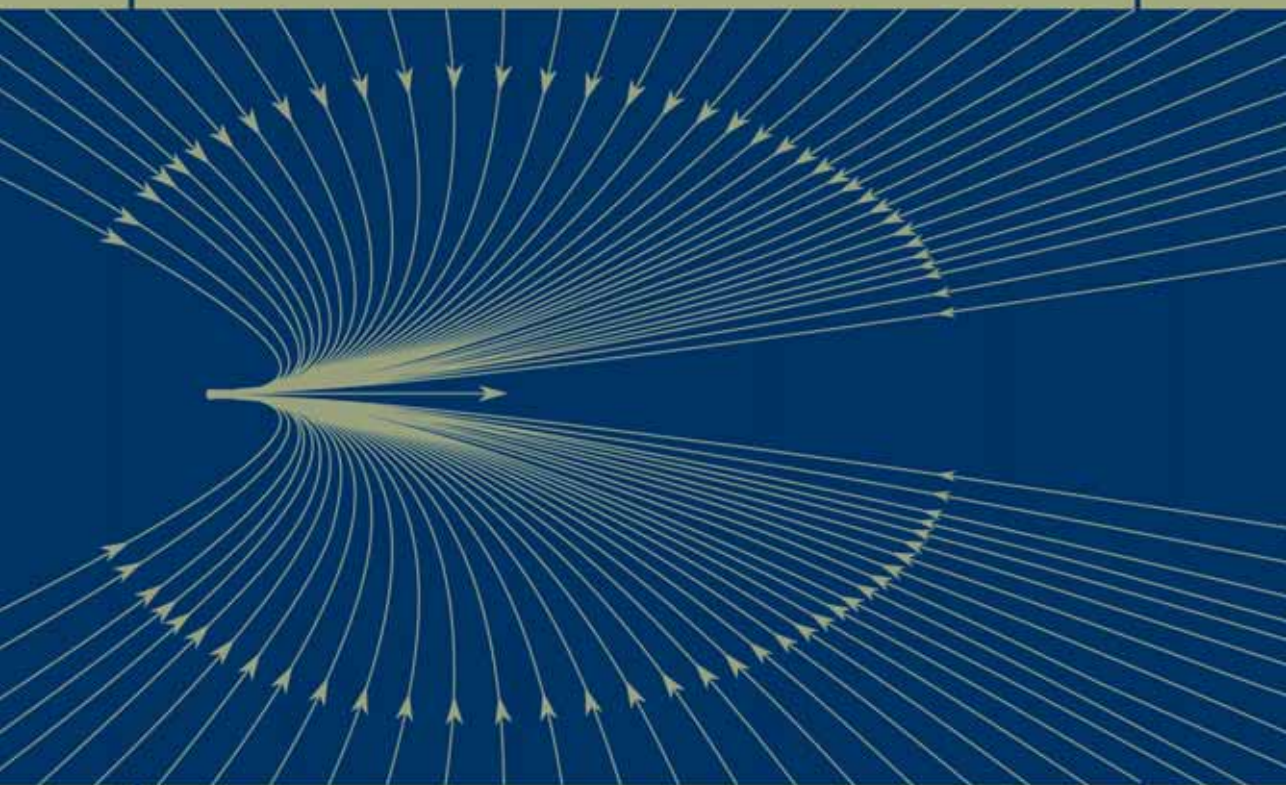




САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

А. С. Матвеев

ВВЕДЕНИЕ В МАТЕМАТИЧЕСКУЮ ТЕОРИЮ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ



УДК 519.71
ББК 32.81я7
М33

Рецензенты: д-р физ.-мат. наук проф. *В. Б. Смирнова*
(С.-Петербургский гос. архитектурно-строительный ун-т)
д-р физ.-мат. наук доц. *Н. В. Кузнецов*
(С.-Петербургский гос. ун-т)

*Рекомендовано к изданию
Редакционно-издательским советом
математико-механического факультета
Санкт-Петербургского государственного университета*

Матвеев А. С.

М33 Введение в математическую теорию оптимального управления:
Учебник. — СПб.: Изд-во С.-Петерб. ун-та, 2018. — 194 с.

ISBN 978-5-288-05809-7

Цель данного учебника — ознакомить читателя с математической теорией оптимального управления, её связями с другими разделами теории экстремальных задач, спецификой типичных математических задач оптимального управления и вытекающих из неё проблем, а также с основными положениями и базовыми подходами этой теории и их применением к решению конкретных задач. В изложении материала упор сделан на подходе, основанном на применении функционального анализа, который был разработан и развит Санкт-Петербургской (Ленинградской) школой математической кибернетики, созданной профессором СПбГУ В. А. Якубовичем, позволяющем не только рассматривать с единой точки зрения экстремальные задачи разных типов, но и унифицировать необходимые условия экстремума первого и более высокого порядка.

В основу учебника положен материал курса лекций, читаемых автором на математико-механическом факультете Санкт-Петербургского государственного университета.

Предназначен для студентов и аспирантов, обучающихся по направлению «Прикладная математика и информатика» и на смежных с ним специальностях.

УДК 519.71
ББК 32.81я7

ISBN 978-5-288-05809-7

© Санкт-Петербургский
государственный
университет, 2018
© А. С. Матвеев, 2018

ОГЛАВЛЕНИЕ

Используемые обозначения	6
Некоторые используемые понятия	11
Предисловие	13
Глава 1. Классические экстремальные задачи	
и задачи оптимального управления	15
1.1. Классические экстремальные задачи	16
1.1.1. Условие Ферма в гладкой задаче безусловной оптимизации	—
1.1.2. Правило множителей Лагранжа в гладкой задаче математического программирования	17
1.1.3. Задачи выпуклого и линейного программирования. Усиленный принцип оптимальности Лагранжа. Теорема двойственности	24
1.1.4. Задачи вариационного исчисления	29
1.1.5. Уравнение Эйлера	30
1.1.6. Каноническая форма уравнения Эйлера. Преобразование Лежандра	31
1.1.7. Гладкость экстремалей	32
1.1.8. Условие Лежандра	—
1.1.9. Условие Вейерштрасса	—
1.1.10. Условие Якоби	33
1.1.11. Вычисление сопряжённых точек	35
1.1.12. Глобальные условия Лежандра и Якоби как достаточные условия глобального экстремума	40
1.2. Простейший пример задачи оптимального управления и её решения	42
1.2.1. Постановка задачи об оптимальном успокоении гармонического осциллятора	—
1.2.2. Решение задачи об оптимальном успокоении гармонического осциллятора	44
1.3. Классический метод вариаций и задачи оптимального управления	50
1.3.1. Классический метод вариаций вывода необходимых условий оптимальности	—

1.3.2. Пример общей постановки задачи оптимального управления	54
1.3.3. Классический метод вариаций и задачи оптимального управления	58
Глава 2. Метод пучков (неклассических вариаций)	
вывода необходимых условий оптимальности	61
2.1. Основные определения, идеи и факты	—
2.1.1. Пучки кривых: основные определения	—
2.1.2. Простейшие необходимые условия оптимальности, связанные с пучками кривых	64
2.2. Примеры пучков и связанных с ними необходимых условий оптимальности	68
2.2.1. Пучок классических вариаций	—
2.2.2. Пучок анизотропных вариаций	69
2.2.3. Стандартное пространство управлений	79
2.2.4. Пучок простых игольчатых вариаций	81
2.2.5. Дифференцирование интегрального функционала по пучку простых игольчатых вариаций	83
2.2.6. Необходимые условия экстремума в задаче минимизации интегрального функционала	84
2.2.7. Пучок сложных игольчатых вариаций	85
2.2.8. Выпуклое дифференцирование интегрального функционала по пучку сложных игольчатых вариаций	89
2.2.9. Необходимые условия экстремума в задаче минимизации интегрального функционала при интегральных ограничениях	91
Глава 3. Абстрактная теория оптимального управления	93
3.1. Постановка абстрактной задачи оптимального управления	94
3.2. Задачи без дополнительных ограничений. План вывода условий оптимальности	96
3.3. Теорема о неявной функции	98
3.4. Разложение и дифференцирование по пучку сложной функции	103
3.5. Необходимые условия оптимальности в абстрактной задаче оптимального управления	108
Глава 4. Принцип максимума Понтрягина	112
4.1. Задача оптимального управления с фиксированным интервалом времени	—
4.1.1. Формализация задачи	114
4.1.2. Проверка условий теоремы 3.5.1	115
4.1.3. Расшифровка заключения абстрактной теоремы и формулировка принципа максимума Понтрягина	125
4.1.4. Завершение доказательства принципа максимума: случай неограниченного множества допустимых управлений ...	130

4.2. Принцип максимума для задачи оптимального управления с фиксированным интервалом времени	134
4.2.1. Постановка задачи	—
4.2.2. Принцип максимума Понтрягина для задачи с нефиксированным временем	136
4.2.3. О применении принципа максимума для решения задач оптимального управления	139
4.2.4. Принцип максимума Понтрягина для задачи с фиксированным временем как частный случай теоремы 4.2.1 ...	142
4.2.5. Принцип максимума Понтрягина для задачи с фиксированным временем и фиксированным начальным и конечным состоянием	145
4.2.6. Замечания о локальном экстремуме	148
4.3. Принцип максимума и вариационное исчисление	149
4.3.1. Постановка простейшей задачи вариационного исчисления и её сведение к задаче оптимального управления ...	—
4.3.2. Принцип максимума для задачи вариационного исчисления	150
4.4. Пример применения принципа максимума: оптимальное по быстродействию успокоение гармонического осциллятора	154
4.4.1. Постановка задачи об оптимальном успокоении гармонического осциллятора	—
4.4.2. Выписывание принципа максимума для рассматриваемой задачи	155
4.4.3. Предварительный анализ принципа максимума	157
4.4.4. Выводы из принципа максимума	158
4.4.5. Коленчатое управление, приводящее осциллятор в состояние покоя	160
4.4.6. Оптимальный закон управления	167
4.4.7. Заключительные замечания	168
4.5. Принцип максимума для задачи оптимального управления системой, описываемой интегральным уравнением	—
4.5.1. Формализация задачи	170
4.5.2. Проверка условий теоремы 3.5.1	171
4.5.3. Расшифровка заключения абстрактной теоремы и формулировка аналога принципа максимума Понтрягина ..	181
4.6. Задачи для самостоятельного решения	182
Литература	191
Список иллюстраций	193

ПРЕДИСЛОВИЕ

В середине прошлого века, в основном в связи с потребностями, связанными с механикой космического полёта и системами автоматического управления различного назначения, возникло новое направление в кибернетике — математическая теория оптимального управления. Она интенсивно развивалась и развивается многими научными коллективами, находящими новые приложения и мотивации, причём не только в области техники, но и в экономике, биологии, медицине, экологии, демографии и в других областях.

Типичная математическая модель оптимального управления — это модель поиска экстремума некоторого функционала. Следовательно, математические задачи оптимального управления родственны задачам на поиск экстремума, которые рассматриваются в классических разделах математики, например задачам математического программирования и вариационного исчисления. Математическая теория оптимального управления использует и частично содержит в себе результаты, идеи и методы этих разделов, но не сводится к ним. Специфика типичных задач оптимального управления такова, что для их систематического комфортного решения потребовались новые результаты, идеи и методы, развивающие фундамент, созданный ранее. Ключевым звеном этого развития было открытие в конце 1950-х годов знаменитого принципа максимума Понтрягина, который указывает необходимые условия оптимальности управляющего воздействия. Среди других важных звеньев отметим разработку принципа динамического программирования применительно к задачам оптимального управления, а также развёрнутое исследование проблемы существования решений неклассических вариационных задач и достаточных условий оптимальности управления. Появление очередного класса задач на поиск экстремума с потребностью в существенной доработке математической теории стимулировало интерес к поиску идейного фундамента, объединяющего экстремальные задачи разных классов и позволяющего

возводить основную часть относящихся к ним теорий едиными методами вплоть до её полного погружения в рамки более общей теории в качестве частного случая. Одним из результатов этого течения «универсализма» стало признание функционального анализа в качестве мощного источника унификации в математической теории оптимального управления.

Цель данного учебника — первоначально ознакомить читателя с математической теорией оптимального управления, её связями с другими разделами теории экстремальных задач, спецификой типичных математических задач оптимального управления и вытекающих из неё проблем, а также с основными положениями и базовыми подходами этой теории и их применением к решению конкретных задач. В изложении материала упор сделан на подходе, основанном на применении функционального анализа, который разработан и развит санкт-петербургской (ленинградской) школой математической кибернетики, созданной профессором СПбГУ В. А. Якубовичем. Этот подход позволяет не только рассматривать с единой точки зрения экстремальные задачи разных типов и, в частности, выводить аналоги принципа максимума для различных задач оптимального управления, в том числе описываемых уравнениями разных типов (обыкновенными дифференциальными, в частных производных, интегральными, с запаздывающим аргументом и т. п.), но и унифицировать необходимые условия экстремума первого и более высокого порядка. Материал книги частично пересекается с содержанием монографии [1], которая ориентирована в первую очередь на научных работников; её можно рекомендовать читателю, заинтересованному в расширенном и углублённом знакомстве с затронутым направлением.

В основе учебника лежит материал курса лекций, читаемых автором на математико-механическом факультете Санкт-Петербургского государственного университета. Автор признателен Российскому научному фонду (грант 14-21-00041п) за поддержку.

Глава 1

КЛАССИЧЕСКИЕ ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ И ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Объект теории оптимального управления — управляемые системы, т. е. системы, на состояние и поведение которых можно целенаправленно воздействовать. Выбор воздействия, именуемого *управлением*, при котором состояние и поведение удовлетворяют определённым требованиям, — основная задача *теории управления*. Об *оптимальном управлении* говорят в тех случаях, когда эти требования включают достижение наилучшего в том или ином смысле результата. *Математическая теория оптимального управления* занимается математическими методами построения оптимального управления. При этом она систематически отталкивается от математического анализа оптимума; необходимые условия оптимальности играют в ней важную роль.

Применение математических методов требует предварительной *математической формализации задачи* или, другими словами, построения её *математической модели*. Здесь прежде всего даётся математическое описание как состояния системы, так и управляющего воздействия на неё. Следующий шаг — математический закон, описывающий влияние управляющего воздействия на состояние. Три упомянутые компоненты объединяют термином *модель объекта управления*, или, для краткости, просто *объект управления*; пару «управление — отвечающее ему состояние» назовём *процессом*. Далее необходимо описать имеющиеся требования математическими формулами. И наконец, нужно математическими средствами ответить на вопрос, что такое наилучший результат. Теория оптималь-

ного управления в основном ориентирована на случай, когда ответ даётся *функционалом качества*: он сопоставляет каждому процессу определённое число, причём процесс тем лучше, чем это число меньше¹.

Итак, типичная математическая задача оптимального управления — это задача на поиск экстремума функционала. Классические результаты теории экстремальных задач уже на младших курсах занимают видное место в вузовских программах и по фундаментальной, и по прикладной математике и рассматриваются как обязательный компонент базовой культуры математика. Встреча с теорией оптимального управления обычно происходит позже на фоне определенного багажа знаний в области экстремальных задач. *Требуяют ли задачи оптимального управления дополнить этот багаж?* Ответ на этот вопрос начнём с краткого обзора упомянутых классических результатов. В его фокусе будут описание классов затронутых задач и условия экстремума. Ввиду обзорного характера следующего параграфа изложенные в нем факты приводятся без доказательства; с ними можно ознакомиться, например, в работах [1–7].

1.1. КЛАССИЧЕСКИЕ ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ

1.1.1. Условие Ферма в гладкой задаче безусловной оптимизации

Оно касается задачи поиска аргумента, оптимизирующего значение гладкой функции $\varphi : O \rightarrow \mathbb{R}$ на открытом подмножестве O конечномерного евклидова пространства $\mathbb{R}^n$: в точке минимума $x^0 \in O$ градиент обращается в нуль:

$$\nabla \varphi(x^0) = 0 \Leftrightarrow \frac{\partial \varphi}{\partial x_i}(x^0) = 0 \quad \forall i \in [1 : n]. \quad (1.1.1)$$

Здесь число n скалярных уравнений Ферма совпадает с числом неизвестных компонент $x_1^0, \dots, x_n^0$ точки экстремума x^0 . Поэтому решение этих уравнений обычно позволяет кардинально снизить неопределённость относительно точки минимума, так как в типичном случае множество решений невелико, по крайней мере дискретно. Вместе с тем решения системы уравнений Ферма необязательно доставляют минимум, например к ним относятся и все точки локального максимума. В случае нескольких решений полезно необходимое условие второго порядка: в точке минимума

¹ Случай, когда «лучше» означает «больше», сводится к указанному путем умножения функционала на -1 .

гессиан — неотрицательно определённая матрица

$$\varphi''(x^0) \geq 0. \quad (1.1.2)$$

Указанные условия выполнены не только в точке глобального, но и локального минимума. Если в (1.1.2) заменить нестрогое неравенство строгим и объединить результат с условием Ферма, получаем достаточное условие локального минимума: если

$$\nabla\varphi(x^0) = 0 \quad \text{и} \quad \varphi''(x^0) > 0,$$

то функция $\varphi(\cdot)$ достигает локального минимума в точке $x^0 \in O$.

1.1.2. Правило множителей Лагранжа

в гладкой задаче математического программирования

Задача заключается в минимизации функции $\varphi(\cdot)$ на множестве решений конечной системы скалярных уравнений и неравенств

$$\varphi(x) \rightarrow \min_{x \in D}, \quad \text{где}$$

$$\begin{aligned} D := \{x \in O : g_1(x) = \dots = g_m(x) = 0, \\ p_1(x) \leq 0, \dots, p_k(x) \leq 0\}. \end{aligned} \quad (1.1.3)$$

Все функции гладкие и заданы на открытом множестве $O \subset \mathbb{R}^n$. В (1.1.3) могут отсутствовать либо равенства ($m = 0$), либо неравенства ($k = 0$), либо все ограничения ($m = k = 0$). Они накладывают условия на выбор точек x , среди которых ищется наилучшая точка; в таких случаях говорят об *условной оптимизации*. Опуская ограничения в (1.1.3), получаем задачу безусловной оптимизации $\varphi(x) \rightarrow \min_{x \in O}$.

Сопоставим каждому уравнению $g_i(x) = 0$ закононеопределенный *множитель Лагранжа* $\lambda_i^g \in \mathbb{R}$, каждому неравенству $p_j(x) \leq 0$ — неотрицательный *множитель Лагранжа* $\lambda_j^p \geq 0$ и такой же множитель $\lambda^\varphi \geq 0$ — минимизируемому функционалу. Умножая левую часть каждого уравнения и неравенства, а также минимизируемый функционал на ассоциированный множитель и суммируя результаты, получаем *функцию Лагранжа* или, иначе, *лагранжиан*

$$L(x) = L(x|\lambda) := \lambda^\varphi \varphi(x) + \sum_{i=1}^m \lambda_i^g g_i(x) + \sum_{j=1}^k \lambda_j^p p_j(x). \quad (1.1.4)$$

Здесь

$$\lambda = \mathbf{col}(\lambda^\varphi, \lambda_1^g, \dots, \lambda_m^g, \lambda_1^p, \dots, \lambda_k^p) \in \mathbb{R}^{k+m+1}$$

— набор множителей Лагранжа и любая сумма вида $\sum_{r=1}^0 \dots$ считается равной нулю.

Правило множителей Лагранжа опирается на следующий факт, который при наличии в (1.1.3) неравенств называют *условиями Куна—Таккера*. Если $x^0 \in O$ доставляет локальный минимум в задаче (1.1.3), то существует набор множителей Лагранжа со следующими свойствами, которые включают обращение градиента лагранжиана в нуль в точке x^0 :

$$\nabla L(x^0) = 0, \quad \lambda^0 \geq 0, \quad \lambda_1^p \geq 0, \dots, \lambda_k^p \geq 0, \quad (1.1.5)$$

$$\lambda_1^p p_1(x^0) = 0, \dots, \lambda_k^p p_k(x^0) = 0, \quad \lambda^0 + \sum_{j=1}^k \lambda_j^p + \sum_{i=1}^m |\lambda_i^g| > 0. \quad (1.1.6)$$

В (1.1.6) последнее неравенство исключает тривиальный случай, когда все множители равны нулю, и именно поэтому все остальные соотношения из (1.1.5) и (1.1.6) выполнены. Равенства из (1.1.6) называют *условиями дополняющей нежёсткости*. Они означают, что множитель, отвечающий неравенству из постановки задачи (1.1.3), может (но не обязан) отличаться от нуля, только если в точке экстремума это неравенство $p_j(x^0) \leq 0$ выполнено как равенство $p_j(x^0) = 0$. Такие неравенства и индексы j называют *критическими* в точке x^0 . В достаточно малой окрестности x^0 прочие неравенства, для которых $p_j(x^0) < 0$, заведомо выполнены по непрерывности и поэтому не влияют на множество D и, следовательно, на саму задачу (1.1.3). Они фактически не участвуют и в обсуждаемом результате, так как входят в лагранжиан с нулевыми множителями.

Условия (1.1.5), (1.1.6) служат основой *правила множителей Лагранжа*: определить все решения (x, λ) системы уравнений

$$\begin{aligned} \frac{\partial L}{\partial x_i}(x) = 0, \quad i \in [1 : n], \quad \lambda_j^p p_j(x) = 0, \quad j \in [1 : k], \\ g_r(x) = 0, \quad r \in [1 : m] \end{aligned} \quad (1.1.7)$$

с $\lambda \neq 0$ и отобрать среди них те, для которых верны неравенства из (1.1.3) и (1.1.5). В (1.1.7) число неизвестных равно $n + m + k + 1$ (а именно: n компонент x и $m + k + 1$ множитель Лагранжа), а уравнений — на единицу меньше. Этот дефицит можно компенсировать, заметив, что соотношения (1.1.5), (1.1.6) инвариантны относительно умножения всех множителей Лагранжа на общую положительную константу. В силу последнего неравенства из (1.1.6) её всегда можно подобрать так, что после умножения

$$\lambda^0 + \sum_{j=1}^k \lambda_j^p + \sum_{i=1}^m |\lambda_i^g| = 1. \quad (1.1.8)$$

Таким образом, это уравнение можно добавить в систему (1.1.7) и сравнить число уравнений с числом неизвестных. Однако на практике так поступают редко. Во-первых, ввиду нехватки одного уравнения система (1.1.7) оставляет решению (x, λ) одну «степень свободы», и именно одна такая степень связана с обсуждавшимся умножением вектора λ на скаляр. В итоге можно надеяться (в типичных случаях оправданно), что из (1.1.7) предмет основного интереса x удастся определить с точностью до дискретного множества вариантов, а континуальная неоднозначность решения (x, λ) (ассоциированная с понятием «степень свободы») сконцентрируется на λ . Во-вторых, имеется более удобный и содержательный способ добавить уравнение, которым, однако, не всегда можно воспользоваться. К его обсуждению и перейдём.

Точку $x^0 \in D$ назовём *особой*, если для неё система (1.1.5), (1.1.6) имеет решение λ с $\lambda^\varphi = 0$. Очевидно, в *неособой точке локального минимума обязательно $\lambda^\varphi > 0$ и, более того, всегда можно обеспечить $\lambda^\varphi = 1$* за счёт деления всех множителей Лагранжа на λ^φ . В этом случае к (1.1.7) обычно добавляют уравнение $\lambda^\varphi = 1$.

Это, однако, некорректно, если точка экстремума является особой. Присмотримся повнимательнее к этой особенности. Очевидно, это свойство ограничений из (1.1.3) или, эквивалентно, области поиска минимума D ; минимизируемая функция $\varphi(\cdot)$ здесь ни при чём. Это подчёркивает и конкретизирует следующий результат.

Лемма 1.1.1. *Точка $x^0 \in D$ не является особой в том и только в том случае, когда выполнены следующие два условия:*

- i) *линейно независимы градиенты $\nabla g_1(x^0), \dots, \nabla g_m(x^0)$ функций в левых частях уравнений (1.1.3);*
- ii) *существует решение $h \in \mathbb{R}^n$ следующей линейной системы уравнений и неравенств:*

$$\langle \nabla g_1(x^0); h \rangle = \dots = \langle \nabla g_m(x^0); h \rangle = 0, \quad (1.1.9)$$

$$\langle \nabla p_j(x^0); h \rangle < 0 \quad \text{для критических индексов } j. \quad (1.1.10)$$

Для задачи без ограничений в виде равенств следует опустить условие i) и систему (1.1.9); а для задачи без ограничений в виде неравенств — условие ii).

Как хорошо известно, условие i) гарантирует, что множество M решений системы уравнений из (1.1.3) является гладким многообразием размерности $n - m$ в окрестности точки x^0 . При этом система (1.1.9) описывает касательное к M пространство: её решениями являются те и только

те векторы h , для которых

$$x(\varepsilon) := x^0 + \varepsilon h + o(\varepsilon) \in M \quad \forall \varepsilon \geq 0, \varepsilon \approx 0$$

при надлежащем выборе остаточного члена $o(\varepsilon)$ порядка выше первого. Дополнительные соотношения (1.1.10) означают, что параметрическую кривую $x(\varepsilon) \in M$ можно построить так, что на ней выполнены все ограничения из (1.1.3), причём неравенства — с «запасом» первого порядка:

$$p_j[x(\varepsilon)] \leq -\delta\varepsilon \quad \text{при некотором } \delta > 0.$$

(Для критических индексов это верно в силу (1.1.10), а для некритических по непрерывности ввиду строгого неравенства $p_j(x^0) < 0$.)

Таким образом, «неособость» точки x^0 означает, что вблизи неё равенства из постановки задачи определяют гладкое многообразие, а наложенные на x ограничения совместны в усиленном смысле.

Из изложенного следует, что, выполнив инструкции правила множителей Лагранжа, мы получаем:

- все особые точки множества D ;
- все точки, для которых (1.1.5), (1.1.6) верно с $\lambda^\varphi = 1$.

Приведём полезный пример неособой точки. Символом $J_{\text{cr}}(x^0)$ обозначим множество индексов, критических в точке $x^0 \in D$:

$$J_{\text{cr}}(x^0) := \{j \in [1 : k] : p_j(x^0) = 0\} \quad (J_{\text{cr}}(x^0) := \emptyset, \text{ если } k = 0).$$

Определение 1.1.1. Точку $x^0 \in D$ назовём регулярной, если линейно независима следующая совокупность градиентов:

$$\nabla g_i(x^0), i \in [1 : m], \quad \nabla p_j(x^0), j \in J_{\text{cr}}(x^0).$$

Лемма 1.1.2. Регулярная точка $x^0 \in D$ является неособой. Для регулярной точки локального минимума существует, причём единственный, набор множителей Лагранжа с $\lambda^\varphi = 1$, удовлетворяющий (1.1.5) и (1.1.6).

Нерегулярная неособая точка может иметь много таких наборов.

Как и в случае правила Ферма, поиск экстремума по правилу множителей Лагранжа может привести к целому множеству ответов, среди которых в общем случае могут быть и ложные. Для их отсекаания могут быть полезны необходимые условия экстремума второго порядка. Они указывают, какими свойствами обладают вторые производные функций из (1.1.3)

в точке экстремума, в то время как в (1.1.5) и (1.1.6) участвуют только производные первого порядка и значения функций в точке x^0 .

Для формулировки условий второго порядка потребуется ряд определений. Точку $x^0 \in D$, для которой существует набор множителей Лагранжа λ со свойствами (1.1.5) и (1.1.6), назовём *подозрительной на оптимальность*, а совокупность всех таких наборов λ , нормированных условием (1.1.8), обозначим через $\Lambda(x^0)$:

$$\Lambda(x^0) := \{\lambda : \text{выполнено (1.1.5), (1.1.6), (1.1.8)}\}. \quad (1.1.11)$$

Для подозрительной на оптимальность точки x^0 вектор $h \in \mathbb{R}^n$ и связанное с ним направление назовём *критическими*, если

$$\langle \nabla g_1(x^0); h \rangle = \dots = \langle \nabla g_m(x^0); h \rangle = 0, \quad (1.1.12)$$

$$\langle \nabla p_j(x^0); h \rangle \leq 0 \quad \forall j \in J_{\text{cr}}(x^0), \quad \langle \nabla \varphi(x^0); h \rangle \leq 0. \quad (1.1.13)$$

От (1.1.9), (1.1.10) система (1.1.12), (1.1.13) отличается тем, что в ней все неравенства нестрогие и добавлено ещё одно соотношение — последнее неравенство в (1.1.13). При этом в (1.1.13) некоторые неравенства обязаны выполняться как равенства: выбирая $\lambda \in \Lambda(x^0)$, умножая левую часть каждого соотношения из (1.1.12) и (1.1.13) на соответствующий множитель Лагранжа и замечая, что сумма всех этих произведений равна нулю в силу первого равенства из (1.1.5), убеждаемся в справедливости следующего факта.

Замечание 1.1.1. Следующие изменения в (1.1.12), (1.1.13) приводят к эквивалентной системе соотношений относительно h :

- замена неравенства $\langle \nabla p_j(x^0); h \rangle \leq 0$ равенством

$$\langle \nabla p_j(x^0); h \rangle = 0,$$

если для индекса $j \in J_{\text{cr}}(x^0)$ существует набор множителей $\lambda \in \Lambda(x^0)$, у которого $\lambda_j^p > 0$;

- замена неравенства $\langle \nabla \varphi(x^0); h \rangle \leq 0$ равенством

$$\langle \nabla \varphi(x^0); h \rangle = 0,$$

если существует такое $\lambda \in \Lambda(x^0)$, что $\lambda^\varphi > 0$.

Следствие 1.1.1. При выполнении условия i) леммы 1.1.1 хотя бы одно неравенство в (1.1.13) можно заменить равенством.

Для неособой точки x^0 в (1.1.13) неравенство $\langle \nabla \varphi(x^0); h \rangle \leq 0$ можно заменить равенством $\langle \nabla \varphi(x^0); h \rangle = 0$.

Введём обозначения для множества всех критических векторов:

$$\mathfrak{K}(x^0) := \{h : \text{выполнено (1.1.12) и (1.1.13)}\}$$

и максимума по множителям из $\Lambda(x^0)$ гессиана лагранжиана:

$$\Upsilon_{x^0}(h) := \max_{\lambda \in \Lambda(x^0)} h^\top L''(x^0 | \lambda) h. \quad (1.1.14)$$

Здесь максимум достигается, так как в силу (1.1.14) максимизируемое выражение линейно, а значит и непрерывно по λ , а множество $\Lambda(x^0)$ компактно в силу (1.1.11). Множество $\mathfrak{K}(x^0)$ очевидно является выпуклым конусом.

Теорема 1.1.1. Если точка $x^0 \in D$ доставляет локальный минимум в задаче (1.1.3), то множество $\Lambda(x^0)$ непусто и

$$\Upsilon_{x^0}(h) \geq 0 \quad \forall h \in \mathfrak{K}(x^0). \quad (1.1.15)$$

Здесь первая часть заключения $\Lambda(x^0) \neq \emptyset$ является переформулировкой условий Куна — Таккера и означает, что точка x^0 подозрительна на оптимальность. Вторая часть даёт необходимые условия экстремума второго порядка.

Замечание 1.1.2. Обратим внимание на следующее:

1. Приведённые условия первого порядка — (1.1.5), (1.1.6), а также условия второго порядка — (1.1.15) выполнены во всех точках $x^0 \in D$, где градиенты $\nabla g_1(x^0), \dots, \nabla g_m(x^0)$ линейно зависимы. Действительно, любой ненулевой набор коэффициентов $\lambda_1^g, \dots, \lambda_m^g$, обращающий линейную комбинацию этих градиентов в нуль, после дополнения нулями $\lambda_j^p := 0, \lambda^q := 0$ и умножения на общую константу ради (1.1.8) даёт набор $\lambda \in \Lambda(x^0)$. Одновременно $-\lambda \in \Lambda(x^0)$: и поэтому (1.1.15) тривиально выполнено для любого $h \in \mathbb{R}^n$. Таким образом, упомянутые условия полезны только для точек, где градиенты линейно независимы.
2. В (1.1.14) $\Lambda(x^0)$ можно заменить любым компактным множеством $\Upsilon'(x^0)$ со следующими свойствами:

- элементами $\Lambda'(x^0)$ являются удовлетворяющие (1.1.5), (1.1.6) наборы λ множителей Лагранжа;
 - любой набор $\lambda \in \Lambda(x^0)$ является положительным кратным $\lambda = \rho\lambda'$, $\rho > 0$ некоторого набора $\lambda' \in \Lambda'(x^0)$;
 - любой набор $\lambda' \in \Upsilon'(x^0)$ является положительным кратным $\lambda' = \rho'\lambda$, $\rho' > 0$ некоторого набора $\lambda \in \Lambda(x^0)$.
3. Для неособой точки x^0 предыдущее замечание позволяет в определении (1.1.14) множества $\Lambda(x^0)$ заменить условие нормировки (1.1.8) условием $\lambda^\varphi = 1$.

4. Поэтому в силу леммы 1.1.2 в случае регулярной точки x^0 множество $\Lambda(x^0)$ содержит единственный набор λ , а условие (1.1.15) упрощается, избавляясь от $\max$ из (1.1.14) и принимая вид

$$h^\top L''(x^0|\lambda)h \geq 0 \quad \forall h \in \mathfrak{K}(x^0). \quad (1.1.16)$$

5. Если в (1.1.13) все неравенства замещаются равенствами, конус $\mathfrak{K}(x^0)$ является линейным подпространством. Выбирая его базис $h_1, \dots, h_s$ и используя параметрическое описание

$$h = h(c) := c_1 h_1 + \dots + c_s h_s, c = \mathbf{col}(c_1, \dots, c_s) \in \mathbb{R}^s$$

этого подпространства, сводим условие (1.1.16) к неотрицательной определённости квадратичной формы $q(c)$ от переменных c_i

$$q(c) := h(c)^\top L''(x^0|\lambda)h(c) \geq 0 \quad \forall c \in \mathbb{R}^s.$$

Последнее свойство обслуживается стандартными алгебраическими тестами, например критерием Сильвестра. В соответствии с замечанием 1.1.1, рассматриваемый случай имеет место при пустоте множества критических индексов, в частности если в задаче (1.1.3) нет ограничений в виде неравенств.

6. Если в (1.1.13) все неравенства, за исключением одного, замещаются равенствами, конус $\mathfrak{K}(x^0)$ является замкнутым полупространством. Так как в (1.1.16) левая часть принимает одинаковые значения на векторах h и $-h$, само неравенство распространяется на все $h \in \mathfrak{L}(x^0) := \mathfrak{K}(x^0) \cup [-\mathfrak{K}(x^0)]$. В рассматриваемом случае $\mathfrak{L}(x^0)$ является линейным подпространством. Поэтому, как и в предыдущем пункте 5, условие (1.1.16) сводится к неотрицательной определённости некоторой квадратичной формы.

Как и в случае безусловной оптимизации, усиление необходимых условий второго порядка до строгого неравенства делает их достаточными, т. е. имеет место следующий факт.

Теорема 1.1.2. *Предположим, что неособая точка $x^0 \in D$ подозрительна на оптимальность $\Lambda(x^0) \neq \emptyset$ и удовлетворяет (1.1.15) в следующем усиленном смысле*

$$\Upsilon_{x^0}(h) > 0 \quad \forall h \in \mathfrak{K}(x^0), h \neq 0. \quad (1.1.17)$$

Тогда точка x^0 доставляет локальный минимум в задаче (1.1.3).

Пункты 2–6 замечания 1.1.2 с очевидными модификациями распространяются на (1.1.17).

1.1.3. Задачи выпуклого и линейного программирования.

Усиленный принцип оптимальности Лагранжа.

Теорема двойственности

Эти задачи имеют прежний вид (1.1.3), где, однако, все функции обладают следующими специальными свойствами и заданы на всём пространстве $\mathbb{R}^n$:

1. Левые части равенств $g_i(x) = 0$ аффинны по x , т. е.:

$$g_i(x) = \langle a_i; x \rangle + b_i, \quad \text{где } a_i \in \mathbb{R}^n, b_i \in \mathbb{R}. \quad (1.1.18)$$

2. Левые части неравенств $p_j(x) \leq 0$ и минимизируемый функционал $\varphi(x)$ — выпуклые функции в случае задачи выпуклого программирования и аффинные функции в случае задачи линейного программирования.

Задача линейного программирования — частный случай и задачи выпуклого программирования, и гладкой задачи математического программирования. Единственная причина, по которой задача выпуклого программирования может не быть частным случаем гладкой задачи математического программирования — негладкость какой-то из выпуклых функций $p_j(\cdot)$ или $\varphi(\cdot)$. Действительно, не все выпуклые функции гладкие, т. е. непрерывно дифференцируемые; простейший примером является норма $x \mapsto \|x\|$: в начале координат она не дифференцируема.

Теория выпуклого программирования уделяет значительное внимание негладким выпуклым функциям. В частности, для них развито обобщён-

ное дифференциальное исчисление, которое в определённых отношениях позволяет обращаться с такими функциями как с гладкими. Поскольку подобные обобщения не представляют особого интереса с точки зрения темы данной книги, для знакомства с ними мы отсылаем интересующегося читателя к специальной литературе [2; 7–9]. Соответственно, ограничимся обсуждением гладких задач выпуклого программирования.

Для них необходимые условия оптимальности из предыдущего раздела выполнены в усиленной форме. Общий смысл этих условий в их исходной форме в том, что надлежащим выбором множителей Лагранжа точке условного экстремума x^0 можно придать по отношению к лагранжиану свойства, тождественные или родственные свойствам его безусловного минимума. Для условий первого порядка — это обращение в нуль градиента в соответствии с правилом Ферма (1.1.1), т. е. выполняется первое уравнение из (1.1.5). Для условий второго порядка — это неотрицательная определённость гессиана (1.1.2). Однако здесь имеет место только аналог (1.1.14), (1.1.15) указанного свойства: соответствующая квадратичная форма неотрицательна не на всех векторах $h \in \mathbb{R}^n$, а в случаях, не охваченных п. 4 замечания 1.2, разным векторам h могут отвечать разные наборы множителей.

Усиление, о котором идёт речь, состоит в том, что точка условного экстремума доставляет безусловный минимум лагранжиана, а не просто разделяет с точками безусловного минимума некоторые общие свойства. Приведём точную формулировку этого результата, который назовём *усиленным принципом оптимальности Лагранжа*.

Теорема 1.1.3. *Если точка x^0 доставляет локальный минимум в задаче выпуклого программирования (1.1.3), то существует набор множителей Лагранжа λ со следующими свойствами:*

- i) *точка x^0 доставляет глобальный безусловный минимум соответствующего лагранжиана на всём пространстве;*
- ii) *выполнены соотношения*

$$\lambda^0 \geq 0, \quad \lambda_1^p \geq 0, \dots, \lambda_k^p \geq 0, \quad (1.1.19)$$

$$\lambda_j^p p_j(x^0) = 0 \quad \forall j \in [1 : k], \quad \lambda^0 + \sum_{j=1}^k \lambda_j^p + \sum_{i=1}^m |\lambda_i^g| > 0. \quad (1.1.20)$$

Они совпадают с условиями (1.1.5), (1.1.6) из правила множителей Лагранжа, за исключением первого уравнения в (1.1.5). Оно вытекает из условия i) в силу условия Ферма (1.1.1).

Если лагранжиан из теоремы 1.1.3 игнорирует минимизируемую функцию $\varphi(\cdot)$ в силу равенства $\lambda^\varphi = 0$, то сам результат всего лишь указывает на возможность достижения оптимума в специальных точках множества D из (1.1.3), которые универсальны в смысле независимости от $\varphi(\cdot)$. Особый интерес представляет противоположный случай: $\lambda^\varphi > 0$. Он имеет место при выполнении условий, перечисленных в следующем определении.

Определение 1.1.2. *Задачу выпуклого программирования назовём несингулярной при выполнении следующих двух условий:*

- i) *линейно независимы векторы $a_1, \dots, a_m$ из (1.1.18);*
- ii) *соотношения из (1.1.3) совместны в усиленном смысле: не только $D \neq \emptyset$, но более того, система соотношений из (1.1.3) имеет решение x после замены всех нестрогих неравенств соответствующими строгими неравенствами*

$$g_1(x) = \dots = g_m(x) = 0, \quad p_1(x) < 0, \dots, p_k(x) < 0. \quad (1.1.21)$$

Если в задаче нет ограничений в виде равенств, следует опустить условие i) и уравнения в (1.1.21), а при отсутствии неравенств — условие ii).

Условие i) тождественно условию i) леммы 1.1.1. Фактически оно означает, что в постановке задачи нет лишних уравнений, которые можно опустить, не меняя области D поиска экстремума. Действительно, пусть $D \neq \emptyset$ и i) нарушается. Тогда один из векторов a_i является линейной комбинацией остальных; за счёт перенумерации его можно сделать последним $a_m = \sum_{i=1}^{m-1} \eta_i a_i$. Скалярно умножая это равенство на любой вектор из D и вспоминая (1.1.18) и уравнения из (1.1.3), убеждаемся, что $b_m = \sum_{i=1}^{m-1} \eta_i b_i$. Поэтому

$$g_i(x) = 0 \quad \forall i \in [1 : m-1] \Rightarrow g_m(x) = 0,$$

т. е. последнее уравнение лишнее и его можно отбросить. Итерируя проделанное рассуждение, убеждаемся, что условие i) всегда можно обеспечить, избавляя постановку задачи от избыточных уравнений.

Следующий результат указывает полезные свойства несингулярных задач, одно из которых $\lambda^\varphi > 0$ было уже упомянуто.

Утверждение 1.1.1. *Для несингулярной задачи выпуклого программирования справедливы следующие утверждения:*

- i) *в теореме 1.1.3 можно $\lambda^\varphi \geq 0$ в (1.1.19) заменить на $\lambda^\varphi = 1$, а в (1.1.20) — опустить последнее соотношение;*

- ii) если набор множителей Лагранжа в некоторой точке $x^0 \in D$ обладает свойствами i) и ii) из теоремы 1.1.3, то $\lambda^\varphi > 0$.

Если для произвольной задачи выпуклого программирования и точки $x^0 \in D$ существует набор множителей Лагранжа, для которого верно (1.1.19), (1.1.20), уравнение $\nabla L(x^0|\lambda) = 0$ и неравенство $\lambda^\varphi > 0$, то эта точка является решением задачи, доставляя условный глобальный минимум.

Следовательно, для несингулярной задачи выпуклого программирования условия Куна — Таккера являются исчерпывающими, т.е. необходимыми и достаточными. Это деактуализирует условия второго порядка. Кроме того точки условного локального минимума доставляют условный глобальный минимум.

В несингулярном случае бывает полезна дополнительная характеристика набора λ с $\lambda^\varphi = 1$ и свойствами из теоремы 1.1.3. Для её формулировки заметим, что при $\lambda^\varphi = 1$ набор λ определён своей частью $\bar{\lambda} = \text{col}(\lambda_1^g, \dots, \lambda_m^g, \lambda_1^p, \dots, \lambda_k^p)$, причём

$$L(x|\lambda) = L(x|\bar{\lambda}) = \varphi(x) + \sum_{i=1}^m \lambda_i^g g_i(x) + \sum_{j=1}^k \lambda_j^p p_j(x).$$

Если $\lambda_j^p \geq 0 \forall j$, этот лагранжиан — выпуклая функция. Рассмотрим задачу его безусловной минимизации на всем пространстве,

$$L(x|\bar{\lambda}) \rightarrow \inf_{x \in \mathbb{R}^n}, \quad (1.1.22)$$

ограничившись нахождением величины инфимума

$$S(\bar{\lambda}) := \inf_{x \in \mathbb{R}^n} L(x|\bar{\lambda}).$$

Здесь возможно равенство $S(\bar{\lambda}) = -\infty$, причём функция $S(\bar{\lambda})$ является вогнутой (выпуклой вниз). Рассмотрим задачу её максимизации по всем наборам $\bar{\lambda}$ множителей Лагранжа с $\lambda_j^p \geq 0$:

$$S(\bar{\lambda}) \rightarrow \max_{\bar{\lambda}}, \quad \lambda_1^p \geq 0, \dots, \lambda_k^p \geq 0. \quad (1.1.23)$$

Здесь поиск максимума заведомо можно ограничить наборами, для которых $S(\bar{\lambda}) > -\infty$. Это условие вместе с неравенствами из (1.1.23) описывает выпуклое множество $\text{dom } S$. Задачу (1.1.23) называют *двойственной* по отношению к исходной задаче (1.1.3).