УДК 517.984.64

ОПТИМАЛЬНОЕ ВОССТАНОВЛЕНИЕ ЛИНЕЙНЫХ ФУНКЦИОНАЛОВ И ОПЕРАТОРОВ

Г. Г. Магарил-Ильяев

В лекциях излагаются начала теории оптимального восстановления значений линейных функционалов и операторов на классах множеств, элементы которых известны приближенно. Особое внимание уделяется восстановлению линейных функционалов, где теория существенно опирается на методы выпуклого анализа. Приводятся различные примеры, связанные, в основном, с задачами классической теории приближений. Лекции рассчитаны на широкий круг читателей.

Ключевые слова: оптимальное восстановление, оптимальный метод, экстремальная задача, выпуклая двойственность.

Лекция 1. Обсуждение возможных подходов к постановке задачи оптимального восстановления. Краткий исторический экскурс

На практике часто возникают задачи, связанные с восстановлением какой-либо характеристики объекта по информации (часто не полной и/или не точной) о других характеристиках этого объекта. К примеру, рассматривается задача о восстановлении функции или ее производной в точке, или интеграла от нее по информации о наборе ее значений в других точках, либо по приближенно заданному преобразованию Фурье, или требуется восстановить решение дифференциального уравнения по неточно известным начальным данным и так далее. Применяются различные подходы к решению подобного класса задач. Здесь мы следуем подходу, который предполагает наличие некоторой априорной информации об объекте, характеристики которого подлежат восстановлению. Это дает возможность поставить задачу о нахождении наилучшего метода восстановления данной характеристики среди всех возможных методов. Такой подход к задаче восстановления идеологически восходит к работам А. Н. Колмогорова 30-х гг. XX века, посвященным нахождению наилучших средств приближения для различных классов функций. Математическая теория, в которой рассматриваются задачи восстановления, основывающиеся на этом подходе, плодотворно развивается, начиная с 60-х гг. XX века.

Начнем с рассмотрения простого примера, показывающего, что постановка задачи оптимального восстановления вполне естественна. Пусть на отрезке $[a,b]$ заданы точки $t_1 < \ldots < t_n$, и в этих точках известны значения некоторой непрерывной функции $x(\cdot)$, т. е. известны числа $x(t_1),\ldots,x(t_n)$. Как по этим числам восстановить значение функции $x(\cdot)$ в точке $\tau \in [a,b]$, $\tau \neq t_i$, $i=1,\ldots,n$? Достаточно традиционная рекомендация состоит в том, что нужно провести через точки $x(t_1),\ldots,x(t_n)$ интерполяционный полином Лагранжа (напомним, что это полином степени $n-1$ вида $p(t) = \sum_{i=1}^{n} l_i(t)x(t_i)$, где l_i — полином степени $n-1$, принимающий значение 1 в точке t_i и нулевые значения в остальных точках t_j, $j \neq i$) и в качестве значения функции $x(\cdot)$ в точке τ взять $p(\tau)$.

Такая рекомендация предполагает, что функция $x(\cdot)$ ведет себя достаточно «хорошо», не может резко меняться. Если никакой априорной информации о функции нет, то задача, очевидно, бессмысленна, поскольку ясно, что существуют функции, принимающие в точках $t_1 < \ldots < t_n$ те же значения, что и $x(\cdot)$, а в точке τ — любое наперед заданное значение.

Будем теперь предполагать, что функция $x(\cdot)$ принадлежит некоторому множеству (классу) W в пространстве $C([a,b])$ непрерывных функций на $[a,b]$. Например, пусть W — это класс функций $x(\cdot)$, у которых первая производная $\dot{x}(\cdot)$ кусочно непрерывна и $|\dot{x}(t)| \leqslant 1$ в точках непрерывности $\dot{x}(\cdot)$. Тогда задача становится более осмысленной — значение функции $x(\cdot)$ в точке τ может находиться только в пределах синего отрезка (см. рис. 1.1).

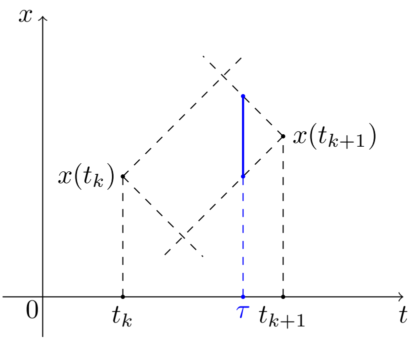

Рис. 1.1.

Итак, считаем, что функция $x(\cdot)$ принадлежит некоторому классу W, и о каждой функции $x(\cdot) \in W$ мы располагаем еще индивидуальной информацией, заключающейся в том, что нам известны значения $x(\cdot)$ в точках $t_1 < \ldots < t_n$. Пусть $I\colon C([a,b]) \to \mathbb{R}^n$ — линейный оператор, действующий по правилу $Ix(\cdot) = (x(t_1),\ldots,x(t_n))$. Тогда индивидуальная информация о функциях из W состоит в том, что нам известны значения этого оператора на W. Оператор I будем называть информационным оператором.

Заметим, что рекомендацию по поводу использования интерполяционного полинома Лагранжа можно интерпретировать так. Возьмите линейную функцию на $\mathbb{R}^n$ вида $\varphi(\xi_1,\ldots,\xi_n) = \sum_{i=1}^n l_i(\tau)\xi_i$ и вместо ξ_i подставьте $x(t_i)$, $i = 1,\ldots,n$. Мы пойдем дальше: возьмем произвольную функцию $\varphi(\xi_1,\ldots,\xi_n)$ на $\mathbb{R}^n$, подставим $x(t_i)$ вместо ξ_i, $i = 1,\ldots,n$, и будем считать, что это и есть оценка $x(\tau)$ для любой функции $x(\cdot) \in W$. Каждой такой функции (методу восстановления) сопоставим число — погрешность этого метода, а затем поставим вопрос о нахождении среди всех таких методов того, у которого погрешность минимальна.

Точная постановка такова. Пусть $\varphi\colon \mathbb{R}^n \to \mathbb{R}$. Сопоставим данной функции (методу восстановления) следующую величину:

$$e(\tau, W, I, \varphi) = \sup_{x(\cdot)\in W} \bigl|x(\tau) - \varphi(x(t_1),\ldots,x(t_n))\bigr|,$$

которую назовем погрешностью метода φ. Ясно, что это «наихудший» результат из всех, которые можно получить, используя данный метод восстановления.

Нас интересует величина

$$E(\tau, W, I) = \inf_{\varphi} e(\tau, W, I, \varphi),$$

где нижняя грань берется по всем отображениям $\varphi\colon \mathbb{R}^n \to \mathbb{R}$, и те методы $\widehat{\varphi}$, на которых эта нижняя грань достигается, т. е.

$$E(\tau, W, I) = e(\tau, W, I, \widehat{\varphi}).$$

Такие методы естественно назвать оптимальными.

Скажем несколько слов об истории вопроса такой постановки. В 1965 г. в кандидатской диссертации С. А. Смоляка [1] была поставлена следующая задача. Пусть X — линейное пространство,

W — класс элементов в X и l_i, $i = 0, 1, \ldots, n$, — линейные функционалы на X. Предполагается, что элементы множества W известны приближенно, а именно, о каждом $x \in W$ известен набор чисел $l_i(x)$, $i = 1, \ldots, n$, т. е. значения на W информационного оператора $I\colon X \to \mathbb{R}^n$, действующего по правилу $Ix = (l_1(x), \ldots, l_n(x))$. Имея эту информацию, мы хотим восстанавливать значения линейного функционала l_0 на W.

Дальше идут те же определения, что и выше. Каждому методу $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ сопоставляется его погрешность

$$e(l_0, W, I, \varphi) = \sup_{x \in W} \bigl| l_0(x) - \varphi(l_1(x), \ldots, l_n(x)) \bigr|,$$

определяется величина

$$E(l_0, W, I) = \inf_{\varphi\colon \mathbb{R}^n \to \mathbb{R}} e(l_0, W, I, \varphi),$$

и ставится вопрос о ее вычислении и нахождении тех методов $\widehat{\varphi}$, на которых нижняя грань достигается.

Отметим, что рассмотренный выше пример частный случай этой постановки, где $X = C([a,b])$, $l_0\colon x(\cdot) \to x(\tau)$ и $l_i\colon x(\cdot) \to x(t_i)$, $i = 1, \ldots, n$.

С. А. Смоляк доказал следующий результата (напомним, что множество $A \subset X$ выпукло, если из того, что $x, y \in A$ следует, что $(1-\alpha)x + \alpha y \in A$ для любого $\alpha \in (0,1)$, и что это множество центрально-симметричного, если $A = -A$).

Лемма 1.1 (Смоляка). *Если в задаче Смоляка множество W выпукло и центрально симметрично, то среди оптимальных методов есть линейный.*

Задача С. А. Смоляка и его лемма послужили началом развития теории оптимального восстановления линейных функционалов и операторов на множествах, элементы которых известно приближенно. Эта теория развивалась и обобщалась в различных направлениях. Значительное внимание уделялось задачам восстановления, в которых начальная информация об элементах W задана неточно (например, в приведенной постановке числа $l_i(x)$, $i = 1, \ldots, n$, известны приближенно) и, быть может, бесконечномерна. Для таких задач находились условия существования линейного оптимального метода, т. е. справедливость для них аналога леммы Смоляка.

Окончательный результат — критерий существования линейного оптимального метода восстановления для достаточно общей постановки задачи оптимального восстановления линейного функционала — получен в работе [2]. Помимо этого, развивались некоторые теоретические аспекты теории восстановления линейных операторов, было решено значительное количество конкретных задач, связанных с нахождением оптимальных методов восстановления (см., например, обзоры [3–5], монографию [6], статьи [7–12]). Подход к нахождению оптимальных методов восстановления линейных функционалов по неточным исходным данным с позиций теории экстремальных задач впервые был предложен в работе [13].

Лекция 2. Постановка общей задачи оптимального восстановления. Начальные сведения из выпуклого анализа

В приведенных в конце лекции 1 обзорах и статьях была поставлена общая задача восстановления, а именно, задача об оптимальном восстановлении значений линейного оператора на классе элементов по неточной и неполной информации о самих элементах. Ее постановка такова.

Пусть X — линейное пространство, Y и Z — нормированные пространства, $\Lambda\colon X \to Z$ и $I\colon X \to Y$ — линейные операторы, W — некоторое непустое множество (класс) элементов из X и $\delta \geqslant 0$. Рассматривается задача о восстановлении значений оператора Λ на множестве W по следующей информации об элементах этого множества: о любом элементе $x \in W$ нам известен элемент $y \in Y$ такой, что $\|Ix - y\|_Y \leqslant \delta$ (если $\delta > 0$, то мы говорим об информации, заданной неточно, а если $\delta = 0$, то известен элемент Ix, и мы говорим об информации, заданной точно). Под методами восстановления понимаются произвольные отображения $\varphi\colon Y \to Z$.

Следующая диаграмма иллюстрирует действия введенных отображений.

$$\begin{tikzcd} W \subset X \arrow[rr, "\Lambda"] \arrow[dr, "I"'] & & Z \\ & Y \arrow[ur, "\varphi"'] & \end{tikzcd}$$

Рис. 2.1.

Погрешностью метода φ называется величина

$$e(\Lambda, W, I, \delta, \varphi) = \sup_{\substack{x\in W,\, y\in Y,\\ \|Ix-y\|_Y\leqslant\delta}} \big\|\Lambda x - \varphi(y)\big\|_Z\,.$$

Если $\delta = 0$, то эта погрешность, очевидно, принимает такой вид:

$$e(\Lambda, W, I, 0, \varphi) = \sup_{x\in W} \big\|\Lambda x - \varphi(Ix)\big\|_Z\,.$$

Нас интересуют величина

$$E(\Lambda, W, I, \delta) = \inf_{\varphi} e(\Lambda, W, I, \delta, \varphi),$$

где нижняя грань берется по всем отображениям (методам) $\varphi\colon Y \to Z$, которая называется *погрешностью оптимального восстановления*, и те методы $\widehat{\varphi}$, на которых нижняя грань достигается, т. е.

$$E(\Lambda, W, I, \delta) = e(\Lambda, W, I, \delta, \widehat{\varphi}).$$

Такие методы будем называть *оптимальными методами* восстановления (значений оператора Λ на W по данной информации).

Рассмотрим экстремальную задачу

$$\|\Lambda x\|_Z \to \max, \quad \|Ix\|_Y \leqslant \delta, \quad x \in W, \tag{2.1}$$

заключающуюся в нахождении тех допустимых (т. е. удовлетворяющих ограничениям задачи) элементов $\widehat{x}$, на которых максимизируемый функционал достигает своего максимального значения. Обозначим через $S(\Lambda, W, I, \delta)$ *значением* задачи (2.1), т. е. точную верхнюю грань максимизируемого функционала на допустимых элементах.

Предложение 2.1. *Если множество W центрально симметрично, то*

$$E(\Lambda, W, I, \delta) \geqslant S(\Lambda, W, I, \delta). \tag{2.2}$$

$\lhd$ Пусть x_0 — допустимый элемент в (2.1), тогда элемент $-x_0$ также допустим, и мы имеем для любого метода $\varphi\colon Y \to Z$

$$\begin{aligned}2\|\Lambda x_0\|_Z &= \|\Lambda x_0 - \varphi(0) - (\Lambda(-x_0) - \varphi(0))\|_Z \leqslant \|\Lambda x_0 - \varphi(0)\|_Z + \\ &+ \|\Lambda(-x_0) - \varphi(0)\|_Z \leqslant 2 \sup_{x\in W,\, \|Ix\|_Y\leqslant\delta} \|\Lambda x - \varphi(0)\|_Z \leqslant \\ &\leqslant 2 \sup_{\substack{x\in W,\, y\in Y,\\ \|Ix-y\|_Y\leqslant\delta}} \|\Lambda x - \varphi(y)\|_Z = 2e(\Lambda, W, I, \delta, \varphi).\end{aligned}$$

Переходя слева к верхней грани по всем допустимым x в задаче (2.1), получаем, что

$$S(\Lambda, W, I, \delta) \leqslant e(\Lambda, W, I, \delta, \varphi).$$

Теперь, переходя справа к нижней грани по всем методам φ, приходим к неравенству (2.2). $\triangleright$

Задача (2.1), значение которой дает оценку снизу для погрешности оптимального восстановления, внешне вполне обозрима. Как будет видно из дальнейшего, для многих частных случаев можно найти ее решение, используя стандартные методы теории экстремума. Поэтому важен вопрос о взаимоотношении задачи (2.1) и исходной задачи оптимального восстановления. Когда неравенство (2.2) превращается в равенство? Что можно сказать в этом случае о существовании оптимальных методов восстановления?

Если оператор Λ — линейный функционал ($Z = \mathbb{R}$), а W — выпуклое множество, то задача (2.1) является выпуклой. Исследование таких задач представляет собой одну из основных частей выпуклого анализа — раздела математики, где изучают выпуклые множества, выпуклые функции и выпуклые экстремальные задачи (см. [14]). Важнейшим явлением, сопутствующим выпуклости, является феномен двойственности, о котором сейчас и будет рассказано.

Пусть X — вещественное векторное пространство. Напомним, что непустое множество $A \subset X$ называется *выпуклым,* если вместе с любыми двумя своими точками x и y оно содержит отрезок $[x, y] = \{z \in X : z = (1-\alpha)x + \alpha y, 0 \leqslant \alpha \leqslant 1\}$, соединяющий точки x и y.

Пустое множество выпукло по определению.

Следующие свойства выпуклых множеств непосредственно следуют из определений.

Если $\{A_i\}_{i \in \mathscr{I}}$ — произвольное семейство выпуклых подмножеств X, то $\bigcap_{i \in \mathscr{I}} A_i$ — выпуклое множество.

Если $\{A_i\}_{i=1}^n$ — конечное семейство выпуклых подмножеств X, то их алгебраическая сумма $A_1 + \ldots + A_n := \{x \in X : x = x_1 + \ldots + x_n, x_i \in A_i, i = 1, \ldots, n\}$ — выпуклое множество.

Если A — выпуклое подмножество X и $\lambda \in \mathbb{R}$, то множество $\lambda A := \{\lambda x : x \in A\}$ выпукло.

Пусть теперь X — нормированное пространство и X^* — его сопряженное. Значение линейного функционала $x^* \in X^*$ на элементе X обозначаем $\langle x^*, x \rangle$.

Пусть $x^* \in X^*$, $x^* \neq 0$ и $\gamma \in \mathbb{R}$. Множество $H = H(x^*, \gamma) = \{x \in X : \langle x^*, x\rangle = \gamma\}$ называется *гиперплоскостью*. Гиперплоскость порождает два полупространства $H_+(x^*, \gamma) = \{x \in X : \langle x^*, x\rangle \leqslant \gamma\}$ и $H_-(x^*, \gamma) = \{x \in X : \langle x^*, x\rangle \geqslant \gamma\}$.

Ясно, что гиперплоскости и полупространства — выпуклые замкнутые множества.

Пусть A и B — непустые подмножества X. Говорят, что эти множества *отделимы*, если существует такая гиперплоскость, что A и B принадлежат различным полупространствам, порожденным этой гиперплоскостью.

Данное геометрическое определение отделимости равносильно, очевидно, следующему аналитическому: множества A и B отделимы, если существует такой ненулевой элемент $x^* \in X^*$, что

$$\sup_{a \in A} \langle x^*, a\rangle \leqslant \inf_{b \in B} \langle x^*, b\rangle.$$

Если неравенство строгое, то говорят, что множества A и B *строго отделимы*.

Нам понадобится одна теорема отделимости, которую обычно называют *второй теоремой отделимости*.

Теорема 2.1 (вторая теорема отделимости). *Пусть A — непустое выпуклое замкнутое подмножество X и $b \notin A$. Тогда множество A и точка b строго отделимы.*

В основе выпуклой двойственности лежит то обстоятельство, что выпуклость можно определить двояко: непосредственно (как сделано выше) и используя двойственное (сопряженное) пространство. Это можно оформить различными способами. Например, в виде следующего утверждения.

Предложение 2.2. *Множество $A \subset X$ выпукло и замкнуто тогда и только тогда, когда оно есть пересечение всех полупространств его содержащих.*

$\lhd$ Пусть A — замкнутое выпуклое подмножество X и пусть A_1 есть пересечение всех полупространств, содержащих A. Ясно, что множество A_1 выпукло замкнуто и $A \subset A_1$. Покажем, что $A = A_1$. Если это не так, то найдется элемент $x_0 \in A_1 \setminus A$. По теореме отделимости существует ненулевой функционал $x^* \in X^*$ такой, что

$$\langle x^*, x_0\rangle < \inf_{x \in A} \langle x^*, x\rangle.$$

Отсюда следует, что $A \subset H_-(x^*, \gamma)$, где $\gamma = \inf_{x \in A} \langle x^*, x \rangle$. Поскольку $x_0 \in A_1$, то этот элемент должен также принадлежать $H_-(x^*, \gamma)$ согласно определению A_1. Но это противоречит полученному неравенству. Следовательно, $A = A_1$.

Если $A = A_1$, то A выпукло и замкнуто как пересечение выпуклых замкнутых множеств. $\rhd$

Лекция 3. Вопросы двойственности выпуклых множеств, функций и экстремальных задач

В конце предыдущей лекции было доказано предложение 2.2. Наиболее интересен и полезен аналог этого утверждения для выпуклых функций. Напомним некоторые определения.

Пусть $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ — расширенная прямая, т. е. прямая, пополненная символами $+\infty$, продолжающими естественное отношение порядка: $a \leqslant +\infty$, $a \in \mathbb{R}$. Кроме того, предполагается, что $a + \infty = +\infty$ для всех $a \in \mathbb{R}$, $a(+\infty) = +\infty$, если $a > 0$ и $+\infty + \infty = +\infty$.

С каждой функцией $f\colon X \to \overline{\mathbb{R}}$ свяжем два множества

$$\operatorname{dom} f = \bigl\{x \in X : \ f(x) < +\infty\bigr\}$$

и

$$\operatorname{epi} f = \bigl\{(x, \alpha) \in X \times \mathbb{R} : \ \alpha \geqslant f(x), x \in \operatorname{dom} f\bigr\},$$

которые называются соответственно *эффективным множеством* и *надграфиком* (или *эпиграфом*) функции f.

Функция $f\colon X \to \overline{\mathbb{R}}$ называется *выпуклой*, если ее надграфик — выпуклое множество в $X \times \mathbb{R}$.

Вот примеры выпуклых функций на прямой: $x \mapsto ax^2 + bx + c$, $a \geqslant 0$; $x \mapsto e^{\alpha x}$, $\alpha \in \mathbb{R}$; $x \mapsto |x|^p$, $p \geqslant 1$; $x \mapsto -\ln x$, если $x > 0$ и $+\infty$, если $x \leqslant 0$; $x \mapsto x \log_2 x + (1 - x)\log_2(1 - x)$, если $0 < x < 1$ и $+\infty$ в остальных случаях.

Легко проверить, что функция $f\colon X \to \overline{\mathbb{R}}$ выпукла тогда и только тогда, когда для любых $x_1, x_2 \in X$ и любых $\alpha_i > 0$, $i = 1, 2$, таких, что $\alpha_1 + \alpha_2 = 1$ справедливо неравенство

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leqslant \alpha_1 f(x_1) + \alpha_2 f(x_2),$$

которое называется *неравенством Иенссена.*

Функция $f\colon X \to \overline{\mathbb{R}}$ называется *замкнутой*, если $\operatorname{epi} f$ — замкнутое множество в $X \times \mathbb{R}$.

Легко проверить, что сумма конечного числа выпуклых функций — выпуклая функция и что, если f_i, $i \in \mathscr{J}$, — произвольное семейство выпуклых функций, то функция $f(x) = \sup_{i\in\mathscr{J}} f_i(x)$ выпукла (это равносильно равенству $\operatorname{epi} f = \bigcap_{i\in\mathscr{J}} \operatorname{epi} f_i$, которое легко проверяется и из которого следует, что если f_i — замкнутые функции, то f — замкнутая функция).

Пусть $f\colon X \to \overline{\mathbb{R}}$ — замкнутая выпуклая функция, т. е. $\operatorname{epi} f$ — выпуклое замкнутое множество в $X \times \mathbb{R}$. Согласно предложению 2.2 это множество есть пересечение всех полупространств, его содержащих. Но гиперплоскости, определяющие эти полупространства суть графики аффинных функций $x \mapsto \langle x^*, x\rangle + \alpha$ (см. рис. 3.1), не превосходящих функцию f.

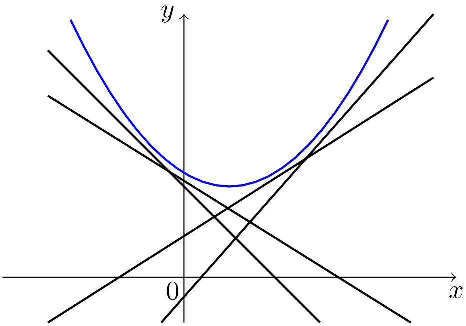

Рис. 3.1.

Как видно из рис. 3.1 (подробного доказательства приводить не будем), это означает, что функция f есть поточечная верхняя грань таких аффинных функций.

Придадим теперь этому утверждению аналитическую форму. Для этого введем некоторые понятия.

Пусть $f\colon X \to \overline{\mathbb{R}}$. Функция $f^*\colon X^* \to \overline{\mathbb{R}}$, определенная равенством

$$f^*(x^*) = \sup_{x\in X}(\langle x^*, x\rangle - f(x)),$$

называется *сопряженной функцией* к f или *преобразованием Лежандра — Юнга — Фенхеля*, а функция $f^{**}\colon X \to \overline{\mathbb{R}}$, определенная по правилу

$$f^{**}(x) = \sup_{x \in X^*} (\langle x^*, x \rangle - f^*(x^*)),$$

называется *второй сопряженной функцией к* f.

Ясно, что функции f^* и f^{**} выпуклы и замкнуты, как верхние грани аффинных функций. Из определений легко следует, что всегда

$$f^{**}(x) \leqslant f(x), \quad \forall\, x \in X.$$

Действительно, $f^*(x^*) = \sup_{x \in X}(\langle x^*, x \rangle - f(x)) \geqslant \langle x^*, x \rangle - f(x)$ для любого $x \in X$. Следовательно, $f(x) \geqslant \langle x^*, x \rangle - f^*(x^*)$ для всех $x \in X$ и $x^* \in X^*$. Но тогда $f^{**}(x) = \sup_{x \in X^*}(\langle x^*, x \rangle - f^*(x^*)) \leqslant f(x)$.

Теорема 3.1 (Фенхеля — Моро)**.** *Функция* $f\colon X \to \overline{\mathbb{R}}$ *выпукла и замкнута тогда и только тогда, когда* $f^{**} = f$.

$\triangleleft$ Если $f^{**} = f$, то функция f выпукла и замкнута, так как таковой является f^{**}. Пусть f выпукла и замкнута. Если $f(x) \equiv +\infty$, то, очевидно, $f^{**}(x) \equiv +\infty$. Пусть $f \not\equiv +\infty$. Согласно предыдущему существует такая аффинная функция $x \mapsto \langle x^*, x \rangle - \alpha$, что $\langle x^*, x \rangle - \alpha \leqslant f(x)$ для всех $x \in X$ или, равносильно, $\alpha \geqslant \sup_{x \in X}(\langle x^*, x \rangle - f(x)) = f^*(x^*)$. Так как f — верхняя грань таких функций, то для всех $x \in X$

$$f(x) = \sup_{\substack{x^* \in X^*,\, \alpha \in \mathbb{R}, \\ \alpha \geqslant f^*(x^*)}} (\langle x^*, x \rangle - \alpha). \tag{3.1}$$

Поскольку функция f не равна тождественно $+\infty$, то в (3.1) вместо α можно взять $f^*(x^*)$. Следовательно, $f(x) = \sup_{x^* \in X^*}(\langle x^*, x \rangle - f^*(x^*)) = f^{**}(x)$ для всех $x \in X$. $\triangleright$

Двойственность экстремальных задач. Применим теорему Фенхеля — Моро к построению двойственных задач. Пусть $f_0\colon X \to \overline{\mathbb{R}}$ и C — непустое подмножество X. Рассмотрим задачу

$$f_0(x) \to \min, \quad x \in C.$$

Определим так называемую *индикаторную функцию* δ множества C по правилу: $\delta C(x) = 0$, если $x \in C$, и $\delta C(x) = +\infty$, если $x \notin C$, и положим $f(x) = f_0(x) + \delta C(x)$, $x \in X$. Тогда выписанная задача, очевидно, равносильна следующей задаче:

$$f(x) \to \min, \quad x \in X. \tag{3.2}$$

Включим ее в серию «подобных» ей задач (или, как говорят, «возмутим» данную задачу). Пусть Y — другое нормированное пространство и функция $F\colon X \times Y \to \overline{\mathbb{R}}$ такова, что $F(x,0) = f(x)$ для всех $x \in X$. Каждому $y \in Y$ сопоставим задачу

$$F(x,y) \to \min, \quad x \in X. \tag{3.3}$$

Семейство таких задач называется *возмущением задачи* (3.2), а функция $S\colon Y \to \overline{\mathbb{R}}$, сопоставляющая $y \in Y$ значение задачи (3.3), называется S*-функцией* данного семейства. Ясно, $S(0)$ — значение исходной задачи (3.2).

Как уже было отмечено, $S^{**}(0) \leqslant S(0)$, а если S-функция выпукла и замкнута (отметим, что несложная проверка показывает, что S-функция выпукла, если функция F выпукла), то по теореме Фенхеля — Моро $S^{**}(0) = S(0)$. Выпишем задачу, значением которой является величина $S^{**}(0)$. По определению $S^{**}(0) = \sup_{y^* \in Y^*}(-S^*(y^*))$. Далее

$$\begin{aligned} S^*(y^*) &= \sup_{y \in Y}(\langle y^*, y\rangle - \inf_{x \in X} F(x,y)) = \\ &= \sup_{x \in X,\, y \in Y}(\langle 0, x\rangle + \langle y^*, y\rangle - F(x,y)). \end{aligned}$$

Справа стоит сопряженная функция к F в точке $(0, y^*)$. Таким образом, задача, значение которой равно $S^{**}(0)$ имеет вид

$$-F^*(0,y^*) \to \max, \quad y^* \in Y^*. \tag{3.4}$$

Задача (3.4) называется *двойственной задачей к* (3.2) (относительно заданного возмущения), и значение ее не больше значения задачи (3.2).

Если исходная задача (3.2) на максимум, то двойственной к ней задачей естественно считать задачу

$$F^*(0,y^*) \to \min, \quad y^* \in Y^*.$$

Здесь значение двойственной задачи, как легко понять, не меньше значения исходной задачи.

Лекция 4. Двойственные задачи к задаче линейного программирования и к задаче оптимального восстановления линейного функционала

Найдем двойственные к задаче линейного программирования и задаче оптимального восстановления линейного функционала.

4.1. Задача линейного программирования. Отметим сначала, что элементы пространства $\mathbb{R}^n$ мы рассматриваем как вектор-столбцы. Пространство $(\mathbb{R}^n)^*$ — сопряженное к $\mathbb{R}^n$ — можно отождествить с самим $\mathbb{R}^n$, где элементы суть вектор-строки. Если функционал (вектор-строка) $a^* = (a_1, \ldots, a_n) \in (\mathbb{R}^n)^*$, а $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ (T — знак транспонирования), то значение a^* на x — это $\langle a^*, x\rangle = \sum_{i=1}^n a_i x_i$.

Пусть $c^* \in (\mathbb{R}^n)^*$, A — матрица размера $m \times n$ и $b \in \mathbb{R}^m$. Рассмотрим следующую задачу линейного программирования:

$$\langle c^*, x\rangle \to \min, \quad Ax \geqslant b, \quad x \geqslant 0.^1 \tag{4.1}$$

В терминах общей постановки из предыдущей лекции, здесь $f(x) = \langle c^*, x\rangle$, когда $Ax \geqslant b$, $x \geqslant 0$ и $f(x) = +\infty$ в противном случае.

Выпишем двойственную задачу к (4.1) относительно возмущения

$$\langle c^*, x\rangle \to \min, \quad Ax \geqslant b + y, \quad x \geqslant 0,$$

где $y \in \mathbb{R}^m$ (т. е. $F(x, y) = \langle c^*, x\rangle$, когда $Ax \geqslant b + y$, $x \geqslant 0$ и $F(x, y) = +\infty$ в противном случае). Имеем

$$F^*(0, y^*) = \sup_{x \in \mathbb{R}^n,\, y \in \mathbb{R}^m} (\langle y^*, y\rangle - F(x, y))$$

Если $F(x, y) = +\infty$, то $-F(x, y) = -\infty$, и этот случай нам не интересен, поскольку нас интересует верхняя грань по $x \in \mathbb{R}^n$ и $y \in \mathbb{R}^m$. Таким образом,

$$F^*(0, y^*) = \sup_{\substack{Ax \geqslant b + y,\\ x \geqslant 0,\, y \in \mathbb{R}^m}} (\langle y^*, y\rangle - \langle c^*, x\rangle).$$

Если хотя бы одна компонента вектора y^* отрицательна, то можно взять вектор y, удовлетворяющий неравенству $Ax \geqslant b + y$ и такой, что у него соответствующая компонента также отрицательна

[1] Неравенства понимаются покоординатно.

и сколь угодно большая по модулю. Следовательно, в этом случае $F^*(0, y^*) = +\infty$. Если $y^* \geqslant 0$, то надо подставить $Ax - b$ вместо y и тогда мы имеем

$$F^*(0, y^*) = \begin{cases} \sup\limits_{x \geqslant 0}(\langle y^*, Ax - b\rangle - \langle c^*, x\rangle), & \text{если} \ \ y^* \geqslant 0; \\ +\infty, & \text{если не так.} \end{cases}$$

Далее

$$\sup_{x \geqslant 0}(\langle y^* A - c^*, x\rangle - \langle y^*, b\rangle) = \begin{cases} -\langle y^*, b\rangle, & \text{если} \ \ y^* A - c^* \leqslant 0; \\ +\infty, & \text{если не так.} \end{cases}$$

Таким образом,

$$F^*(0, y^*) = \begin{cases} -\langle y^*, b\rangle, & \text{если} \ \ y^* A \leqslant c^*, \ y^* \geqslant 0; \\ +\infty, & \text{если не так,} \end{cases}$$

и следовательно, двойственная задача, обозначая $\lambda = y^*$, имеет вид

$$\langle \lambda, b\rangle \to \max, \quad \lambda A \leqslant c^*, \quad \lambda \geqslant 0. \tag{4.2}$$

Это, очевидно, снова задача линейного программирования.

4.2. Двойственная к задаче оптимального восстановления линейного функционала. Задача (2.1) из второй лекции в случае, когда $\Lambda = x^*$ — линейный функционал на X (и тем самым $Z = \mathbb{R}$) и множество W центрально симметрично, имеет вид

$$\langle x^*, x\rangle \to \max, \quad \|Ix\|_Y \leqslant \delta, \quad x \in W. \tag{4.3}$$

Найдем двойственную к этой задаче относительно возмущения

$$\langle x^*, x\rangle \to \max, \quad \|Ix - y\|_Y \leqslant \delta, \quad x \in W.$$

В соответствии с общей постановкой здесь $f(x) = \langle x^*, x\rangle$, если $\|Ix\|_Y \leqslant \delta$, $x \in W$, и $f(x) = +\infty$ в противном случае, а $F(x, y) = \langle x^*, x\rangle$, если $\|Ix - y\|_Y \leqslant \delta$, $x \in W$, и $F(x, y) = +\infty$ в противном случае.

Так как здесь задача на максимум, то двойственная задача имеет вид

$$F^*(0, y^*) \to \min, \quad y^* \in Y^*.$$

По определению (учитывая, что множество $\{(x,y) \in X \times Y : x \in W,\ \|Ix-y\|_Y \leqslant \delta\}$ центрально симметрично)

$$F^*(0,y^*) = \sup_{x\in X,\, y\in Y} (\langle y^*,y\rangle - F(x,y)) =$$
$$= \sup_{\substack{x\in W,\, y\in Y,\\ \|Ix-y\|_Y\leqslant\delta}} (\langle y^*,y\rangle - \langle x^*,x\rangle) = \sup_{\substack{x\in W,\, y\in Y,\\ \|Ix-y\|_Y\leqslant\delta}} |\langle x^*,x\rangle - \langle y^*,y\rangle|,$$

т. е. двойственная задача такова:

$$\sup_{\substack{x\in W,\, y\in Y,\\ \|Ix-y\|_Y\leqslant\delta}} \bigl|\langle x^*,x\rangle - \langle y^*,y\rangle\bigr| \to \min, \quad y^*\in Y^*. \tag{4.4}$$

Значение этой задачи есть погрешность оптимального восстановления функционала x^* на классе W по информации I, но не по всем методам $\varphi\colon Y \to \mathbb{R}$, а только по линейным. Поскольку задача (4.3) на максимум, то ее значение не меньше значения задачи (4.4), и мы получили аналог неравенства (2.2) для данного случая.

Можно показать, что задача (4.3) является двойственной к задаче (4.4) относительно некоторого возмущения, так что задачи (4.3) и (4.4) двойственны друг к другу.

Для того, чтобы описать как связаны между собой решение задачи (2.1) и оптимальный метод в соответствующей задаче восстановления линейного функционала, нам понадобятся некоторые дополнительные сведения из выпуклого анализа.

4.3. Выпуклые задачи. Введем понятие субдифференциала для функций на нормированном пространстве X, которое является аналогом производной для выпуклых функций. Пусть выпуклая функция $f\colon X \to \overline{\mathbb{R}}$ конечна в точке $\widehat{x}$. *Субдифференциалом* функции f в точке $\widehat{x}$ называется множество (возможно пустое)

$$\partial f(\widehat{x}) = \bigl\{x^*\in X^* : f(x) - f(\widehat{x}) \geqslant \langle x^*, x-\widehat{x}\rangle,\ \forall\, x\in X\bigr\}.$$

Следующее предложение показывает, что субдифференциал достаточно естественное обобщение понятия производной на выпуклые функции.

Предложение 4.1. *Пусть выпуклая функция $f\colon X \to \overline{\mathbb{R}}$ дифференцируема в точке $\widehat{x}$. Тогда $\partial f(\widehat{x}) = \{f'(\widehat{x})\}$.*

$\lhd$ Пусть $x \in X$. Для любого $0 < \alpha < 1$ имеем по неравенству Иенссена: $f((1-\alpha)\widehat{x} + \alpha x) \leqslant (1-\alpha)f(\widehat{x}) + \alpha f(x)$, откуда

$$\frac{f(\widehat{x} + \alpha(x - \widehat{x})) - f(\widehat{x})}{\alpha} \leqslant f(x) - f(\widehat{x}).$$

В силу дифференцируемости f в точке $\widehat{x}$

$$\frac{f(\widehat{x} + \alpha(x - \widehat{x})) - f(\widehat{x})}{\alpha} = \frac{f(\widehat{x}) + \alpha\langle f'(\widehat{x}), x - \widehat{x}\rangle + o(\alpha) - f(\widehat{x})}{\alpha}.$$

Переходя здесь к пределу при $\alpha \to 0$ и учитывая предыдущее неравенство, получаем, что $\langle f'(\widehat{x}), x - \widehat{x}\rangle \leqslant f(x) - f(\widehat{x})$, т. е. $f'(\widehat{x}) \in \partial f(\widehat{x})$.

Обратно, если $x^* \in \partial f(\widehat{x})$, то для любого $x \in X$ и любого $t > 0$ имеем $f(\widehat{x} + tx) - f(\widehat{x}) \geqslant t\langle x^*, x\rangle$. Следовательно,

$$\frac{f(\widehat{x} + tx) - f(\widehat{x})}{t} \geqslant \langle x^*, x\rangle,$$

и аналогично предыдущему, отсюда следует, что $\langle f'(\widehat{x}), x\rangle \geqslant \langle x^*, x\rangle$ для любого x и значит, $x^* = f'(\widehat{x})$. $\rhd$

Приведем два примера вычисления субдифференциалов функций, которые нам понадобятся ниже. Первый пример — это норма в X, т. е. функция $x \mapsto \|x\|_X$. Нетрудно проверить, что

$$\partial\|0\|_X = BX^* = \big\{x^* \in X^* : \|x^*\|_{X^*} \leqslant 1\big\}$$

и если $x \neq 0$, то

$$\partial\|x\|_X = \big\{x^* \in X^* : \|x^*\|_{X^*} = 1,\ \langle x^*, x\rangle = \|x\|_X\big\}.$$

Второй пример — это индикаторная функция δ подпространства $L \subset X$, т. е. напомним, $\delta L(x) = 0$, если $x \in L$ и $\delta L(x) = +\infty$, если $x \neq L$. В этом случае, если $\widehat{x} \in L$, то

$$\partial\delta L(\widehat{x}) = L^{\perp},$$

где $L^{\perp} = \{x^* \in X^* : \langle x^*, x\rangle = 0, \forall\, x \in L\}$ — аннулятор L. Действительно, по определению

$$\partial\delta L(\widehat{x}) = \big\{x^* \in X^* : \delta L(x) - \delta L(\widehat{x}) \geqslant \langle x^*, x - \widehat{x}\rangle,\ \forall\, x \in X\big\}.$$

Для $x \in L$ имеем $\langle x^*, x - \widehat{x}\rangle \leqslant 0$. Пусть $x = 2\widehat{x}$, тогда $\langle x^*, \widehat{x}\rangle \leqslant 0$, и тем самым $\langle x^*, x\rangle \leqslant \langle x^*, \widehat{x}\rangle \leqslant 0$ для любого $x \in L$. Взяв $-x$, получим противоположное неравенство и значит, $\partial\delta L(\widehat{x}) \subset L^{\perp}$. Обратно, пусть $x^* \in L^{\perp}$. Если $x \notin L$, то неравенство $\delta L(x) - \delta L(\widehat{x}) \geqslant \langle x^*, x - \widehat{x}\rangle$ очевидным образом выполняется. Если $x \in L$, то $\delta L(x) - \delta L(\widehat{x}) = 0 = \langle x^*, x - \widehat{x}\rangle$ (так как $x - \widehat{x} \in L$) и тем самым $L^{\perp} \subset \partial\delta L(\widehat{x})$.

Лекция 5. Теорема Каруша — Куна — Таккера в субдифференциальной форме и существование оптимальных методов

Перейдем теперь к рассмотрению вопросов, связанных с необходимыми и достаточными условиями минимума для выпуклых задач. Пусть X — вещественное нормированное пространство и $f\colon X \to \overline{\mathbb{R}}$ — выпуклая функция. Рассмотрим задачу

$$f(x) \to \min, \quad x \in X, \tag{5.1}$$

которая называется *выпуклой задачей без ограничений.*

Отметим, что в этой задаче нет смысла говорить о локальных минимумах, поскольку любой локальный минимум является и глобальным. Действительно, пусть $\widehat{x}$ — локальный минимум, т. е. существует такая окрестность U точки $\widehat{x}$, что $f(\widehat{x}) \leqslant f(x)$ для всех $x \in U$. Пусть теперь x — произвольная точка из X. Для достаточно малых $0 < \alpha \leqslant 1$ точки $(1-\alpha)\widehat{x}+\alpha x = \widehat{x}+\alpha(x-\widehat{x})$, очевидно, принадлежат U и поэтому (по неравенству Иенссена) $f(\widehat{x}) \leqslant f((1-\alpha)\widehat{x}+\alpha x) \leqslant (1-\alpha)f(\widehat{x})+\alpha f(x)$, откуда следует, что $f(\widehat{x}) \leqslant f(x)$.

Теорема 5.1 (Ферма для выпуклых функций)**.** *Точка $\widehat{x}$ является минимумом в задаче* (5.1) *тогда и только тогда, когда $0 \in \partial f(\widehat{x})$.*

$\lhd$ Если $\widehat{x}$ — минимум в (5.1), то $f(x)-f(\widehat{x}) \geqslant 0 = \langle 0, x\rangle$ для любого $x \in X$, т. е. $0 \in \partial f(\widehat{x})$. Если $0 \in \partial f(\widehat{x})$, то $f(x) - f(\widehat{x}) \geqslant \langle 0, x\rangle = 0$, т. е. $f(x) \geqslant f(\widehat{x})$ для любого $x \in X$. $\rhd$

Рассмотрим теперь выпуклые задачи с ограничениями. Пусть $f_i\colon X \to \overline{\mathbb{R}}$, $i = 0, 1, \ldots, m$, — выпуклые функции и A — непустое выпуклое подмножество X. Задача

$$f_0(x) \to \min, \quad f_i(x) \leqslant 0, \quad 1 \leqslant i \leqslant m, \quad x \in A, \tag{5.2}$$

называется *выпуклой задачей* с ограничениями или *задачей выпуклого программирования.*

Теорема 5.2 (Каруша — Куна — Таккера в субдифференциальной форме)**.** *Пусть $\widehat{x}$ — решение задачи* (5.2)*. Если функции f_i, $i = 0, 1, \ldots, m$, непрерывны в точке $\widehat{x}$, то найдутся числа $\lambda_i \geqslant 0$, $i = 0, 1, \ldots, m$, такие, что $\sum_{i=0}^{m} \lambda_i = 1$, $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \ldots, m$, и функционал $x_0^* \in \partial\delta A(\widehat{x})$, для которых справедливо включение*

$$0 \in \sum_{i=0}^{m} \lambda_i \partial f_i(\widehat{x}) + x_0^*.$$

Пусть в задаче (5.2) функции принимают только конечные значения. Сопоставим этой задаче *функцию Лагранжа*, определенную формулой

$$\mathscr{L}(x, \lambda) = \sum_{i=0}^{m} \lambda_i f_i(x),$$

где $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ — набор множителей Лагранжа.

Необходимые условия минимума в следующей теореме Каруша — Куна — Таккера (в традиционной форме) являются непосредственным следствием только что сформулированной теоремы.

Теорема 5.3 (Каруша — Куна — Таккера)**.**

1) *Необходимость. Пусть $\widehat{x}$ — решение задачи* (5.2). *Если функции f_i, $i = 0, 1, \ldots, m$, непрерывны в точке $\widehat{x}$, то найдется ненулевой набор множители Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ такой, что*

(*a*) $\min\limits_{x \in A} \mathscr{L}(x, \lambda) = \mathscr{L}(\widehat{x}, \lambda)$;

(*b*) $\lambda_i \geqslant 0$, $i = 0, 1, \ldots, m$;

(*c*) $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \ldots, m$.

Если существует такая точка $\overline{x} \in A$, что $f_i(\overline{x}) < 0$, $1 \leqslant i \leqslant m$, то $\lambda_0 \neq 0$ (условие Слейтера).

2) *Достаточность. Если существует допустимая в* (5.2) *точка $\widehat{x}$ и набор множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ с $\lambda_0 > 0$, удовлетворяющие условиям* (*a*), (*b*) *и* (*c*), *то $\widehat{x}$ — решение задачи* (5.2).

Условие (*c*) называют *условием дополняющей нежесткости.*

Следует сказать, что эта теорема верна и тогда, когда X — вещественное векторное пространство, $f_i \colon X \to \mathbb{R}$, $i = 0, 1, \ldots, m$, — выпуклые функции и A — непустое выпуклое подмножество X.

Тех сведений из выпуклого анализа, которые были приведены выше, уже достаточно, чтобы исследовать различные конкретные ситуации. Мы начинаем с классических задач теории приближений.

Экстремальные задачи классической теории приближений и оптимальное восстановление. Сначала докажем одно общее утверждение.

Пусть X — нормированное пространство, L — подпространство в X, X^* — сопряженное пространство к X, $x^* \in X^*$ и $\delta > 0$. Рассмотрим следующую задачу:

$$\langle x^*, x \rangle \to \max, \quad \|x\|_X \leqslant \delta, \quad x \in L. \tag{5.3}$$

Эта задача имеет вид задачи (2.1), в которой $Z = \mathbb{R}$, $\Lambda = x^*$, $Y = X$, I — оператор вложения L в X и $W = L$. Таким образом, если поставить задачу оптимального восстановления значений линейного функционала x^* на подпространстве L по следующей информации: элементы $x \in L$ известны с точностью до δ в метрике X, т. е. о каждом $x \in L$ известен элемент $y \in X$ такой, что $\|x - y\|_X \leqslant \delta$, то мы получаем (согласно (2.2)), что величина $E(x^*, L, I, \delta)$ — погрешность оптимального восстановления — не меньше величины $S(x^*, L, I, \delta)$ — значения задачи (5.3).

Естественно считать, что L не принадлежит ядру функционала x^*, так как в противном случае задача не представляет интереса. Если это не так, то функционал принимает и ненулевые значения. Тогда, если у задачи (5.3) есть решение, то оно, очевидно, ненулевое.

Справедливо следующее

Предложение 5.1. *Если* $\widehat{x}$ *— решение задачи* (5.3), *то найдется линейный функционал* $x_1^* \in \partial\|\widehat{x}\|_X$ *такой, что линейный метод* $\widehat{\varphi}\colon X \to \mathbb{R}$, *действующий по правилу*

$$\widehat{\varphi}(y) = \frac{\langle x^*, \widehat{x}\rangle}{\|\widehat{x}\|_X}\,\langle x_1^*, y\rangle,$$

является оптимальным.

Метод $\widehat{\varphi}$ *точен на* L, *т. е.* $\widehat{\varphi}(x) = \langle x^*, x\rangle$ *для любого* $x \in L$.

$\lhd$ Сформулированные теоремы о необходимых условиях касались выпуклых задач на минимум. Чтобы формально этому соответствовать, сопоставим задаче (5.3) следующую задачу:

$$-\langle x^*, x\rangle \to \min, \quad \|x\|_X \leqslant \delta, \quad x \in L. \tag{5.4}$$

Ясно, что это выпуклая задача и что, если $\widehat{x}$ — решение задачи (5.3), то $\widehat{x}$ — решение и задачи (5.4).

Согласно теореме Каруша — Куна — Таккера в субдифференциальной форме, примененной к задаче (5.4), найдутся $\lambda_i \geqslant 0$, $i = 0, 1$, $\lambda_0 + \lambda_1 = 1$ и $x_0^* \in \partial\delta L(\widehat{x})$ такие, что (субдифференциал линейного линейной функции $x \mapsto \langle x^*, x\rangle$ есть x^*)

$$0 \in -\lambda_0 x^* + \lambda_1 \partial\|\widehat{x}\|_X + x_0^*, \tag{5.5}$$

т. е. найдется такой функционал $x_1^* \in \partial\|\widehat{x}\|_X$, что

$$0 = -\lambda_0\langle x^*, x\rangle + \lambda_1\langle x_1^*, x\rangle + \langle x_0^*, x\rangle, \quad \forall\, x \in X. \tag{5.6}$$

Покажем, что $\lambda_0 \neq 0$. Предположим, что $\lambda_0 = 0$. Подставляя $x = \widehat{x}$ в (5.5) и учитывая, что $\langle x_0^*, \widehat{x}\rangle = 0$, поскольку $\widehat{x} \in L$, а $\partial\delta L(\widehat{x}) = L^{\perp}$, и что $\langle x^*, \widehat{x}\rangle = \|\widehat{x}\|_X$ в силу вида субдифференциала нормы, приходим к противоречию: $0 = \|\widehat{x}\|_X \neq 0$.

Деля тождество (5.6) на λ_0 и обозначая $\lambda = \lambda_1/\lambda_0$ и $\widetilde{x}^* = \lambda_0^{-1}x_0^*$, получим, что

$$\langle x^*, x\rangle = \lambda\langle x_1^*, x\rangle + \langle \widetilde{x}^*, x\rangle, \quad \forall\, x \in X,$$

или

$$\langle x^*, x\rangle = \lambda\langle x_1^*, x\rangle, \quad \forall\, x \in L. \tag{5.7}$$

Подставляя сюда $x = \widehat{x}$, получаем, что $\lambda = \langle x^*, \widehat{x}\rangle/\|\widehat{x}\|_X$.

Покажем, что линейный функционал $\widehat{\varphi}$: $y \to \lambda\langle x_1^*, y\rangle$ является оптимальным методом. Оценим его погрешность. Для любых $x \in L$ и $y \in X$ таких, что $\|x-y\|_X \leqslant \delta$ имеем, используя соотношение (5.7) и то, что $\|x_1^*\|_{X^*} = 1$

$$\begin{gathered}\big|\langle x^*, x\rangle - \widehat{\varphi}(y)\big| = \big|\langle x^*, x\rangle - \lambda\langle x_1^*, y\rangle\big| = \\ = \big|\langle x^*, x\rangle - \lambda\langle x_1^*, x\rangle + \lambda\langle x_1^*, x-y\rangle\big| = \lambda\big|\langle x_1^*, x-y\rangle\big| \leqslant \lambda\big\|x-y\big\|_X \leqslant \lambda\delta.\end{gathered}$$

Переходя слева к верхней грани по всем указанным x и y, получим, что

$$e(x^*, L, I, \delta, \widehat{\varphi}) \leqslant \lambda\delta. \tag{5.8}$$

Далее, так как $\lambda_1(\|\widehat{x}\|_X - \delta) = 0$ (согласно теореме), то отсюда и (5.7) следует, что

$$\langle x^*, \widehat{x}\rangle = \lambda\langle x_1^*, \widehat{x}\rangle = \lambda\|\widehat{x}\|_X = \lambda_0^{-1}\big(\lambda_1(\|\widehat{x}\|_X - \delta) + \lambda_1\delta\big) = \lambda\delta,$$

т. е. $\lambda\delta$ — значение задачи (5.3). Тогда, в силу неравенства (2.2), справедлива оценка

$$E(x^*, L, I, \delta) \geqslant \lambda\delta.$$

Из этого неравенства и (5.8) получаем, что

$$\lambda\delta \leqslant E(x^*, L, I, \delta) \leqslant e(x^*, L, I, \delta, \widehat{\varphi}) \leqslant \lambda\delta,$$

и тем самым $\widehat{\varphi}$ — оптимальный метод.

Точность метода $\widehat{\varphi}$ сразу следует из соотношения (5.7). $\triangleright$

Лекция 6. Оптимальное восстановление и задачи классической теории приближений

После доказательства общего результата (предложение 5.1), перейдем непосредственно к классическим задачам теории приближений.

6.1. Задача П. Л. Чебышева об экстраполяции. В 1881 г. Чебышев поставил следующую задачу. Среди всех полиномов степени не выше n, которые на отрезке $[-1, 1]$ не превосходят по модулю δ, найти тот, чье значение в точке $\tau > 1$ максимально. Если обозначить через $\mathscr{P}_n$ множество полиномов степени не выше n, то формальная формулировка этой задачи, очевидно, такова

$$x(\tau) \to \max, \quad \|x(\cdot)\|_{C([-1,1]]} \leqslant \delta, \quad x(\cdot) \in \mathscr{P}_n. \tag{6.1}$$

Данная задача имеет вид задачи (5.3). Соответствующая задача оптимального восстановления состоит в том, что надо восстановить значения линейного функционала $x(\cdot) \mapsto x(\tau)$ на подпространстве $\mathscr{P}_n$ по следующей информации: полиномы $x(\cdot) \in \mathscr{P}_n$ известны с точностью до δ в метрике $C([-1, 1])$, т. е. о каждом $x(\cdot) \in \mathscr{P}_n$ известна функция $y(\cdot) \in C([-1, 1])$ такая, что $\|x(\cdot) - y(\cdot)\|_{C([-1,1])} \leqslant \delta$.

Решением этой задачи является полином $\delta T_n(\cdot)$, где $T_n(\cdot)$ — полином Чебышева степени n. Напомним, что этот полином для всех $t \in \mathbb{R}$ имеет вид $T_n(t) = \cos(n \arccos t)$, $T_1(t) = t$, $T_2(t) = 2t^2 - 1$, $T_3(t) = 4t^3 - 3t$ и т. д. На отрезке $[-1, 1]$ полином Чебышева принимает поочередно свои максимальные (равные 1) и минимальные (равные -1) значения в $(n + 1)$-й точке, включая точки -1 и 1. Эти точки называют точками *альтернанса* полинома Чебышева.

Согласно предложению 5.1 в задаче оптимального восстановления, соответствующей задаче (6.1), существует линейный оптимальный метод $\widehat{\varphi}$, который определяется функционалом из субдифференциала нормы в $C([-1, 1])$ в точке $\delta T_n(\cdot)$. Сопряженное к $C([-1, 1])$ можно отождествить с пространством регулярных борелевских мер на $[-1, 1]$. В соответствии с описанием субдифференциала нормы в $C([-1, 1])$ (см. [14]), мера должна быть сосредоточена в точках альтернанса $\delta T_n(\cdot)$. Тогда в силу точности метода $\widehat{\varphi}$ справедливо равенство

$$\widehat{\varphi}(x(\cdot)) = x(\tau) = \lambda \sum_{i=1}^{n+1} \mu_i x(\tau_i), \quad \forall\, x(\cdot) \in \mathscr{P}_n,$$

где $\lambda = \delta T_n(\tau)/\delta = T_n(\tau)$ и τ_i, $i = 1, \ldots, n+1$, — точки альтернанса $T_n(\cdot)$. Подставляя в это равенство полиномы n-го порядка $l_1(\cdot), \ldots, l_{n+1}$ такие, что $l_i(\tau_i) = 1$ и $l_i(\tau_j) = 0$, $j \neq i$, получаем равенства $\lambda\mu_i = l_i(\tau)$, $i = 1, \ldots, n+1$, и значит, оптимальный метод действует по правилу

$$\widehat{\varphi}(y(\cdot)) = \sum_{i=1}^{n+1} l_i(\tau) y(\tau_i).$$

Это означает, что надо взять значения функции $y(\cdot)$ (которую мы наблюдаем вместо полинома) в точках τ_i, $i = 1, \ldots, n+1$, провести через них интерполяционный полином Лагранжа и взять значение этого полинома в точке τ. Это и будет наилучший метод восстановления функционала $x(\cdot) \mapsto x(\tau)$ на пространстве $\mathscr{P}_n$ по указанной приближенной информации.

Отметим, что оптимальный метод использует не всю информацию, т. е. не всю функцию $y(\cdot)$, а только ее значения в $(n+1)$-й точке.

6.2. Неравенство С. Н. Бернштейна для тригонометрических полиномов. Обозначим через $\mathbb{T}$ отрезок $[-\pi, \pi]$ с индентифицированными концами, а через $\mathscr{T}_n$ — пространство тригонометрических полиномов степени не выше n. В 1912 году С. Н. Бернштейн доказал следующее точное неравенство:

$$\|\dot{x}(\cdot)\|_{C(\mathbb{T})} \leqslant n\|x(\cdot)\|_{C(\mathbb{T})},$$

справедливое для всех $x(\cdot) \in \mathscr{T}_n$. Равенство достигается на функции $\sin nt$.

В силу инвариантности нормы относительно сдвига, вопрос о нахождении точной константы в этом неравенстве равносилен нахождению значения следующей экстремальной задачи:

$$\dot{x}(0) \to \max, \quad \|x(\cdot)\|_{C(\mathbb{T})} \leqslant \delta, \quad x(\cdot) \in \mathscr{T}_n. \tag{6.2}$$

Вместо δ можно поставить любую положительную константу, но мы взяли δ для связи с задачей оптимального восстановления. Эта задача состоит в том, что требуется восстановить значения линейного функционала $x(\cdot) \mapsto \dot{x}(0)$ на подпространстве $\mathscr{T}_n$ по следующей информации: полиномы $x(\cdot) \in \mathscr{T}_n$ известны с точностью до δ в метрике $C(\mathbb{T})$, т. е. о каждом $x(\cdot) \in \mathscr{T}_n$ известна функция $y(\cdot) \in C(\mathbb{T})$ такая, что $\|x(\cdot) - y(\cdot)\|_{C(\mathbb{T})} \leqslant \delta$. Рассуждая как и в предыдущем примере, получаем, что оптимальный метод восстановления имеет вид

$$\widehat{\varphi}(y(\cdot)) = \frac{1}{2n}\sum_{i=1}^{2n}\frac{(-1)^{i+1}}{2\sin^2\frac{\tau_i}{2}}\,y(\tau_i), \quad \tau_i = \frac{(2i-1)\pi}{2n}, \quad i = 1,\ldots,2n.$$

И снова, используется не вся функция $y(\cdot)$, а лишь ее значения в $2n$ точках.

6.3. Неравенства Ландау — Колмогорова. Под такими неравенствами понимают неравенства вида

$$\big\|x^{(k)}(\cdot)\big\|_{L_q(T)} \leqslant K\big\|x(\cdot)\big\|^{\alpha}_{L_p(T)}\big\|x^{(n)}(\cdot)\big\|^{\beta}_{L_r(T)}, \tag{6.3}$$

где $0 \leqslant k < n$ — целые, $1 \leqslant p,q,r \leqslant \infty$, $\alpha,\beta \geqslant 0$ и $T = \mathbb{R}$ или $\mathbb{R}_+$, справедливые для всех функций $x(\cdot) \in L_p(T)$, у которых $(n-1)$-я производная локально абсолютно непрерывна на T и $x^{(n)}(\cdot) \in L_r(T)$. Это пространство функций обозначим через $\mathscr{W}^n_{p,r}(T)$.

При фиксированном T неравенство (6.3) зависит от пяти параметров k, n, p, q, r (числа α и β определяются ими однозначно). Первое точное (т. е. с наименьшей возможной константой K) неравенство подобного вида для случая, когда $T = \mathbb{R}_+$, $k = 1$, $n = 2$ и $p = q = r = \infty$ было доказано Эдмундом Ландау в 1913 г. В 1938 г. А. Н. Колмогоров доказал точное неравенство для любых $1 \leqslant k < n$ и $p = q = r = \infty$. До сих пор этот результат является одним из самых ярких в этой тематике.

Нахождение точной константы в неравенстве (6.3) равносильно нахождению значения следующей экстремальной задачи:

$$\big\|x^{(k)}(\cdot)\big\|_{L_q(T)} \to \max, \quad \big\|x(\cdot)\big\|_{L_p(T)} \leqslant \delta, \quad \big\|x^{(n)}(\cdot)\big\|_{L_r(T)} \leqslant 1.$$

Здесь также вместо δ и 1 можно поставить любые положительные числа.

Если $q = \infty$, то нахождение значения этой задачи равносильно нахождению значения такой задачи

$$x^{(k)}(0) \to \max, \quad \big\|x(\cdot)\big\|_{L_p(T)} \leqslant \delta, \quad \big\|x^{(n)}(\cdot)\big\|_{L_r(T)} \leqslant 1, \tag{6.4}$$

которая имеет вид задачи (2.1), где $X = \mathscr{W}^n_{p,r}(T)$, Λ — линейный функционал $x(\cdot) \mapsto x^{(k)}(0)$, I — оператор вложения $\mathscr{W}^n_{p,r}(T)$ в $L_p(T)$ и

$$W = W^n_{p,r}(T) = \big\{x \in \mathscr{W}^n_{p,r}(T) :\, \big\|x^{(n)}(\cdot)\big\|_{L_r(T)} \leqslant 1\big\}.$$

Таким образом, соответствующая задача восстановления состоит в том, чтобы восстановить значения функционала $x(\cdot) \mapsto x^{(k)}(0)$ на классе W по следующей информации: известны значения функций из W с точностью до δ в метрике $L_p(T)$, т. е. о каждой функции $x(\cdot) \in W$ известна функция $y(\cdot) \in L_p(T)$ такая, что $\|x(\cdot) - y(\cdot)\|_{L_p(T)} \leqslant \delta$.

Лекция 7. Оптимальное восстановление и неравенства для производных Ландау — Колмогорова

Если известно решение задачи (6.4) из предыдущей лекции, то в соответствующей задаче оптимального восстановления существует линейный оптимальный метод, и соображения здесь аналогичны тем, что использовались при доказательстве предложения 5.1 из лекции 5. Не будем здесь доказывать соответствующего общего результата, а проиллюстрируем это на примере, когда в задаче (6.4) $T = \mathbb{R}_+$, $n = 1$, $k = 0$ и $p = r = 2$, т. е. рассматриваем задачу

$$x(0) \to \max, \quad \|x(\cdot)\|_{L_2(\mathbb{R}_+)} \leqslant \delta, \quad \|\dot{x}(\cdot)\|_{L_2(\mathbb{R}_+)} \leqslant 1,$$

которую можно записать так:

$$x(0) \to \max, \quad \int\limits_{\mathbb{R}_+} x^2(t)\,dt \leqslant \delta^2, \quad \int\limits_{\mathbb{R}_+} \dot{x}^2(t)\,dt \leqslant 1, \tag{7.1}$$

где $x(\cdot) \in \mathscr{W}^1_{2,2}(\mathbb{R}_+)$.

Найдем решение этой задачи, используя достаточные условия минимума в стандартной теореме Каруша — Куна — Таккера (см. лекцию 5). Для этого сопоставим задаче (7.1) задачу

$$-x(0) \to \min, \quad \int\limits_{\mathbb{R}_+} x^2(t)\,dt \leqslant \delta^2, \quad \int\limits_{\mathbb{R}_+} \dot{x}^2(t)\,dt \leqslant 1. \tag{7.2}$$

Очевидно, что решение этой задачи является решением и задачи (7.1). Запишем функцию Лагранжа задачи (7.2), считая, что $\lambda_0 = 1$ (мы можем рассуждать эвристически, не обосновывая каждый шаг, поскольку цель — удовлетворить достаточным условиям, т. е. предъявить допустимую функции и нужные множители

Лагранжа):

$$\mathscr{L}(x(\cdot),\lambda_1,\lambda_2) = -x(0) + \lambda_1\Bigg(\int\limits_{\mathbb{R}_+} x^2(t)\,dt - \delta^2\Bigg) + \lambda_2\Bigg(\int\limits_{\mathbb{R}_+} \dot{x}^2(t)\,dt - 1\Bigg).$$

Это гладкая функция, и если в некоторой точке $\widehat{x}(\cdot)$ она достигает минимума, то в этой точке ее производная равна нулю, т. е. для всех $x(\cdot) \in \mathscr{W}^1_{2,2}(\mathbb{R}_+)$ справедливо равенство

$$-x(0) + 2\lambda_1 \int\limits_{\mathbb{R}_+} \widehat{x}(t)x(t)\,dt + 2\lambda_2 \int\limits_{\mathbb{R}_+} \dot{\widehat{x}}(t)\dot{x}(t)\,dt = 0. \tag{7.3}$$

Интегрируя по частям во втором интеграле (считая, что функции стремятся к нулю на бесконечности), получаем, что необходимо

$$\lambda_2\ddot{\widehat{x}}(t) - \lambda_1\widehat{x}(t) = 0, \quad 2\lambda_2\dot{\widehat{x}}(0) = -1. \tag{7.4}$$

Отсюда ясно, что $\lambda_i > 0$, $i = 1, 2$. Общее решение уравнения есть сумма двух экспонент. Коэффициент при экспоненте с положительной степенью должен быть нулевым, и тем самым

$$\widehat{x}(t) = ce^{-\sqrt{\lambda_1/\lambda_2}\,t}.$$

Используя второе соотношение в (7.4) и условия дополняющей нежесткости

$$\int\limits_{\mathbb{R}_+} x^2(t)\,dt = \delta^2, \quad \int\limits_{\mathbb{R}_+} \dot{x}^2(t)\,dt = 1, \tag{7.5}$$

находим, что $c = 1/(2\sqrt{\lambda_1\lambda_2})$, $\lambda_1 = 1/(2\delta\sqrt{2\delta})$ и $\lambda_2 = \sqrt{\delta}/(2\sqrt{2})$ и значит,

$$\widehat{x}(t) = \sqrt{2\delta}\,e^{-t/\delta}.$$

Итак, функция $\widehat{x}(\cdot)$ допустима в задаче (7.2), так как она удовлетворяет соотношениям (7.5). Эта функция с набором $\lambda = (\lambda_0, \lambda_1, \lambda_2)$, где $\lambda_0 = 1$, а λ_1 и λ_2 только что найдены, удовлетворяют соотношениям (a), (b) и (c) теоремы Каруша — Куна — Таккера. Действительно, соотношения (b) и (c) выполняются очевидны образом. Функция Лагранжа выпукла по $x(\cdot)$ и ее производная в точке $\widehat{x}(\cdot)$ обращается в ноль. Следовательно, эта точка является глобальным минимумом функции Лагранжа, т. е. выполнено условие (a).

Таким образом, в силу достаточных условий в теореме Каруша — Куна — Таккера, $\widehat{x}(\cdot)$ — решение задачи (7.2), а значит и задачи (7.1), и тем самым ее значение равно $\widehat{x}(0) = \sqrt{2\delta}$. Тогда согласно оценке (2.2)

$$E(x(0), W_2^1(\mathbb{T}), I, \delta) \geqslant \sqrt{2\delta}. \tag{7.6}$$

Подставляя в (7.3) найденные $\widehat{x}(\cdot)$, λ_1 и λ_2, получим, что для всех $x(\cdot) \in \mathscr{W}^1_{2,2}(\mathbb{R}_+)$ выполняется равенство

$$x(0) = \frac{1}{\delta}\int\limits_{\mathbb{R}_+} e^{-t/\delta}\, x(t)\, dt - \int\limits_{\mathbb{R}_+} e^{-t/\delta}\, \dot{x}(t)\, dt, \tag{7.7}$$

справедливость которого можно проверить и непосредственно.

Опираясь на это равенство и рассуждая аналогично тому как это было сделано при доказательстве предложения 5.1 (см. лекцию 5), нетрудно показать, что линейный метод $\widehat{\varphi}\colon L_2(\mathbb{R}_+) \to \mathbb{R}$, действующий по правилу

$$\widehat{\varphi}(y(\cdot)) = \frac{1}{\delta}\int\limits_{\mathbb{R}_+} e^{-t/\delta}\, y(t)\, dt,$$

является оптимальным. Тем не менее, проведем здесь эти рассуждения.

Оценим погрешность метода $\widehat{\varphi}$. Пусть $x(\cdot) \in W_2^1(\mathbb{T})$ и $y(\cdot) \in L_2(\mathbb{R}_+)$ такое, что $\|x(\cdot) - y(\cdot)\|_{L_2(\mathbb{R}_+)} \leqslant \delta$. Тогда, в силу (7.7), имеем

$$x(0) - \widehat{\varphi}(y) = x(0) - \frac{1}{\delta}\int\limits_{\mathbb{R}_+} e^{-t/\delta}\, y(t)\, dt =$$

$$= x(0) - \frac{1}{\delta}\int\limits_{\mathbb{R}_+} e^{-t/\delta}\, x(t)\, dt + \frac{1}{\delta}\int\limits_{\mathbb{R}_+} e^{-t/\delta}\, (x(t) - y(t))\, dt =$$

$$= -\int\limits_{\mathbb{R}_+} e^{-t/\delta}\, \dot{x}(t)\, dt + \frac{1}{\delta}\int\limits_{\mathbb{R}_+} e^{-t/\delta}\, (x(t) - y(t))\, dt.$$

Оценивая интегралы справа по неравенству Коши — Буняковского, получим, что

$$\bigl|x(0) - \widehat{\varphi}(y)\bigr| \leqslant \sqrt{\frac{\delta}{2}}\, \|\dot{x}(\cdot)\|_{L_2(\mathbb{R}_+)} + \frac{1}{\delta}\sqrt{\frac{\delta}{2}}\, \|x(\cdot) - y(\cdot)\|_{L_2(\mathbb{R}_+)} \leqslant \sqrt{2\delta}$$

и значит, $e(x(0), W_2^1(\mathbb{T}), I, \delta, \widehat{\varphi}) \leqslant \sqrt{2\delta}$. Отсюда и из оценки (7.6) следует, что

$$\sqrt{2\delta} \leqslant E(x(0), W_2^1(\mathbb{T}), I, \delta) \leqslant e(x(0), W_2^1(\mathbb{T}), I, \delta, \widehat{\varphi}) \leqslant \sqrt{2\delta},$$

т. е. $\widehat{\varphi}$ — оптимальный метод.

До сих пор рассматривались задачи восстановления линейных функционалов. Теперь мы займемся задачами восстановления линейных операторов. Начнем с задач восстановления периодических функций по конечному набору их коэффициентов Фурье.

Сначала сделаем несколько предварительных замечаний. Обозначим через $\mathscr{W}_2^1(\mathbb{T})$ (напомним, $\mathbb{T}$ обозначает отрезок $[-\pi, \pi]$ с идентифицированными концами) совокупность абсолютно непрерывных 2π-периодических функций, производные которых принадлежит $L_2(\mathbb{T})$. Пусть $x(\cdot) \in \mathscr{W}_2^1(\mathbb{T})$. Тогда, как хорошо известно, эта функция в каждой точке t единственным образом разлагается в ряд Фурье

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos kt + b_k \sin kt),$$

который сходится к ней равномерно. Коэффициенты Фурье a_k и b_k вычисляются по формулам

$$a_k = a_k(x(\cdot)) = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos kt\, dt, \quad k = 0, 1, 2, \ldots,$$

$$b_k = b_k(x(\cdot)) = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin kt\, dt, \quad k = 1, 2, \ldots$$

Восстановление в среднеквадратической метрике по точным значениям коэффициентов Фурье. Определим следующий класс функций:

$$W_2^1(\mathbb{T}) = \left\{x(\cdot) \in \mathscr{W}_2^1(\mathbb{T}) : \|\dot{x}(\cdot)\|_{L_2(\mathbb{T})} \leqslant 1\right\}.$$

Пусть n_1 и n_2 — натуральные числа. Допустим, что о каждой функции $x(\cdot) \in W_2^1(\mathbb{T})$ нам известны ее коэффициенты Фурье

$a_k = a_k(x(\cdot))$, $k = 0, 1, \ldots, n_1$, и $b_k = b_k(x(\cdot))$, $k = 1, \ldots, n_2$. Мы хотим по этой информации восстановить функции из класса $W_2^1(\mathbb{T})$.

В терминах общей постановки здесь $X = \mathscr{W}_2^1(\mathbb{T})$, $Z = L_2(\mathbb{T})$, Λ — оператор вложения $\mathscr{W}_2^1(\mathbb{T})$ в $L_2(\mathbb{T})$, информационный оператор $I = I_{n_1,n_2}\colon \mathscr{W}_2^1(\mathbb{T}) \to \mathbb{R}^{n_1+n_2+1}$, $I_{n_1,n_2}x(\cdot) = (a_0, a_1, \ldots, a_{n_1}, b_1, \ldots, b_{n_2})$ и $\delta = 0$.

В данной задаче погрешность оптимального восстановления обозначим так: $E(W_2^1(\mathbb{T}), I_{n_1,n_2})$.

Теорема 7.1. *Пусть $n_0 = \min(n_1, n_2)$. Тогда*

$$E\big(W_2^1(\mathbb{T}), I_{n_1,n_2}\big) = \frac{1}{n_0+1}.$$

Метод

$$\widehat{\varphi}(I_{n_1,n_2}x(\cdot))(t) = \frac{a_0}{2} + \sum_{k=1}^{n_0}(a_k \cos kt + b_k \sin kt), \quad t \in \mathbb{T},$$

где $a_k = a_k(x(\cdot))$, $k = 0, 1, \ldots, n_0$ и $b_k = b_k(x(\cdot))$, $k = 1, \ldots, n_0$, является оптимальным.

Отметим, что хотя коэффициенты Фурье функции известны точно, оптимальный метод использует не все коэффициенты Фурье, а только с номерами до n_0 (а их может быть очень мало).

Лекция 8. Оптимальное восстановление периодических функций и решений дифференциальных уравнений

Доказательство теоремы, сформулированной в конце предыдущей лекции, достаточно просто, и поэтому приведем его здесь.

$\lhd$ Согласно оценке (2.2) из первой лекции погрешность оптимального восстановления не меньше значения следующей задачи:

$$\|x(\cdot)\|_{L_2(\mathbb{T})} \to \max, \quad I_{n_1,n_2}x(\cdot) = 0, \quad \|\dot{x}(\cdot)\|_{L_2(\mathbb{T})} \leqslant 1 \tag{8.1}$$

на пространстве $\mathscr{W}_2^1(\mathbb{T})$. Покажем, что значение $S(W_2^1(\mathbb{T}), I_{n_1,n_2})$ этой задачи не меньше $1/(n_0+1)$.

Действительно, пусть $n_0 = n_1$. Рассмотрим функцию $x_0(t) = \cos(n_0+1)t/(n_0+1)$. Тогда ясно, что $I_{n_1,n_2}x_0(\cdot) = 0$. Элементарный подсчет показывает, что $\|\dot{x}_0(\cdot)\|_{L_2(\mathbb{T})} = 1$ и $\|x_0(\cdot)\|_{L_2(\mathbb{T})} = 1/(n_0+1)$.

Если $n_0 = n_1$, то беря функцию $x_0(t) = \sin(n_0+1)t/(n_0+1)$, приходим к тем же выводам. Таким образом, функция x_0 допустима в задаче (8.1) и, следовательно, $S(W_2^1(\mathbb{T}), I_{n_1,n_2}) \geqslant 1/(n_0+1)$. Отсюда вместе с оценкой (2.2) получаем, что

$$E\big(W_2^1(\mathbb{T}), I_{n_1,n_2}\big) \geqslant \frac{1}{n_0+1}. \tag{8.2}$$

Оценим погрешность метода $\widehat{\varphi}$ из формулировки теоремы. Предварительно отметим, что если $x(\cdot) \in \mathscr{W}_2^1(\mathbb{T})$, то для этой функции и ее производной справедливы соответственно равенства Парсеваля

$$\frac{1}{\pi}\int_{-\pi}^{\pi} x^2(t)\,dt = \frac{a_0^2}{2} + \sum_{k=1}^{\infty}\big(a_k^2 + b_k^2\big)$$

и

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \dot{x}^2(t)\,dt = \sum_{k=1}^{\infty} k^2\big(a_k^2 + b_k^2\big).$$

Пусть $x(\cdot) \in W_2^1(\mathbb{T})$. Используя равенство Парсеваля для $x(\cdot)$, элементарные оценки и равенство Парсеваля для $\dot{x}(\cdot)$, будем иметь

$$\big\|x(\cdot) - \varphi(I_{n_1,n_2}x(\cdot))(\cdot)\big\|_{L_2(\mathbb{T})}^2 = \sum_{k=n_0+1}^{\infty}\big(a_k^2 + b_k^2\big) \leqslant$$

$$\leqslant \frac{1}{(n_0+1)^2}\sum_{k=n_0+1}^{\infty} k^2\big(a_k^2 + b_k^2\big) \leqslant \frac{1}{(n_0+1)^2}\sum_{k=1}^{\infty} k^2\big(a_k^2 + b_k^2\big) =$$

$$= \frac{1}{(n_0+1)^2}\big\|\dot{x}(\cdot)\big\|_{L_2(\mathbb{T})}^2 \leqslant \frac{1}{(n_0+1)^2}.$$

Извлекая квадратный корень из обеих частей этого неравенства и переходя слева к верхней грани по всем $x(\cdot) \in W_2^1(\mathbb{T})$, получаем оценку

$$e\big(W_2^1(\mathbb{T}), I_{n_1,n_2}, \widehat{\varphi}\big) \leqslant \frac{1}{n_0+1}.$$

Отсюда и (8.2) следует, что

$$\frac{1}{n_0+1} \leqslant E\big(W_2^1(\mathbb{T}), I_{n_1,n_2}\big) \leqslant e\big(W_2^1(\mathbb{T}), I_{n_1,n_2}, \widehat{\varphi}\big) \leqslant \frac{1}{n_0+1},$$

т. е. $\widehat{\varphi}$ — оптимальный метод и $E(W_2^1(\mathbb{T}), I_{n_1,n_2}) = 1/(n_0+1)$. $\triangleright$

8.1. Восстановление в среднеквадратической метрике по неточным значениям коэффициентов Фурье. Рассмотрим теперь задачу восстановления функции $x(\cdot)$ на том же классе $W_2^1(\mathbb{T})$, но в ситуации, когда вместо точных значений ее коэффициентов Фурье $a_k = a_k(x(\cdot))$, $k = 0, 1, \ldots, n_1$ и $b_k = b_k(x(\cdot))$, $k = 1, \ldots, n_2$, известны лишь их приближенные значения, т. е. известны числа $\widetilde{a}_0, \widetilde{a}_1, \ldots, \widetilde{a}_{n_1}$ и $\widetilde{b}_1, \ldots, \widetilde{b}_{n_2}$ такие, что

$$|a_k - \widetilde{a}_k| \leqslant \delta,\ k = 0, 1, \ldots, n_1, \quad |b_k - \widetilde{b}_k| \leqslant \delta,\ k = 1, \ldots, n_2,$$

где $\delta > 0$.

Погрешность оптимального восстановления в этой задаче обозначаем так $E(W_2^1(\mathbb{T}), I_{n_1,n_2}, \delta)$.

Для каждого $\delta > 0$ положим

$$p(\delta) = \max\bigl\{p \in \mathbb{Z}_+ : \delta^2 p(p+1)(2p+1) < 3\bigr\}.$$

Теорема 8.1. *Пусть* $p_0 = p_0(\delta, n_1, n_2) = \min(p(\delta), n_1, n_2)$. *Тогда*

$$E\bigl(W_2^1(\mathbb{T}), I_{n_1,n_2}, \delta\bigr) = \frac{1}{p_0+1}\sqrt{1 + \frac{\delta^2}{6}\,(p_0+1)(8p_0^2 + 13p_0 + 3)}\,.$$

Метод

$$\widehat{\varphi}(\widetilde{a}_0, \ldots, \widetilde{a}_{n_1}, \widetilde{b}_1, \ldots, \widetilde{b}_{n_2})(t) = \\ = \frac{\widetilde{a}_0}{2} + \sum_{k=1}^{p_0}\left(1 - \left(\frac{k}{p_0+1}\right)^2\right)\bigl(\widetilde{a}_k \cos kt + \widetilde{b}_k \sin kt\bigr), \quad t \in \mathbb{T},$$

является оптимальным.

Заметим, что оптимальный метод использует далеко не всю имеющуюся информацию о коэффициентах Фурье, а ту, которую использует, фильтрует («сглаживает»). Это соответствует тому, что происходит на практике: высокие частоты отбрасывают, а низкие сглаживают.

Перейдем теперь к задаче оптимального восстановления решения дифференциального уравнения по неточным начальным данным.

8.2. Восстановление решения уравнения теплопроводности в данный момент времени по неточным его измерениям в другие моменты времени. Распространение тепла на прямой $\mathbb{R}$ описывается, как хорошо известно, уравнением теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial t^2},$$

где $u(\cdot,\cdot)$ — функция на $[0,\infty)\times\mathbb{R}$ с заданным начальным распределением температуры

$$u(0,\cdot) = u_0(\cdot).$$

Мы предполагаем, что $u_0(\cdot) \in L_2(\mathbb{R})$. Единственным решением выписанного уравнения является интеграл Пуассона

$$u(t,x) = u(t,x;u_0(\cdot)) = \frac{1}{2\sqrt{\pi t}} \int\limits_{\mathbb{R}} e^{-\frac{|x-\xi|^2}{4t}}\, u_0(\xi)\, d\xi, \tag{8.3}$$

и при этом $u(t,\cdot) \to u_0(\cdot)$ при $t \to 0$ в метрике $L_2(\mathbb{R})$.

Ставится следующая задача. Пусть в моменты времени $0 \leqslant t_1 < \ldots < t_n$ приближенно известны распределения температур $u(t_1,\cdot),\ldots,u(t_n,\cdot)$. Точнее говоря, известны функции $y_i(\cdot) \in L_2(\mathbb{R})$ такие, что

$$\big\|u(t_i,\cdot) - y_i(\cdot)\big\|_{L_2(\mathbb{R})} \leqslant \delta_i, \quad i = 1,\ldots,n,$$

где $\delta_i > 0$, $i = 1,\ldots,n$. Мы хотим каждому такому набору функций сопоставить функцию из $L_2(\mathbb{R})$, которая бы наилучшим образом аппроксимировала истинное распределение температуры на $\mathbb{R}$ в фиксированный момент времени t.

Формально эта задача не попадает под общую постановку задачи оптимального восстановления, сформулированную в первой лекции, но которую легко обобщить, чтобы покрывался и этот случай. Мы не стали этого делать, чтобы не загромождать изложение. Непосредственно здесь поставим соответствующую задачу оптимального восстановления.

Каждому методу φ из $(L_2(\mathbb{R}))^n = L_2(\mathbb{R}) \times \ldots \times L_2(\mathbb{R})$ в $L_2(\mathbb{R})$ сопоставляем его погрешность по формуле

$$e\big(t,\overline{\delta},\varphi\big) = \sup_{\substack{u_0(\cdot)\in L_2(\mathbb{R}),\, \overline{y}(\cdot)\in(L_2(\mathbb{R}))^n,\\ \|u(t_i,\cdot)-y_i(\cdot)\|_{L_2(\mathbb{R})}\leqslant\delta_i,\, i=1,\ldots,n}} \big\|u(t,\cdot) - \varphi(\overline{y}(\cdot))(\cdot)\big\|_{L_2(\mathbb{R})},$$

где $\overline{y}(\cdot) = (y_1(\cdot),\ldots,y_n(\cdot))$ и $\overline{\delta} = (\delta_1,\ldots,\delta_n)$.

Как и раньше, нас интересует погрешность оптимального восстановления, которая имеет вид

$$E(t,\overline{\delta}\,) = \inf_{\varphi\colon\, (L_2(\mathbb{R}))^n \to L_2(\mathbb{R})} e(t,\overline{\delta},\varphi),$$

и оптимальные методы, т. е. методы, на которых нижняя грань достигается.

Перед формулировкой теоремы сделаем некоторые построения. Напомним, что если A — некоторое подмножество линейного пространства X, то $\operatorname{co} A$ обозначает выпуклую оболочку A, т. е. наименьшее выпуклое множество, содержащее A. Преобразование Фурье в $L_2(\mathbb{R})$ обозначаем через F.

На двумерной плоскости (t,x) рассмотрим множество (см. рис. 8.1)

$$M = \operatorname{co}\big\{(t_j, \ln(1/\delta_j)),\ 1 \leqslant j \leqslant n\big\} + \big\{(t,0) \mid t \geqslant 0\big\}.$$

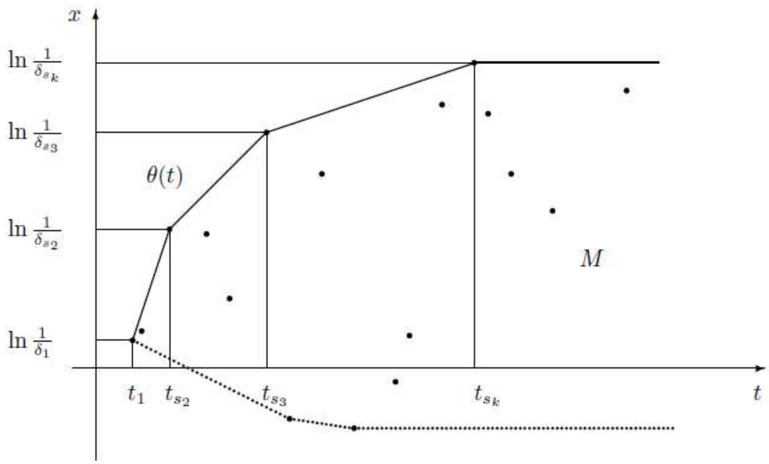

Рис. 8.1.

Определим функцию $\theta(\cdot)$ на $[0,\infty)$ равенством: $\theta(t) = \max\{x \,|\, (t,x) \in M\}$, причем $\theta(t) = -\infty$, если $(t,x) \notin M$ для всех x. Ясно, что на $[t_1,\infty)$ функция $\theta(\cdot)$ — вогнутая ломаная. Обозначим через $t_{s_1} < \ldots < t_{s_k}$ ее точки излома (считая t_1 также точкой излома, т. е. $t_{s_1} = t_1$), которые, очевидно, являются подмножеством точек $\{t_1,\ldots,t_n\}$.

Для $t \in (t_{s_j}, t_{s_{j+1}})$, $1 \leqslant j \leqslant k-1$, мы полагаем

$$\lambda_{s_j}(t) = \frac{t_{s_{j+1}} - t}{t_{s_{j+1}} - t_{s_j}} \left(\frac{\delta_{s_j}}{\delta_{s_{j+1}}}\right)^{\frac{-2(t-t_{s_j})}{t_{s_{j+1}}-t_{s_j}}},$$

$$\lambda_{s_{j+1}}(t) = \frac{t - t_{s_j}}{t_{s_{j+1}} - t_{s_j}} \left(\frac{\delta_{s_j}}{\delta_{s_{j+1}}}\right)^{\frac{2(t_{s_{j+1}}-t)}{t_{s_{j+1}}-t_{s_j}}}.$$

Легко видеть, что это положительные числа и $\lambda_{s_j}(t) < 1$.

Определим также следующее множество:

$$B_j(t) = \left\{\xi \in \mathbb{R} : |\xi| \leqslant \sqrt{-\frac{\ln \lambda_{s_j}(t)}{2(t - t_{s_j})}}\right\}.$$

Теорема 8.2. *Справедливы следующие утверждения:*

1) *Для любого* $t \geqslant 0$

$$E(t, \overline{\delta}\,) = e^{-\theta(t)}.$$

2) *Если* $t \in (t_{s_j}, t_{s_{j+1}})$, $1 \leqslant j \leqslant k-1$, *то множество измеримых функций* $\omega(\cdot)$ *на* $\mathbb{R}$ *таких, что*

$$\begin{cases} \dfrac{\left|e^{-(t-t_{s_j})a(\xi)} - \omega(\xi)e^{-(t_{s_{j+1}}-t_{s_j})a(\xi)}\right|^2}{\lambda_{s_j}(t)} + \dfrac{|\omega(\xi)|^2}{\lambda_{s_{j+1}}(t)} \leqslant 1, & \xi \in B_j(t), \\ 0, & \xi \neq B_j(t) \end{cases}$$

непусто и для каждой такой функции $\omega(\cdot)$ *метод* $\widehat{\varphi}_\omega$, *определенный по формуле*

$$\widehat{\varphi}_\omega(y_1(\cdot), \ldots, y_n(\cdot))(\cdot) = (K_j * y_{s_j})(\cdot) + (K_{j+1} * y_{s_{j+1}})(\cdot),$$

где $F[K_j](\xi) = e^{-(t-t_{s_j})a(\xi)} - \omega(\xi)e^{-(t_{s_{j+1}}-t_{s_j})a(\xi)}$ *и* $F[K_{j+1}](\xi) = \omega(\xi)$ *для п. в.* $\xi \in \mathbb{R}^d$, *является оптимальным.*

3) *Если* $t = t_{s_j}$, *тогда метод* $\widehat{\varphi}$, *определенный формулой* $\widehat{\varphi}(y_1(\cdot), \ldots, y_n(\cdot))(\cdot) = y_{s_j}(\cdot)$, *является оптимальным.*

4) *Если* $t > t_{s_k}$, *то метод* $\widehat{\varphi}$, *определенный формулой* $\widehat{\varphi}(y_1(\cdot), \ldots, y_n(\cdot))(\cdot) = (K * y_{s_k})(\cdot)$, *где* $F[K](\xi) = e^{-(t-t_{s_k})a(\xi)}$ *для п. в.* $\xi \in \mathbb{R}^d$, *является оптимальным.*

Сделаем несколько замечаний по поводу сформулированной теоремы.

1. Если $t_1 > 0$ и $0 \leqslant t_1 < t$, то $\theta(t) = -\infty$ и, следовательно, $E(t) = +\infty$; т. е. прошлое нельзя восстановить по неточному настоящему. В этом случае любой метод можно рассматривать как оптимальный.

2. Формулы для оптимальных методов определены корректно. Действительно, функции $\omega(\cdot)$, очевидно, ограничены, и так как они равны нулю вне некоторого шара, то они принадлежат $L_2(\mathbb{R})$.

3. Оптимальные методы линейны, сглаживают наблюдения и используют информацию о не более, чем двух измерениях.

4. Если $t = t_i$ и t_i не точка излома функции $\theta(\cdot)$, то оптимальный метод позволяет это измерение уточнить.

Литература

1. *Смоляк С. А.* Об оптимальном восстановлении функций и функционалов от них: дис. ... канд. физ.-мат. наук.—М.: МГУ, 1965.
2. *Магарил-Ильяев Г. Г., Осипенко К. Ю.* Об оптимальном восстановлении функционалов по неточным данным // Мат. заметки.—1991.—Т. 50, № 6.—С. 85–93.
3. *Micchelli C. A., Rivlin T. J.* A Survey of Optimal Recovery // Optimal Estimation in Approximation Theory / Eds. C. A. Micchelli and T. J. Rivlin.—P. 1–54.—N. Y.: Plenum Press, 1977.
4. *Melkman A. A., Micchelli C. A.* Optimal Estimation of Linear Operators in Hilbert Spaces from Inaccurate Data // SIAM J. Numer. Anal.—1979.—Vol. 16.—P. 87–105. DOI: 10.1137/0716007.
5. *Micchelli C. A., Rivlin T. J.* Lectures on Optimal Recovery.—Berlin: Springer–Verlag, 1985.—P. 21–93.—(Lect. Notes Math. Vol. 1129).
6. *Traub J. F., Woźniakowski H.* A General Theory of Optimal Algorithms.–N. Y.: Acad. Press, 1980.
7. *Магарил-Ильяев Г. Г., Осипенко К. Ю.* Оптимальное восстановление функций и их производных по приближенной информации о спектре и неравенства для производных // Функц. анализ и его прил.—2003.—Т. 37, № 3.—С. 51–64. DOI: 10.4213/faa157.
8. *Магарил-Ильяев Г. Г., Осипенко К. Ю.* Оптимальное восстановление решения уравнения теплопроводности по неточным измерениям // Мат. сб.—2009.—Vol. 200, № 5.—P. 37–54. DOI: 10.4213/sm7301.
9. *Магарил-Ильяев Г. Г., Осипенко К. Ю.* Об оптимальном гармоническом синтезе по неточно заданному спектру // Функц. анализ и его прил.—2010.—Т. 44, № 3.—С. 76–79. DOI: 10.4213/faa2999.
10. *Магарил-Ильяев Г. Г., Осипенко К. Ю.* Неравенство Харди — Литтлвуда — Полиа и восстановление производных по неточной информации // Докл. АН.—2011.—Т. 438, № 3.—С. 300–302.

11. *Магарил-Ильяев Г. Г., Сивкова Е. О.* Наилучшее восстановление оператора Лапласа функции по ее неточно заданному спектру // Мат. сб.—2012.—Т. 203, № 4.—С. 119–130. DOI: 10.4213/sm7903.
12. *Magaril-Il'yaev G. G., Sivkova E. O.* Optimal Recovery of Semi-Group Operators from Inaccurate Data // Eurasian Math. J.—2019.—T. 10, № 4.—P. 75–84. DOI: 10.32523/2077-9879-2019-10-4-75-84.
13. *Магарил-Ильяев Г. Г., Тихомиров В. М.* О неравенствах для производных колмогоровского типа // Мат. сб.—1997.—Т. 188, № 12.—С 73–106. DOI: 10.4213/sm274.
14. *Магарил-Ильяев Г. Г., Тихомиров В. М.* Выпуклый анализ и его приложения, изд. 5-е, доп.—М.: УРСС, 2020.

Магарил-Ильяев Георгий Георгиевич
Московский государственный университет имени М. В. Ломоносова
РОССИЯ, 119991, Москва, Ленинские горы, 1
E-mail: magaril@mech.math.msu.su, georgii.magaril@math.msu.ru

OPTIMAL RECOVERY OF LINEAR FUNCTIONALS AND OPERATORS

G. G. Magaril-Il'yaev

The lectures present the beginnings of the theory of optimal recovery of the values of linear functionals and operators on classes of sets whose elements are known approximately. Particular attention is paid to the recovery of linear functionals, where the theory relies heavily on the methods of convex analysis. Various examples are given, mainly related to problems of classical approximation theory. The lectures are intended for a wide range of readers.

Key words: optimal recovery, optimal method, extremal problem, convex duality.