторы  $a_i$ , удовлетворяющие ограничениям этой задачи. Пусть положительными являются только переменные  $x_1, \dots, x_k$ . Если  $k \leq m$ , то это решение называется *допустимым базисным решением*. Этому решению соответствует оптимальное решение задачи (1.5), содержащее не более  $m$  импульсов. Если  $k = m$ , то это решение — *невырожденное*, в противном случае — *вырожденное*. Любые  $m$  линейно независимых векторов  $a_1, \dots, a_m$  (нумерация векторов условна), включающих векторы допустимого базисного решения  $a_1, \dots, a_k$ , образуют *базисную матрицу  $B$  (базис)*. Невырожденному допустимому базисному решению соответствует только один базис, а вырожденному допустимому базисному решению может соответствовать некоторое множество базисов.

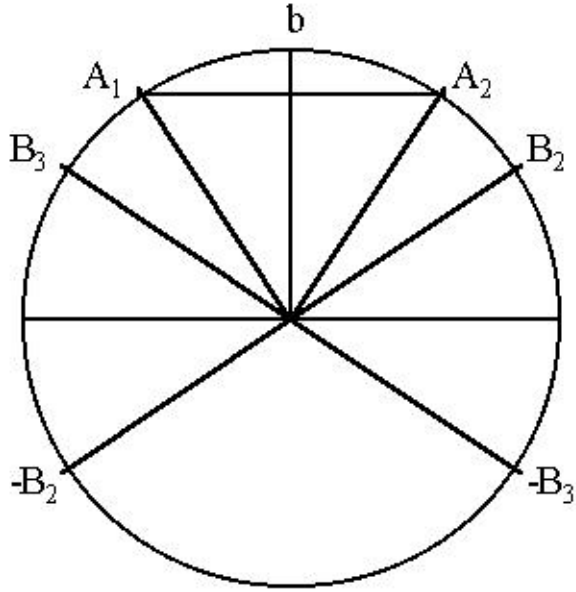


Рис.6. Работа метода ОЛИК для модельной задачи с ограничениями.

Будем пояснять введенные понятия и дальнейшие результаты на следующем примере. Рассмотрим случай (1.3) евклидовой нормы для функции затрат на  $i$ -ю коррекцию. Пусть

$$n = 3, \quad n_1 = 2, \quad n_2 = n_3 = 1,$$

$$B_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}, \quad B_3 = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (1.8)$$

Тогда (рис. 6) единичная окружность – граница множества  $\mathcal{A}_1$ , точки  $(B_2, -B_2)$  и  $(B_3, -B_3)$  — относительные границы множеств  $\mathcal{A}_2$  и  $\mathcal{A}_3$  соответственно. Базис  $(a_1, B_2)$  соответствуют невырожденному допустимому базисному решению, а базис  $(b, B_2)$  — вырожденному.

Согласно [18], существует оптимальное допустимое базисное решение, а также имеет место критерий оптимальности Куна-Таккера.

ТЕОРЕМА 6.1. *Допустимое базисное решение задачи (1.6) является оптимальным тогда и только тогда, когда найдется  $m$ -вектор  $\pi$  такой, что*

$$1 - \pi^T a_i = 0, \quad i = 1, \dots, k, \quad (1.9)$$

$$1 - \pi^T a \geq 0 \quad \forall a \in \bigcup_{i=1}^n \mathcal{A}_i. \quad (1.10)$$

Для примера (1.8) оптимальными являются базисы  $(b, \pm B_1)$  и  $(b, \pm B_2)$ , которым соответствует одно оптимальное вырожденное допустимое базисное решение  $x_1 = 1$  с соответствующим вектором условий  $b$ . Это следует из того, что условия теоремы 6.1 выполняются при  $\pi = (0, 1)^T$ . В случае вырожденного допустимого базисного решения из теоремы 6.1 не вытекает однозначного алгоритма нахождения вектора  $\pi$ . Поэтому для описываемого метода генерации столбцов теоретически можно было бы использовать следующее достаточное условие оптимальности (раздел 3.2):

ТЕОРЕМА 6.2. *Допустимое базисное решение является оптимальным, если выполняются условия (2.2) (сравни с (1.11)) :*

$$1 - \pi^T a \geq 0 \quad \forall a \in \bigcup_{i=1}^n \mathcal{A}_i, \quad (1.11)$$

где вектор  $\pi$  определяется из уравнений

$$1 - \pi^T a_i = 0, \quad i = 1, \dots, m \quad \Leftrightarrow \quad \pi^T B = e^T, \quad (1.12)$$

в которых базисная матрица соответствует этому допустимому базисному решению.

Далее мы покажем, что это достаточное условие нужно заменить на более реалистичное.

### 6.1.2. Геометрическое описание алгоритма решения задачи (1.6)

Теорема 6.2 аналогична соответствующему достаточному условию оптимальности в обычной задаче линейного программирования. Однако, здесь есть существенное отличие. В обычной задаче линейного программирования всегда существует базис, удовлетворяющий этим условиям. В обобщенной задаче линейного программирования такого базиса может не быть (см. раздел 3.2). Действительно, в рассматриваемом примере наличие такого базиса означает, что существуют две точки на окружности такие, что выше прямой, проходящей через них, нет других точек окружности [44] (так как выше прямой не выполняется условие (1.12)). Но такого базиса нет, хотя, как указывалось выше, любой базис  $(b, a)$  является оптимальным, если точка  $a$  лежит в круге. Однако, согласно ([18]), существует множество базисов, для которых в задаче (1.6) вместо условия (1.11) выполняется достаточное условие  $\varepsilon$ -оптимальности:

$$1 - \pi^T a \geq -\varepsilon \quad \forall a \in \bigcup_{i=1}^n \mathcal{A}_i, \quad (1.13)$$

где  $\varepsilon > 0$  — заданное число. Один из таких базисов ищется симплекс-методом, геометрическая интерпретация которого на данном примере следующая. Если взять, например, базис  $(B_1, B_2)$ , то выше прямой, проходящей через концы векторов  $B_1$  и  $B_2$  окажутся точки окружности, каждая из которых может быть введена в базис и целевая функция уменьшится. Вводя наиболее возвышающуюся точку  $b = (0, 1)$  вместо точки  $B_1$ , получим оптимальный базис  $(b, B_2)$ . Однако этот базис не удовлетворяет условию (1.12). Мы можем удовлетворить лишь условию (1.13) при любом малом  $\varepsilon$ , проводя далее описанную итерацию симплекс-метода многократно. При этом каждая итерация будет *вырожденной*, т.е. не меняющей значения целевой функции.

### 6.1.3. Аналитическое описание алгоритма решения задачи (1.6) и его обоснование

Представим в формульном виде описанный выше геометрически метод генерации столбцов. Запишем достаточные условия  $\varepsilon$ -оптимальности в виде

$$\Delta \geq -\varepsilon, \quad (1.14)$$

где

$$\Delta = \min \left\{ 1 - \pi^T a : a \in \bigcup_{i=1}^n \mathcal{A}_i \right\} = 1 - \beta, \quad (1.15)$$

$$\beta = \max_i \{ \beta_i : i = 1, \dots, n \},$$

$$\begin{aligned} \beta_i = \max \{ \pi^T a : a \in \mathcal{A}_i \} = \\ = \max \{ \pi^T B_i \gamma : p(\gamma) = 1 \} = p_i^*(B_i^T \pi), \end{aligned} \quad (1.16)$$

где функция

$$p_i^*(x) = \max \left\{ \frac{x^T u}{p_i(u)} : u \in \mathbb{R}^n \right\} = \max \{ x^T u^0 : p_i(u^0) = 1 \} \quad (1.17)$$

— так называемая двойственная норма. Эта норма в случае (1.3) также является евклидовой, т.е.  $\|x\|^* = \|x\|$  и она достигается при  $u^0 = \|x\|/x$ , при этом  $\beta_i = \|B_i^T \pi\|$ . В случае (1.4) двойственная норма равна максимальному модулю  $|x_r|$  компонент вектора  $x$  и достигается, когда  $r$ -я компонента вектора  $u_0$  равна  $\text{sign}(x_r)$ , а остальные компоненты равны нулю.

Если условие (1.14) не выполняется, то в текущий базис вводится тот вектор  $a$ , на котором достигается оптимум в (1.15). При этом согласно разделу 3.2 существует полезная оценка близости текущего значения  $L(x) = \sum_{i=1}^n x_i$  целевой функции задачи (1.6) и ее оптимального значения  $L$  (см. также [44]):

$$\frac{L(x)}{1 + \Delta} \leq L \leq L(x). \quad (1.18)$$

Предлагаемый алгоритм может привести к базисному решению, которое не является допустимым для задачи (1.6), так как содержит несколько векторов из одного множества (1.7), с положительными коэффициентами в условиях задачи (1.6). Это связано с тем, что фактически решается вспомогательная обобщенная задача линейного программирования [18, 9]

$$L = \min_{\mathcal{F}_i, x_i(a)} \left\{ \sum_{i=1}^n \sum_{a \in \mathcal{F}_i} x_i(a) : \sum_{i=1}^n \sum_{a \in \mathcal{F}_i} x_i(a)a = b \right\}, \quad (1.19)$$

где  $\mathcal{F}_i$  — совокупность конечного числа крайних точек множества  $\mathcal{A}_i$ . При этом значения задач (1.6) и (1.19) совпадают, а решение задачи (1.6) получается из решения задачи (1.19) по формулам

$$x_i = \sum_{a \in \mathcal{F}_i} x_i(a),$$

$$a_i = \sum_{a \in \mathcal{F}_i} \lambda_i(a)a, \quad (1.20)$$

где

$$\lambda_i(a) = x_i(a)/x_i.$$

Задача (1.19) получается из задачи (1.6) преобразованием (1.20), где

$$\sum_{a \in \mathcal{F}_i} \lambda_i(a) = 1, \quad (1.21)$$

и введением переменных

$$x_i(a) = \lambda_i(a)x_i. \quad (1.22)$$

Преобразование (1.20) при условии (1.21) есть запись теоремы, согласно которой выпуклая оболочка крайних точек выпуклого множества совпадает с этим множеством. Нетрудно видеть, что в (1.19) каждое множество  $\mathcal{F}_i$  можно считать состоящими из точек множества  $\mathcal{A}_i$ , так как любую выпуклую комбинацию крайних точек выпуклого множества можно заменить на одну точку этого множества.

Отметим, что если множество  $\mathcal{A}_i$  не содержит прямолинейных участков (например это имеет место в случае (1.3), когда это множество — эллипсоид), то решение задачи (1.19), удовлетворяющее условиям оптимальности (1.11) может содержать не более одного вектора из  $\mathcal{A}_i$  с соответствующей положительной компонентой  $x_i$ . В этом случае полученный базис является допустимым для задачи (1.6). Если же множество  $\mathcal{A}_i$  содержит прямолинейные участки (как в случае (1.4)), то возможно, что в базисе окажется, к примеру, два вектора  $a_{i1}, a_{i2} \in \mathcal{A}_i$  с положительными коэффициентами  $x_{i1}, x_{i2}$  в разложении вектора  $b$ . В этом случае базис не будет допустимым для задачи (1.6). Однако вектор  $a_{i1}$  может быть заменен на выпуклую комбинацию (1.20):

$$a_i = \lambda a_{i1} + (1 - \lambda) a_{i2} \in \mathcal{A}_i,$$

$$\lambda \doteq x_{i1} / (x_{i1} + x_{i2}).$$

Новый базис дает то же значение целевой функции, но из множества  $\mathcal{A}_i$  содержит уже только один вектор  $a_i$  с положительным коэффициентом  $x_i = x_{i1} + x_{i2}$ . Поэтому он будет допустимым.

Таким образом, в случае отсутствия ограничений на импульсы метод генерации столбцов сводится к обычным симплексным итерациям, на каждой из которых для ввода в базис вектора и проверки достаточного условия оптимальности нужно дополнительно решать подзадачи (1.16), не представляющие затруднений.

**ЗАМЕЧАНИЕ 6.1.** Как мы видели при геометрическом описании алгоритма, наличие вырожденных итераций в обобщенной задаче линейного программирования не является исключительным случаем. Исключить такие итерации позволяет метод, изложенный в разделе 3.2.



#### 6.1.4. Неприменимость описанного алгоритма для задачи обобщенной линейной импульсной коррекции с ограничениями

В отличие от задачи обобщенной линейной импульсной коррекции без ограничений на импульсы при некоторых способах задания множеств  $\mathcal{U}_i$  указанный алгоритм может не привести к допустимому решению задачи, так как преобразование типа (1.20) может нарушать условие (1.2). В описанном выше модельном примере рассмотрим случай, когда имеются ограничения на величины импульсов:  $\|u_i\| \leq c_i, \quad i = 1, 2, 3$ .

Положим  $c_1 = 1/\sqrt{3}, \quad c_2 = c_3 = 1$ . Тогда задача обобщенной линейной импульсной коррекции сводится к задаче (1.6) с дополнительными ограничениями

$$x_i \leq c_i, \quad (1.23)$$

однако преобразование (1.20) с заменой переменных (1.22), вообще говоря, уже не приводит к задаче (1.19) с дополнительными ограничениями (1.23). В связи с этим решение последней задачи может не быть допустимым для исходной задачи обобщенной линейной импульсной коррекции. Действительно, как будет показано в разделе 6.2, для рассматриваемого примера (1.8) в случае (1.23) оптимальная программа коррекции включает три импульса: двумерный вектор  $u_1^T = (0, \quad 1/\sqrt{3})$ , соответствующий матрице влияния  $B_1$ , и два одномерных импульса  $u_2 = u_3 = 1 - 1/\sqrt{3}$ , соответствующих матрицам влияния  $B_2$  и  $B_3$ . Значение целевой функции при этом равно  $2 - 1/\sqrt{3}$ . С другой стороны, для задачи (1.19) с дополнительными ограничениями (1.23) допустимым является решение, которое соответствует базису  $(a_1, a_2)$  и дает значение целевой функции, равное  $1/\sqrt{3}$ , меньшее оптимального значения в исходной задаче (1.5). Следовательно, решение этой задачи не может быть решением задачи (1.5).

Таким образом рассмотренный выше вариант симплекс-метода, вообще говоря, не позволяет решать задачи обобщенной линейной импульсной коррекции с ограничениями на импульсы и утверждения работ [44, 18] о его применимости в

этом случае ошибочны.

## 6.2. Преобразование задачи с ограничениями к обобщенной задаче линейного программирования

Без ограничения общности можно считать, что задача (1.5) решается при дополнительном ограничении

$$p_i(u_i) \leq c_0. \quad (2.1)$$

Здесь постоянную  $c_0$  нужно положить равной значению целевой функции в (1.5) на некотором допустимом решении. Введем дополнительные переменные  $y_i$ ,  $i = 1, \dots, n$ , удовлетворяющие условиям  $y_i \geq p_i(u_i)$  и обозначим

$$N = \sum_{i=1}^n n_i, \quad Y^T = (y_1, \dots, y_n) \in \mathbb{R}^n, \quad U^T = (u_1, \dots, u_n) \in \mathbb{R}^N,$$

$e^T = (1, \dots, 1)$  — вектор из  $n$  единиц,  $B_0 = (B_1, \dots, B_n)$  — матрица размерности  $N \times m$ .

В пространстве  $\mathbb{R}^{N+n}$  определим множество  $\mathcal{M}$  векторов  $(Y^T, U^T)$ :

$$\mathcal{M} = \left\{ (Y^T, U^T) \in \mathbb{R}^{N+n} : \right. \\ \left. y_i \geq p_i(u_i), \quad y_i \leq c_0, \quad u_i \in \mathcal{U}_i, \quad i = 1, \dots, n \right\}.$$

В принятых обозначениях задача (1.5) принимает вид

$$L = \min_{Y, U} \left\{ e^T Y : B_0 U = b, \quad (Y^T, U^T) \in \mathcal{M} \right\}, \quad (2.2)$$

так как решение задачи (2.2) всегда удовлетворяет условию  $y_i = p_i(u_i)$ , при котором задача (1.5) совпадает с задачей (2.2).

Очевидно, что множество  $\mathcal{M}$  выпукло и ограничено. Используем преобразования, аналогичные (1.20):

$$\begin{pmatrix} Y \\ U \end{pmatrix} = \sum_{(Y^T, U^T) \in \mathcal{F}} \lambda(Y, U) \begin{pmatrix} Y \\ U \end{pmatrix}, \quad (2.3)$$

где  $\mathcal{F} \in \mathcal{M}$  — некоторое конечное множество, а переменные  $\lambda(\cdot)$  удовлетворяют условиям

$$\sum_{(Y^T, U^T) \in \mathcal{F}} \lambda(Y, U) = 1, \quad \lambda(Y, U) \geq 0. \quad (2.4)$$

Подставляя (2.3) в (2.2), получим

$$L = \min_{\lambda(Y, U), \mathcal{F} \in \mathcal{M}} \left\{ \sum_{(Y^T, U^T) \in \mathcal{F}} \lambda(Y, U) e^T Y : \sum_{(Y^T, U^T) \in \mathcal{F}} \lambda(Y, U) B_0 U = b, \begin{pmatrix} Y \\ U \end{pmatrix} \in \mathcal{F} \right\}, \quad (2.5)$$

где  $\lambda(Y, U)$  подчиняются условиям (2.4).

Задача (2.5) есть обобщенная задача линейного программирования, в которой целевая функция и ограничения линейны по переменным  $\lambda(Y, U)$ , а коэффициенты целевой функции и вектора условий выбираются из выпуклого множества. Для этой задачи базисная матрица содержит все те векторы условий  $((B_0 U_i)^T, 1)$ ,  $i = 1, \dots, k$ , которые соответствуют положительным компонентам  $\lambda_1, \dots, \lambda_k$ ,  $k \leq m + 1$ , т.е. эта  $(m + 1) \times (m + 1)$ -матрица имеет вид

$$B = \begin{pmatrix} B_0 U_1 & , \dots, & B_0 U_{m+1} \\ 1 & , \dots, & 1 \end{pmatrix}. \quad (2.6)$$

После того, как найдена оптимальная базисная матрица (2.6), оптимальное значение  $U^{*T} = (u_1^{*T}, \dots, u_n^{*T})$  находится из разложения (2.3):

$$U^* = \sum_{i=1}^{m+1} \lambda_i^* U_i^*, \quad (2.7)$$

где  $\lambda_i^*$  — компоненты базисного вектора  $\lambda_B^*$ , соответствующего базисной матрице. Для задачи (2.5) справедливы теоремы 6.1 и 6.2 с очевидными изменениями,

соответствующими виду этой задачи. В частности, условие оптимальности (1.11) имеет вид

$$e^T Y - \pi^T \begin{pmatrix} B_0 U \\ 1 \end{pmatrix} \geq 0 \quad \forall \begin{pmatrix} Y \\ U \end{pmatrix} \in \mathcal{M}, \quad (2.8)$$

а соотношение (1.12) для определения  $(m+1)$ -вектора  $\pi$  есть

$$(e^T Y_1, \dots, e^T Y_{m+1}) - \pi^T B = 0. \quad (2.9)$$

Применим теорему 6.1 для доказательства оптимальности решения задачи (1.6), приведенного в разделе 6.1 для примера (1.8) с ограничениями (1.23). Возьмем в (2.6)

$$Y_1^T = \left(\frac{1}{\sqrt{3}}, 0, 0\right), \quad U_1^T = \left(0, \frac{1}{\sqrt{3}}, 0, 0\right),$$

$$Y_2^T = \left(\frac{1}{\sqrt{3}}, 1, 0\right), \quad U_2^T = \left(0, \frac{1}{\sqrt{3}}, 1, 0\right),$$

$$Y_3^T = \left(\frac{1}{\sqrt{3}}, 1, 1\right), \quad U_3^T = \left(0, \frac{1}{\sqrt{3}}, 1, 1\right).$$

Тогда из соотношений (2.9), (2.6) и условий задачи (2.5) получаем

$$\pi^T = \left(0, 2, -\frac{1}{\sqrt{3}}\right), \quad \lambda_B^T = \left(\frac{1}{\sqrt{3}}, 0, 1 - \frac{1}{\sqrt{3}}\right).$$

Непосредственно проверяется, что для данного  $\pi$  удовлетворяются условия оптимальности (2.8), т.е. полученное решение оптимально и согласно формулам (2.7) оно совпадает с решением, приведенным в разделе 6.1.

## 6.3. Алгоритм поиска оптимального базисного решения

### 6.3.1. Описание алгоритма

Опишем этот алгоритм в соответствии с теорией, изложенной в разделе 6.1. Представим вектор  $\pi$ , определенный из соотношения (1.12) в виде  $\tilde{\pi}^T =$

$(\tilde{\pi}^T, \pi_{m+1})$ , где  $\pi^T \in \mathbb{R}^m$ ,  $\pi_{m+1}$  — скаляр. Правая часть неравенства (2.8) тогда запишется в виде

$$\Delta = \min_{Y, U} \left\{ e^T Y - \tilde{\pi}^T B_0 U - \pi_{m+1} : \begin{pmatrix} Y \\ U \end{pmatrix} \in \mathcal{M} \right\}. \quad (3.1)$$

Если выполняется неравенство (1.14)  $\Delta \geq -\varepsilon$  для величины (3.1), то матрица  $B$  удовлетворяет условию (1.13)  $\varepsilon$ -оптимальности допустимого базисного решения в задаче (2.5). В противном случае в базис вводится вектор  $((B_0 v)^T, 1)$ , на котором достигается минимум в (3.1) и опять проверяется условие (1.13). Вместо оценки (1.18) в задаче (2.5) существует следующая оценка близости текущего значения  $L(\lambda)$  целевой функции в (2.5) к ее оптимальному значению  $L$  [49]:

$$L(\lambda) + \Delta \leq L \leq L(\lambda),$$

где  $L(\lambda)$  — текущее значение целевой функции, а величина  $\Delta$  определена в (3.1). Эта оценка используется для нахождения оптимального решения с заданной точностью.

Согласно разделу 6.1, итерации в описанном алгоритме генерации столбцов будут так же эффективны, как и итерации симплекс-метода в обычной задаче линейного программирования порядка  $m+1$  с  $m+2$  переменными, если подзадача (3.1) решается эффективно.

### 6.3.2. Решение подзадачи (3.1) в некоторых важных частных случаях

Покажем сначала, что эта подзадача решается аналитически в достаточно общем случае, когда имеются только ограничения на величину функций затрат на коррекцию при произвольно, заданных нормах  $p_i(x)$ , т.е. в случае

$$\mathcal{U}_i = \{u_i \in \mathbb{R}^{n_i} : p_i(u_i) \leq c_i, \quad i = 1, \dots, n\}, \quad (3.2)$$

где постоянные  $c_i$  можно считать не превосходящими величины  $c_0$ , определенной в (2.1). В практических приложениях постоянные  $c_i$  выбираются как из условий

ограничений при аппаратной реализации импульсов коррекции, так и из условий достаточно точного описания цели коррекции линейной зависимостью (1.1). В рассматриваемом случае задача (3.1) имеет вид

$$\Delta = \sum_{i=1}^n \delta_i - \pi_{m+1},$$

где

$$\delta_i = \min_{y, u} \{y - a_i^T u : p_i(u) \leq c_i, y_i \geq p_i(u)\},$$

$$a_i^T \doteq \tilde{\pi}^T B_i.$$

Очевидно, что последний минимум достигается, если  $y = p_i(u)$ . Вводя единичный в норме  $p_i(\cdot)$  вектор  $u^0$ , получим

$$\begin{aligned} \delta_i &= \min_u \{p_i(u) - a_i^T u : p_i(u) \leq c_i\} = \\ &= \min_{u^0, p_i(u)} \{p_i(u)(1 - a_i^T u^0) : p_i(u^0) = 1, p_i(u) \leq c_i\}. \end{aligned} \quad (3.3)$$

Используя далее соотношение (1.17)

$$p_i^*(a_i) = \max \{a_i^T u^0 : p_i(u^0) = 1\}$$

(которое есть форма записи неравенства Коши-Шварца [68]), получим

$$\begin{aligned} \delta_i &= \min_u \{p_i(u)(1 - p_i^*(a_i)) : p_i(u) \leq c_i\} = \\ &= \begin{cases} 0, & p_i^*(a_i) \leq 1, \\ c_i(1 - p_i^*(a_i)), & p_i^*(a_i) > 1. \end{cases} \end{aligned} \quad (3.4)$$

Таким образом, в рассматриваемом частном случае величина (3.1) вычисляется аналитически для любых видов норм в (3.2). При этом не равные нулю подвекторы вводимого в базис вектора согласно (3.3), (3.4) определяются по формуле  $u_i = c_i u_i^0$ , где вектор  $u_i^0$  определяется, как правило, аналитически. Выражения для нахождения этого вектора в практически важных случаях (1.4), (1.3) приведены после формулы (1.17).

**ЗАМЕЧАНИЕ 6.2.** В случае, когда все нормы имеют вид (1.4), каждую компоненту  $u_{ij}$  можно рассматривать как одномерный вектор коррекции, которой соответствует матрица-столбец влияния  $B_{ij}$ . При этом задача (1.5) сводится к обычной задаче линейного программирования с двусторонними ограничениями ([101]).

### 6.3.3. Дополнения к предлагаемому алгоритму

О НАХОЖДЕНИИ ДОПУСТИМОГО БАЗИСНОГО РЕШЕНИЯ ЗАДАЧИ ОБОБЩЕННОЙ ЛИНЕЙНОЙ ИМПУЛЬСНОЙ КОРРЕКЦИИ. Для определенности рассмотрим случай 1 из предыдущего раздела. В этом случае задача обобщенной линейной импульсной коррекции сводится к задаче (1.6) с дополнительными ограничениями (1.23). Пусть переменные  $x_i, a_i, i = 1, \dots, n$  составляют некоторое допустимое решение этой задачи и только  $k$  переменных  $x_i$  удовлетворяют строгим неравенствам

$$0 < x_i < c_i, \quad i = 1, \dots, k$$

(нумерация условна). Это означает, что  $k$  импульсов коррекции лежат внутри областей своих ограничений. Аналогично разделу 6.1, будем называть это решение допустимым базисным решением, если  $k \leq m$ . Справедлив следующий результат.

**ТЕОРЕМА 6.3.** *Если существует оптимальное решение задачи (1.6) с дополнительными ограничениями (1.23), то существует и ее оптимальное допустимое базисное решение.*

**ДОКАЗАТЕЛЬСТВО.** Рассмотрим обычную задачу линейного программирования, столбцами которой являются векторы  $a_i$  данного оптимального допустимого решения задачи (1.6), а правая часть и ограничения на переменные — те же, что у задачи (1.6). Ее оптимальное решение, найденное обычным симплекс-методом, дает то же значение целевой функции, что заданное оптимальное решение задачи (1.6) и является допустимым базисным решением для этой задачи а, следовательно, и для рассматриваемой задачи (1.6).

Теорема 6.3 дает конструктивный способ нахождения допустимого базис-

ного решения задачи коррекции с ограничениями. Обычно это позволяет найти наименьшее число необходимых импульсов.

**О НАЧАЛЬНОМ ПРИБЛИЖЕНИИ.** Работа предлагаемого алгоритма должна начинаться с поиска допустимого базиса, который осуществляется стандартным в линейном программировании методом введения искусственных переменных. В результате этой процедуры либо будет найден начальный базис, либо установлено отсутствие решений исходной задачи при заданных ограничениях.

## 6.4. Некоторые численные результаты

### 6.4.1. Постановка задачи

Для иллюстрации работы предлагаемого алгоритма рассмотрим задачу коррекции кеплеровской орбиты спутника Земли. В качестве корректируемых параметров выберем следующие параметры: радиусы апогея и перигея  $r_a$  и  $r_\pi$ , аргумент перигея  $\omega$ , наклонение  $i$ , долгота восходящего узла  $\Omega$ . Необходимые для коррекции приращения этих параметров равны компонентам вектора  $b$  в задаче (1.5), а корректирующий импульс  $u$  в каждый момент времени будем записывать в орбитальной системе координат, оси которой направлены соответственно по радиусу от Земли, перпендикулярно радиусу в направлении движения и перпендикулярно плоскости орбиты:  $u = (V_r, V_u, V_n)$ .

Будем полагать, что коррекция проводится на одном витке траектории, возможные моменты приложения импульсов соответствуют точкам орбиты со значениями эксцентрической аномалии  $E = 0^0, 1^0, \dots, 359^0$ , а все нормы векторов в задаче (1.5) — евклидовы и ограничены одной константой  $c$ . Для проведения расчетов выбрана орбита с параметрами:  $a = 42163$  км,  $e = 0,00025$ ,  $i = 0,05^0$ ,  $\omega = 80^0$ ,  $\Omega = 270^0$ , где  $a$  — большая полуось орбиты,  $e$  — эксцентриситет орбиты. Значение эксцентрической аномалии в начальный момент времени полагается равным нулю.

Далее предполагается, что соотношения (1.1) получены линеаризацией истинных зависимостей между вектором  $b$  и корректирующими импульсами. При



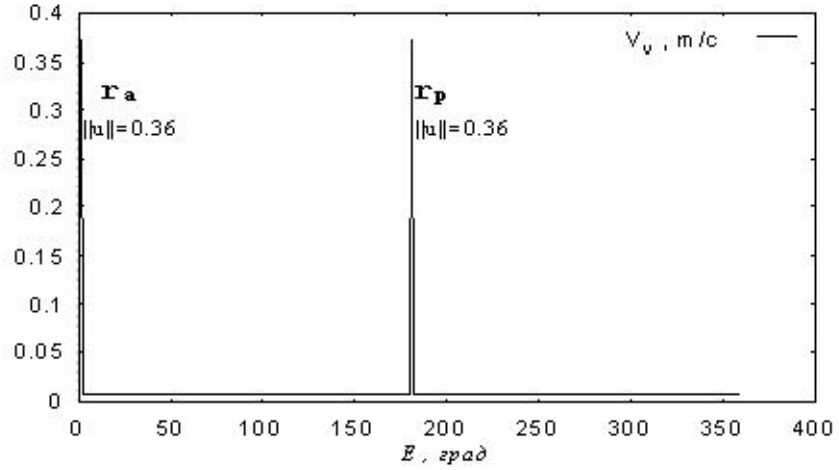


Рис.7. Коррекция радиусов апогея и перигея в плоской задаче без ограничений.

этом матрицы влияния вычисляются по формулам  $B_i = \partial b / \partial V_i$ , где  $V_i$  — вектор скорости, соответствующий  $i$ -й эксцентрической аномалии. Вычисление матриц  $B_i$  основано на формулах работы [62].

### 6.4.2. Плоская задача коррекции

В этой задаче будем рассматривать коррекцию параметров  $r_a$  и  $r_\pi$ . Рассмотрим сначала задачу без ограничений на величины импульсов. Для этого нужно задать достаточно большое  $c$ . Пусть  $\delta r_a = \delta r_\pi = 20\text{ км}$ . На рис. 7 приведены решения задачи (1.5) без ограничений на импульсы для трех случаев, в которых корректируемыми параметрами соответственно являются: только радиус апогея; только радиус перигея; оба радиуса. Во всех решениях ненулевыми оказались только компоненты  $V_u$ , прилагаемые в апогее и перигее (для рассматриваемой почти круговой орбиты это соответствует импульсам вдоль скорости, что согласуется с выводами теории оптимальных маневров [56]). Как видно из рисунка,

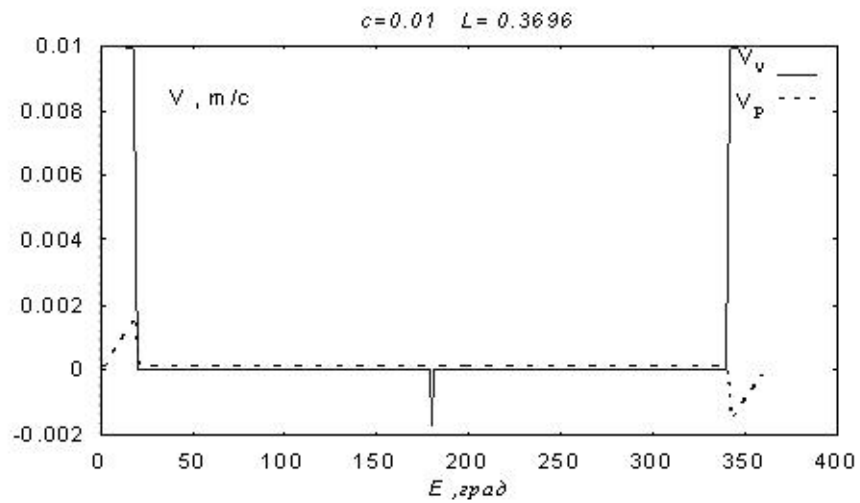


Рис.8. Коррекция радиуса апогея в плоской задаче.

решение задачи в третьем случае является объединением решений для первого и второго случаев.

Выберем теперь  $c = 0,01$  м/с. На рис. 8 — 10 приведены оптимальные импульсы для рассматриваемых трех случаев. Видно, что наличие ограничений привело к появлению радиальных составляющих импульса скорости для случаев отдельной коррекции радиусов апогея и перигея. Кроме того, одноимпульсная коррекция апогея в задаче без ограничений превратилась в трехинтервальную: два больших "размазанных" импульса в перигее и один (в пять раз меньше и в другую сторону) — в апогее. Совместная коррекция радиусов апогея и перигея уже не является здесь точным объединением однопараметрических коррекций, как это было в задаче без ограничений. Кроме того, на рис. 10 отсутствуют радиальные составляющие импульса и импульс в апогее. Отметим также, что при дальнейшем уменьшении величины  $c$  импульсы будут "размазываться" до тех пор, пока при некотором критическом значении  $c^*$  задача (1.5) не перестанет иметь решения из-за несовместности ограничений этой задачи.

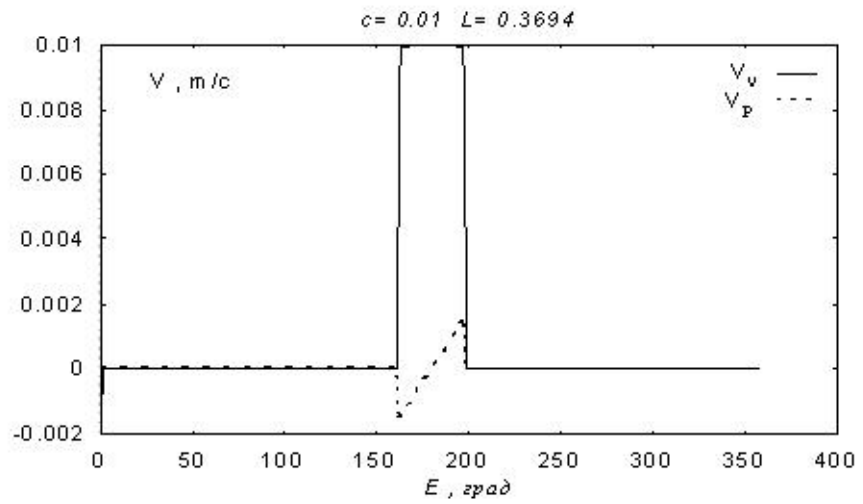


Рис.9. Коррекция радиуса перигея в плоской задаче.

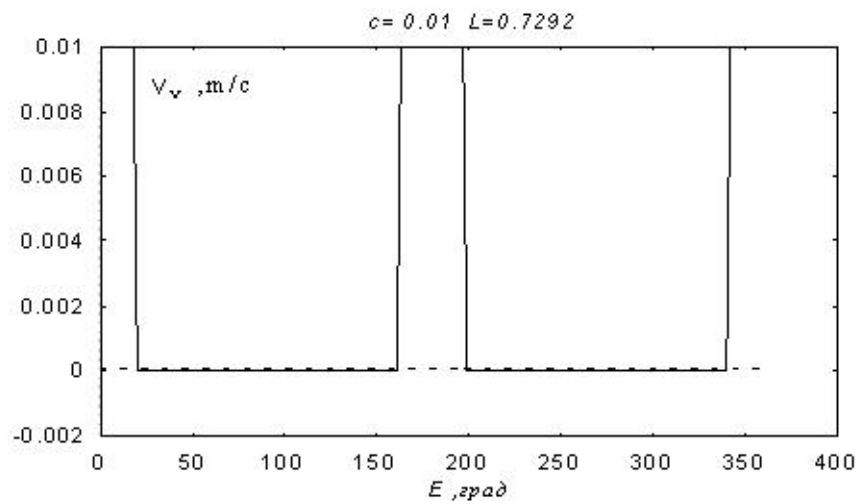


Рис.10. Совместная коррекция радиусов апогея и перигея в плоской задаче.

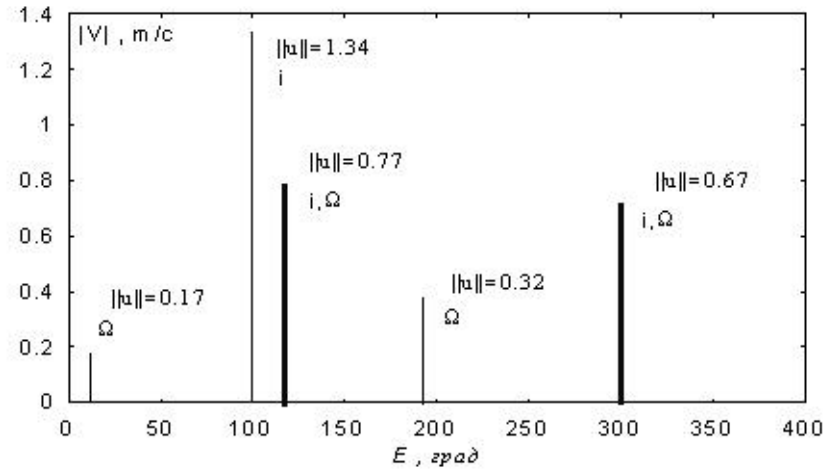


Рис.11. Модули импульсов скорости для оптимальной коррекции узла  $\Omega$ , наклонения  $i$  и совместной коррекции  $i, \Omega$  в задаче без ограничений.

### 6.4.3. Пространственная коррекция

Рассмотрим коррекцию наклонения и узла орбиты. Выберем  $\delta_i = 0,025^0$ ,  $\delta\Omega = 1^0$ . Как и в плоском случае, проведем расчеты для однопараметрической и совместной коррекций наклонения и узла орбиты.

Решение задачи без ограничений приведено на рис. 11. На нем изображен только  $V_n$ , поскольку остальные компоненты корректирующего импульса равны нулю. Отметим, что совместная коррекция не равна объединению однопараметрических, как это было в плоском случае при коррекции радиусов апогея и перигея.

В задаче с ограничением при  $c = 0,01$  для коррекции только наклонения (рис. 12) оптимальным является использование импульсов, перпендикулярных к плоскости орбиты.

Для коррекции узла орбиты (рис. 13) появляются составляющие  $V_u$ ; для совместной коррекции узла орбиты и наклонения (рис. 14) не равны нулю все

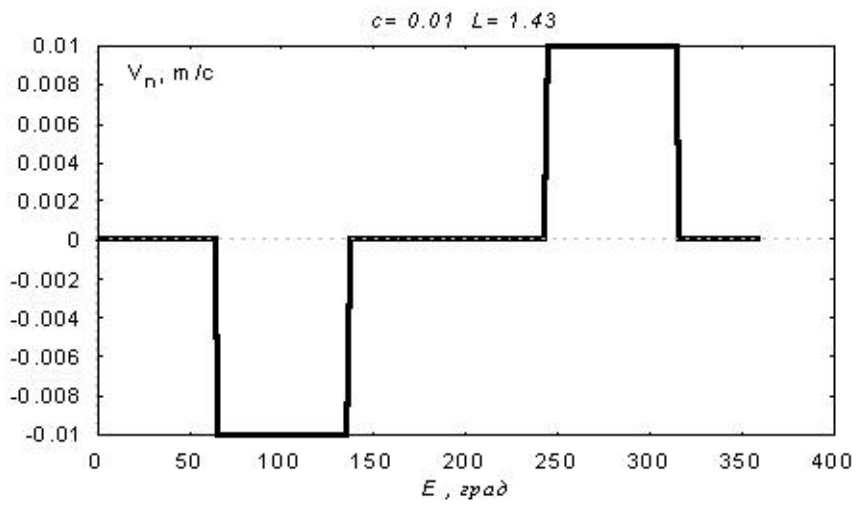


Рис.12. Коррекция наклонения орбиты при ограничениях на модуль импульса.

компоненты импульса в рассматриваемой системе координат.

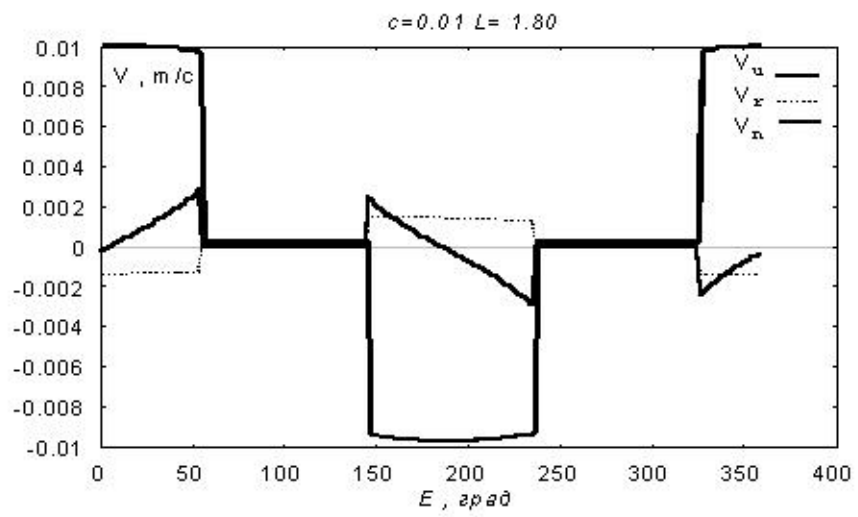


Рис.13. Коррекция узла орбиты при ограничениях на модуль импульса.

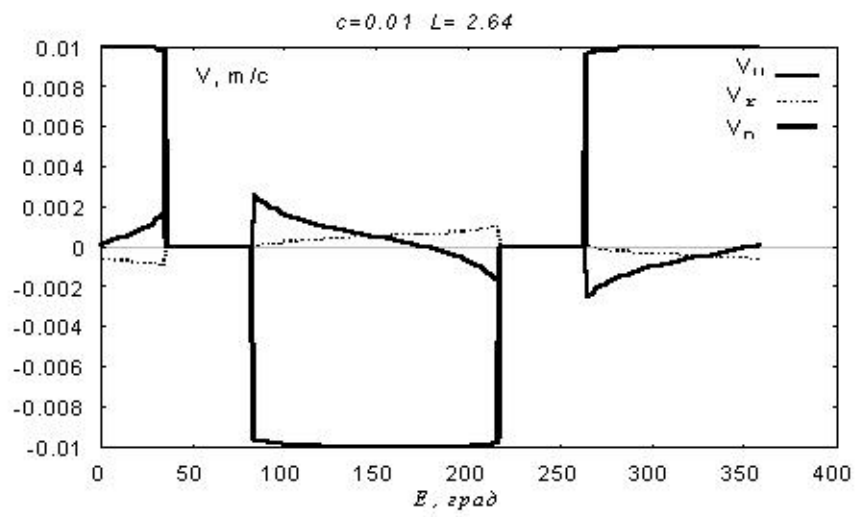


Рис.14. Совместная коррекция наклона и узла орбиты при ограничениях на модуль импульса.

## Заключение

В диссертации получены следующие основные научные результаты.

Получены критерии оптимальности и разработаны строго монотонные алгоритмы решения вырожденных и обобщенных задач линейного программирования, которые естественно обобщают известные симплекс-метод и метод генерации столбцов, строго монотонные лишь в невырожденном случае. Получены оценки для оптимума в важном случае целевой линейной функции с положительными коэффициентами. На примере планирования авиаперевозок показана эффективность нового алгоритма по сравнению с известными алгоритмами, используемыми для решения больших вырожденных задач линейного программирования.

Показана эквивалентность  $L$ - и  $MV$ - задач планирования эксперимента некоторым задачам оптимальной линейной идеальной коррекции. Получены критерии оптимальности и разработаны алгоритмы решения указанных задач. Найдено аналитическое решение  $MV$ - задачи в случае полиномиальной регрессии. Решена практически важная задача о выборе оптимальной программы измерений при оценивании геодинамических параметров Земли.

Показана возможность сведения к эффективно решаемым задачам обобщенного линейного программирования некоторых задач минимаксного оценивания, а также задачи робастного оценивания.

Задача идеальной линейной коррекции с ограничениями на величины импульсов сведена к обобщенной задаче линейного программирования. Получено эффективный алгоритм решения последней задачи для практически важных случаев. Проведены массовые расчеты для коррекции искусственного спутника Земли. Показана возможность моделирования оптимальной непрерывной коррекции с двигателем малой тягой оптимальной программой коррекции при достаточно малом пороге, ограничивающем величины импульсов.



В работе также развита теория линейного несмещенного оценивания – впервые установлено точное описание множества весовых матриц метода наименьших квадратов, определяющих заданную несмещенную оценку, и обобщены теоремы эквивалентности, позволяющие в схеме с мешающими параметрами рассматривать только линейные оценки при условиях несмещенности и без мешающих параметров.

## Список литературы

1. *Бажинов И.К., Почукаев В.Н.* Оптимальное планирование навигационных измерений в космическом полете. – М.: Машиностроение, 1976.
2. *Бахшиян Б.Ц.* Выбор оптимальных моментов независимых траекторных измерений // Косм. исследования. 1970. Т.8. N ° 1. С. 3-7.
3. *Бахшиян Б.Ц.* Некоторые задачи оценки точности прогнозирования параметров траектории и алгоритмы их решения // Косм. исследования. 1974. Т.12. N ° 6. С. 811-818.
4. *Бахшиян Б.Ц.* Комбинаторный метод решения задачи оптимальной коррекции траектории при ограничении на число импульсов // Косм. исследования. 1976. Т.14/ N ° 4. С. 630-632.
5. *Бахшиян Б.Ц.* Представление весовых матриц, определяющих заданную оценку наименьших квадратов // Навигационная привязка и статистическая обработка космической информации. М.: Наука, 1983, С. 81-90.
6. *Бахшиян Б.Ц.* Решение вырожденной и обобщенной задач линейного программирования на основе критериев оптимальности: Препринт N ° 1265. М.: Ин-т космических исследований АН СССР, 1987.
7. *Бахшиян Б.Ц.* Гарантированные характеристики точности линейного оценивания, их свойства и применение. Препринт N ° 1332. М.: Ин-т космических исследований АН СССР, 1987.

8. *Бахшиян Б.Ц.* Симплексный алгоритм решения оптимальной задачи гарантирующего оценивания с немоделируемыми возмущениями // Косм. исследования. 1988. Т.26 N ° 1. С. 127-141.
9. *Бахшиян Б.Ц.* Критерии оптимальности и алгоритмы решения вырожденной и обобщенной задач линейного программирования // Экономика и мат. методы. 1989. Т.28. N ° 2. С. 314-324.
10. *Бахшиян Б.Ц.* Эффективный симплексный алгоритм для некоторых задач минимаксного оценивания. Седьмая Всесоюзная конференция "Управление в механических системах". Тезисы докладов. г.Свердловск. 1990, С.12.
11. *Бахшиян Б.Ц.* Критерий оптимальности и алгоритм решения обобщенной задачи линейного программирования. "Понтрягинские чтения-VII". Тезисы докладов. г.Воронеж. 1996, С.31.
12. *Бахшиян Б.Ц.* О решении проблемы моментов методами линейного программирования. Тезисы докладов конф. "Математическое программирование и приложения." Ин-т математики и механики, г.Екатеринбург. 1999.
13. *Бахшиян Б.Ц.* Теория и симплексные алгоритмы решения задач L-, MV- и E-оптимального планирования эксперимента. Тезисы докладов конф. "Механика, управление и информатика. 1999, <http://www.iki.rssi.ru/seminar>.
14. *Бахшиян Б.Ц., Баюк О.А., Филимонов В.О.* Оптимальное планирование лазерных наблюдений спутников ЛАГЕОС 1, 2 //Космическая геодезия и современная геодинамика. Москва: Наука, 1996, С.205-221.
15. *Бахшиян Б.Ц., М.И.Войсковский, Ч.В.Пак.* Об оптимальной линейной идеальной коррекции при ограничениях на корректирующие импульсы //Космические исследования.1997.Т.35. N ° 4. С. 387-395.
16. *Бахшиян Б.Ц., М.И.Войсковский* О решении проблемы моментов методами линейного программирования // Вестник Тамбовского Университета, 2000, Т.5, вып.4. С. 412-413.

17. *Бахшиян Б.Ц., Матасов А.И., Федяев К.С.* О решении вырожденных задач линейного программирования. Автоматика и телемеханика. 2000. N ° 1. С.105-117.
18. *Бахшиян Б.Ц., Назиров Р.Р., Эльясберг П.Е.* Определение и коррекция движения. М.: Наука, 1980.
19. *Бахшиян Б.Ц., Назиров Р.Р., Эльясберг П.Е.* Оптимизация определения орбиты при неполном знании ковариационной матрицы и математического ожидания ошибок // Космические исследования. 1977. Т.15. N ° 5. С. 658-667.
20. *Бахшиян Б.Ц., Соловьев В.Н.* Применение теоремы двойственности к задаче оптимального гарантирующего оценивания // Космич. исследования. 1990. Т.28. N ° 2. С.163-169.
21. *Бахшиян Б.Ц., Соловьев В.Н.* Теория и алгоритмы решения задач L- и MV-оптимального планирования эксперимента. Автоматика и телемеханика. 1998. N ° 8. С. 80-96.
22. *Бахшиян Б.Ц., Суханов А.А.* Выбор оптимального состава астроизмерений для определения орбит искусственных спутников // Космические исследования. 1977. Т.15. N ° 1. С. 3-7.
23. *Бахшиян Б.Ц., Суханов А.А., Эльясберг П.Е.* Априорная точность прогноза положения кометы Галлея по наземным и бортовым наблюдениям // Космические исследования. 1985. Т.23. N ° 5. С. 876-885.
24. *Баяк О.А.* Определение параметров вращения Земли по лазерным наблюдениям искусственных спутников Земли ЛАГЕОС 1, 2 // Космическая геодезия и современная геодинамика. М: Наука, 1996, С.233-244.
25. *Белоусов Л.Ю.* Определение оптимальных моментов измерений // Космич. исследования, 1969. т. 7. N ° 1. с. 28-34.

26. Белоусов Л.Ю., Комаров В.А. Некоторые общие результаты в задаче определения оптимальных моментов измерения для различных моделей ошибок измерений // Косм. исследования. 1970. Т.8. N° 3. С. 452-453.
27. Гольштейн Е.Г. Выпуклое программирование. Элементы теории. М.: Наука, 1989.
28. Войсковский М.И., Меринов И.Е. Симплексный алгоритм решения минимаксной задачи оценивания // Препринт 1697 Института космических исследований АН СССР. 1990.
29. Войсковский М.И. Симплексный алгоритм решения задачи  $MV$ -оптимального планирования эксперимента // Препринт 1979 Института космических исследований РАН. 1998.
30. Гольштейн Е.Г., Третьяков Н.В. Модифицированные функции Лагранжа. М.: Наука, 1989.
31. Данциг Дж. Линейное программирование, его применения и обобщения. М.: Прогресс, 1966. (*G.B.Dantzig. Linear Programming and Extensions. Princeton U.P., 1963.*)
32. Демиденко Е.З. Линейная и нелинейная регрессии. М.: Финансы и статистика, 1981.
33. Ермаков С.М., Жигляевский А.А. Математическая теория оптимального эксперимента. М.: Наука, 1987.
34. Математическая теория планирования эксперимента / Под ред. Ермакова С.М. М.: Наука, 1983.
35. Ершов В.Г. Об оптимизации программы траекторных измерений // Косм. исследования. 1971. Т.9. N° 1. С. 17.
36. Иоффе А.Д., Тихомиров В.М. Теория экстремальных задач. - М.: Наука, 1974.

37. *Канторович Л.В.* Математические методы организации и планирования производства. - Л.: Изд-во Ленинградского университета, 1939.
38. *Канторович Л.В., Акилов Г.П.* Функциональный анализ в нормированных пространствах. М.: Физматгиз, 1959.
39. *Красовский Н.Н.* Теория управления движением.- М.: Наука, 1968.
40. *Куржанский А.Б.* Управление и оценивание в условиях неопределенности.- М.: Наука, 1977.
41. *Куркин О.М. Коробочкин Ю.Б., Шаталов С.А.* Минимаксная обработка информации. - М.: Энергоатомиздат, 1990.
42. *Леман Э.* Теория точечного оцениванияю - М.: Наука, 1991.
43. *Лидов М.Л.* К априорным оценкам точности определения параметров по методу наименьших квадратов // Косм. исследования. 1964. Т.2. N ° 5. С. 713-718.
44. *Лидов М.Л.* Математическая аналогия между некоторыми оптимальными задачами коррекции траекторий и выбора состава измерений и алгоритмы их решения // Космические исследования. 1971. Т.9. N ° 5. С. 687-706.
45. *Лидов М.Л.* Эффективный алгоритм решения задачи о выборе оптимальной программы измерений // Космические исследования. 1985. Т.23. N ° 4. С. 499-517.
46. *Лидов М.Л.* О модификации симплекс-метода линейного программирования в случае вырождения//Космические исследования. 1991. Т.29. N ° 4. С. 499-508.
47. *Лидов М.Л., Бакума Л.М.* Определение оптимальной программы измерений с ограничениями на ошибки оценки трех параметров движения суточного спутника // Космические исследования. 1986. Т.24 С. 483-496.

48. *Лидов М.Л., Бакума Л.М.* Экспериментальная проверка эффективности нового алгоритма для задачи оценивания с немоделируемыми ускорениями // Космические исследования. 1991. Т.29. N ° 3.
49. *Лидов М.И., Бахшиян Б.Ш., Матасов А.И.* Об одном направлении в проблеме гарантирующего оценивания (обзор) // Космические исследования. 1991. Т.29. N ° 5. С. 659-684.
50. Лидов М.Л., Матасов А.И. Об одном обобщении задачи о "наихудшей корреляции" // Космические исследования, 1989. т. 27, N ° 3, с. 454-456.
51. Липцер Р.Ш., Ширяев А.Н. Статистика случайных процессов. - М.; Наука, 1974.
52. Малышев В.В., Кибзун А.И. Анализ и синтез высокоточных систем управления летательными аппаратами. - М.: Машиностроение, 1987.
53. Марчук А.Г., Осипенко К.Ю. Наилучшее приближение функций, заданных с погрешностью в конечном числе точек // Математические заметки, 1975. т. 17, N ° 3, с. 359-368.
54. Матасов А.И. Об оптимальности линейных алгоритмов гарантирующего оценивания. Часть I // Космические исследования, 1988. т. 26, N ° 5, с. 643-653.
55. Матасов А.И. Об оптимальности линейных алгоритмов гарантирующего оценивания. Часть II // Космические исследования, 1988. т. 26, N ° 6, с. 807-812.
56. *Лоуден Д.Ф.* Оптимальные траектории для космической навигации. М.: Мир, 1966.
57. *Лэсдон Л.С.* Оптимизация больших систем. М: Наука, 1975.
58. *Мину М.* Математическое программирование. Теория и алгоритмы. М.: Наука, 1989.

59. *Муртаф Б.* Современное линейное программирование М.: Мир, 1984. (A.Murtagh. Advanced Linear Programming: Computation and Practice. McGraw-Hill International Book Company, 1981.)
60. *Мелас В.Б.* Одна теорема двойственности и E-оптимальность // Заводская лаборатория. Т.82. N ° 3. С. 48-50.
61. *Мелас В.Б., Крылова Л.А.* E-оптимальные планы для кубической регрессии на симметричном отрезке // Вестник СПбГУ Сер.1, 1996, вып.3 (N ° 15). С. 26-30.
62. *Назирова Р.Р., Тимохова Т.А.* Оптимальная линейная коррекция эллиптических орбит // Автоматика и телемеханика. 1993. N ° 3 С. 93-101.
63. *Обен Ж.-П., Экланд И.* Прикладной нелинейный анализ. М.: Мир, 1988.
64. *Панков А.Р.* Стратегии управления в линейной стохастической системе с негауссовскими возмущениями // Автоматика и телемеханика, 1994. N ° 6, с. 74-83.
65. *Парусников Н.А., Морозов В.М., Борзов В.И.* Задача коррекции в инерциальной навигации. - М.; Изд-во МГУ, 1982.
66. *Поляк Б.Т.* Введение в оптимизацию. М.: Наука, 1983.
67. *Пшеничный Б.Н.* Необходимые условия экстремума. М.: Наука, 1982.
68. *Рокафеллар Р.* Выпуклый анализ. М.: Мир, 1973
69. *Смоляк С.А. Титаренко Б.П.* Устойчивые методы оценивания. М.: Статистика, 1980.
70. *Соловьев В.Н.* Двойственность некоторых невыпуклых экстремальных задач // Ж.вычисл.матем. и мат.физики. 1987. Т.27. N ° 2. С. 459-464.
71. *Соловьев В.Н.* Двойственные алгоритмы минимаксного оценивания параметров движения в непрерывной постановке // Космические исследования, 1991. т. 29, N ° 1, с. 127-132.



72. *Соловьев В.Н.* Об оптимальности линейных алгоритмов в задаче гарантирующего оценивания при случайных ошибках измерений //Космические исследования, 1994. т. 32, N ° 2, с. 122-124.
73. *Соловьев В.Н.* Двойственные экстремальные задачи и их применения к задачам минимаксного оценивания //Успехи математических наук, 1997. т. 52, N ° 4, с. 49-86.
74. *Тихомиров В.М.* Выпуклый анализ. В книге: Современные проблемы математики, фундаментальные направления. Анализ-8, - М.: ВИНТИ, 1987, с. 5- 101.
75. *Тихомиров В.М.* Теория приближений. В книге: Современные проблемы математики. Фундаментальные направления. Анализ-8, - М.: ВИНТИ, 1987, с. 103-260.
76. *Федоров В.В.* Теория оптимального эксперимента. М.: Наука, 1971.
77. *Федоров В.В.* Активные регрессионные эксперименты // Математические методы планирования эксперимента. Новосибирск: Наука, 1983, С.19.
78. *Р. Хорн, Ч.Джонсон.* Матричный анализ.М.:Мир, 1989.
79. Хьюбер П. Робастность в статистике. - М.: Мир, 1984.
80. *Черноузько Ф.Л.* Оценивание фазового состояния динамических систем. - М.: Наука,1988.
81. *Черноузько Ф.Л., Меликян А.А.* Игровые задачи управления и поиска. М.: Наука, 1978.
82. *Черноузько Ф.Л., Колмановский В.Б.* Оптимальное управление при случайных возмущениях. М.: Наука, 1979.
83. *Экланд И., Темам Р.* Выпуклый анализ и вариационные проблемы. М.: Мир, 1979.

84. *Эльясберг П.Е.* Определение движения. М.: Наука, 1976.
85. *Эльясберг П.Е.* Измерительная информация: сколько ее нужно? как ее обрабатывать? М.: Наука, 1983.
86. *B.Ts.Bakhshiyn, I.A.Yastrzhemsky* Comet Halley orbit optimum determination by means of ground based astrometric observations and on board observations from the "Vega" spacecraft//Adv.Space Res. 1987.V.5. N12. P.181-183.
87. *B.Ts.Bakhshiyn, O.A.Bayuk* Optimization of the distribution of centerline the observations of the oscillatory system//Control of Oscillatios and Chaos. Proceedings of 1st International Conference. 1997. St.Peterburg. V.2. P.238-241.
88. *R.G.Bland.* New finite pivot rules for simplex method// Math. Oper. Res. 1977. V. 2. P. 103-107.
89. *Elfvig G.* Optimum allocation in linear programming // Ann.Math.Statist. 1952. V.23. P. 255.
90. *G.B.Dantzig.* Making Progress During a Stall in the Simplex Algorithm // Linear Algebra and its Applications. 1989. V.114/115. P. 251-259.
91. *Huber P.J.* Robust Estimation of a Location Parameter//Annals of Mathamatical Statistics, 1964. V.5,PP.73–101.
92. *Kibzun A.I., Kan Yu.S.* Stochastic Programming Problems (with Probability and Quantile Functions). John Wiley and Sons, Chichester-New York-Brisbane-Toronto- Singapore, 1996.
93. *T.C.T. Kotiun, D.I.Steinberg.* On the Possibility Cycling with the Simplex Method//Oper. Res. 1984. V.26. N ° 2. P. 374-376.
94. *Mitra S.K., Moore J.B.* Gauss-Markoff estimation with an incorrect dispersion matrix // Sankhya. 1973, Ser.A. V.35. P. 139-152.
95. *Pukelsheim F.* Optimal design of experiments. New York: J.Wiley and Sons, 1993.

96. *Rao C.R.* Unified theory of linear estimation// Sankhya. 1971, Ser.A. V.335. P. 371-399.
97. *Rao C.R.* Representations of best linear unbiased estimations in the Gauss-Markoff models with a singular dispersion matrix // J.Multivar.Anal. 1973. V.3. P. 276-292.
98. *Rao C.R.* On a unified theory of estimation in linear models – a review of recent results // Perspectives in probability and statistics.L.etc.: Acad.press, 1975, P. 89-104.
99. *Rao C.R., Mitra S.K.* Generalized inverse of matrices and its applications. N.Y.: Wiley, 1971.
100. *D.M. Ryan, M.R.Osborne.* On the solution of highly degenerate linear programmes // Mathematical Programming. 1988. V.41. P. 385-392.
101. *Waespy C.M.* Linear-programming solutions for orbital-transfer trajectories // Operation Research. 1970. V.18. N ° 5. P. 635-653.
102. *P.Wolfe.* A technique for resolving degeneracy in linear programming. J. Soc. Indust. and Appl. Math. 1963. V.11. P. 205-211.