Пять сюжетов о творчестве Владимира Михайловича Тихомирова

Г. Г. Магарил-Ильяев

Математическое творчество Владимира Михайловича Тихомирова очень разнообразно, но основные интересы его связаны с теорией приближений, теорией экстремальных задач и выпуклым анализом. В своей деятельности Владимир Михайлович всегда руководствуется идеей, что все значительное в математике основано на небольшом числе общих фундаментальных принципов, а конкретные результаты получаются на основе хорошо развитых исчислений. Здесь будет представлено несколько сюжетов, рассказывающих о проблемах, которыми занимался Владимир Михайлович в разные времена и где его влияние было существенным в становлении и/или развитии соответствующей тематики. Этих сюжетов пять, и это связано, в частности, с тем, что Владимир Михайлович любит, чтобы в любом списке (а он их составляет часто и по самым разным поводам) число пунктов было кратно пяти.

1. Поперечники множеств. В 1936 году А. Н. Колмогоров ввел понятие поперечника — величину, которая характеризует наилучшее приближение данного множества пространствами фиксированной размерности. Определение поперечника таково. Пусть X — нормированное пространство, C — подмножество X и n — неотрицательное целое число. Величина

$$d_n(C,X)=\inf_{L_n}\sup_{x\in C}\inf_{\xi\in L_n}\|x-\xi\|_X,$$

где первая нижняя грань берется по всем подпространствам L_n в X размерности не выше n, называется *n-поперечником по Колмогорову множества C в X*.

Для того, чтобы лучше почувствовать смысл этого понятия, рассмотрим следующий пример. Пусть X — плоскость с евклидовой нормой $|\cdot|$ и C — эллипс (см. рис. 1). Мы хотим наилучшим образом приблизить C одномерными подпространствами (т. е. прямыми L, проходящими через ноль). Это можно понимать так. Фиксируем какую-нибудь прямую L и пусть $x\in C$. Расстоянием от x до L в метрике X называется величина $d(x,C,X)=\inf_{\xi\in L}|x-\xi|$, а величина $d(C,L,X)=\sup_{x\in C}d(x,L,X)$

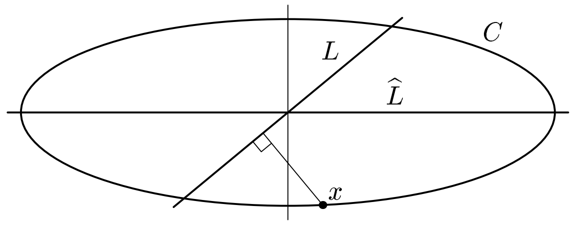

Рис. 1.

— уклонением C от L в метрике X. Нас интересует то подпространство $\widehat{L}$, на котором это уклонение минимально. Мы говорим тогда, что $\widehat{L}$ осуществляет наилучшее приближение эллипса C одномерными подпространствами. Очевидно, что $\widehat{L}$ совпадает с осью абсцисс и минимальное уклонение равно длине малой полуоси эллипса, что в соответствии с определением выше и есть 1-поперечник по Колмогорову эллипса C в евклидовой норме.

Появление работы Колмогорова ознаменовало начало нового этапа развития теории приближений, связанного с поисками наилучшего метода приближения данного функционального класса. С 1936 года и вплоть до шестидесятых годов практически не было публикаций на тему о поперечниках. Но затем, под влиянием исследований, связанных с проблемой Гильберта о суперпозициях функций и модной тогда теорией информации, эта тематика стала активно развиваться.

В этот период В. М. Тихомиров публикует ряд работ, которые стали определяющими для ее дальнейшего развития. Он ввел ряд величин, которые с разных точек зрения характеризуют аппроксимативные возможности данного множества (проекционный поперечник, линейный поперечник, поперечник по Гельфанду, по Бернштейну и др.) и получил ряд конкретных результатов для них, основываясь на разработанных им же методах, которые впоследствии многократно применялись и обобщались в десятках работ. Определенный итог этого этапа своей деятельности Владимир Михайлович подвел в монографии [1].

Здесь мы сформулируем один результат В. М. Тихомирова, который дает общий подход к оценкам поперечников снизу и покажем на примере, как он применяется.

Теорема (Тихомирова о поперечнике шара). *Пусть X — нормированное пространство, L_{n+1} — $(n+1)$-мерное подпространство X, $\gamma > 0$ и $\gamma BL_{n+1} = \{x \in L_{n+1} \mid \|x\|_X \leqslant \gamma\}$. Тогда*

$$d_n(\gamma BL_{n+1}, X) = \gamma.$$

Рассмотрим следующий пример. Пусть $X = C([0,1])$ — пространство непрерывных функций $x(\cdot)$ на отрезке $[0,1]$ с нормой $\|x(\cdot)\|_{C([0,1])} = \max_{t\in[0,1]}|x(t)|$, $C = W^1_\infty([0,1])$ — класс функций $x(\cdot)$ на $[0,1]$, удовлетворяющих условию Липшица (с константой единица), т.е. $|x(t) - x(t')| \leqslant |t - t'|$ для всех $t,t' \in [0,1]$. Покажем, что для любого $n \geqslant 1$

$$d_n(W^1_\infty([0,1]), C([0,1])) = \frac{1}{2n}\,.$$

Оценка снизу. Обозначим через L_{n+1} пространство ломаных на отрезке $[0,1]$ с изломами в точках $t_k = k/n$, $k = 1,\dots,n-1$. Легко проверить, что размерность L_{n+1} равна $n+1$ и что если $x(\cdot) \in L_{n+1}$ и максимум модулей угловых коэффициентов звеньев ломаной $x(\cdot)$ не превосходит единицы, то $x(\cdot) \in W^1_\infty([0,1])$. В частности, если $\|x(\cdot)\|_{C([0,1])} \leqslant 1/2n$, то нетрудно видеть, что этот максимум модулей не превосходит единицы и значит, $(1/2n)BL_{n+1} \subset W^1_\infty([0,1])$. Отсюда, учитывая очевидное свойство поперечника: если $C_1 \subset C_2$, то $d_n(C_1,X) \leqslant d_n(C_2,X)$, будем иметь по теореме Тихомирова о поперечнике шара:

$$d_n(W^1_\infty([0,1]), C([0,1])) \geqslant d_n\Big(\frac{1}{2n}BL_{n+1}, C([0,1])\Big) = \frac{1}{2n}\,.$$

Оценка сверху. Обозначим через L_n подпространство ломаных с изломами в точках $\tau_k = (2k-1)/2n$, $k = 1,\dots,n-1$, и постоянных на отрезках $[0,\tau_1]$ и $[\tau_{n-1},1]$. Легко видеть, что L_n — n-мерное пространство. Пусть $x(\cdot) \in W^1_\infty([0,1])$. Сопоставим этой функции ломаную $\xi(\cdot) \in L_n$ такую, что $\xi(\tau_k) = x(\tau_k)$, $k = 1,\dots,n-1$. Простой подсчет показывает, что $\|x(\cdot) - \xi(\cdot)\|_{C([0,1])} \leqslant 1/2n$. Тогда расстояние от $x(\cdot)$ до L_n в метрике $C([0,1])$ тем более не превосходит $1/2n$, т.е. $d(x(\cdot), L_n, C([0,1])) \leqslant 1/2n$. Переходя слева в этом неравенстве к верхней грани по всем $x(\cdot) \in W^1_\infty([0,1])$, а затем к нижней грани по всем подпространствам размерности не выше n, получаем, что

$$d_n(W^1_\infty([0,1]), C([0,1])) \leqslant \frac{1}{2n}\,.$$

Вместе с оценкой снизу это доказывает нужное утверждение.

2. Средняя размерность и наилучшие приближения на некомпактных многообразиях. В теории приближений, начиная с работ С. Н. Бернштейна, изучаются приближения отдельных функций и классов функций, заданных на всей прямой $\mathbb{R}$. Сам Бернштейн ввел для этого некий аналог пространства тригонометрических полиномов, а именно, пространство $B_\sigma(\mathbb{R})$, $\sigma > 0$, которое представляет собой сужения целых функций f на $\mathbb{R}$, удовлетворяющих условию: для любого $\varepsilon > 0$ найдется

такое число $C_\varepsilon = C_\varepsilon(f) \geqslant 0$, что для всех $z \in \mathbb{C}$ выполняется неравенство $|f(z)| \leqslant C_\varepsilon \exp(\sigma+\varepsilon)|z|$. Эквивалентное определение $B_\sigma(\mathbb{R})$ таково: это совокупность функций f на $\mathbb{R}$, у которых преобразование Фурье (как обобщенной функции) сосредоточено на отрезке $[-\sigma,\sigma]$. В последующие годы появились другие средства приближения классов функций на прямой, например, пространства сплайнов и вейвлетов. Кусочно-постоянные функции и ломаные — простейшие примеры пространств сплайнов. Эти пространства и пространства $B_\sigma(\mathbb{R})$ бесконечномерны, а стандартные классы функций (скажем, соболевские классы функций на прямой), для приближения которых они используются — некомпактны. Как в этой ситуации сравнивать различные средства приближения?

Один их возможных подходов основан на понятии средней размерности пространства, истоки которого восходят к работам К. Шеннона по теории информации, где он дал определение «энтропии на единицу времени» случайного сигнала на прямой с ограниченным спектром (т. е. реализации этого сигнала как раз являются элементами $B_\sigma(\mathbb{R})$). В 1956 году А. Н. Колмогоров модифицировал это определение для подпространств обычных (не случайных) функций.

Первый результат в этом направлении — энтропия на единицу времени ограниченных функций из $B_\sigma(\mathbb{R})$ — был получен В. М. Тихомировым [2]. Впоследствии Владимир Михайлович ввел характеристику подпространства, аналогичную колмогоровской, но отправляющуюся не от энтропии, а от поперечника по Колмогорову, которая и получили название средней размерности пространства. Суть дела такова. Начнем с простой ситуации. Пусть L^h — пространство ломаных на $\mathbb{R}$ с изломами в точках $\{kh\}$, $k \in \mathbb{Z}$, $h > 0$. Для любого $T > 0$ обозначим через L^h_T сужение пространства L^h на отрезок $[-T,T]$. Легко проверить, что размерность L^h_T равна $2T/h+3$, если T/h — целое, и $2[T/h]+3$ (где $[T/h]$ — целая часть T/h), если T/h — нецелое, и отсюда легко следует, что $\lim_{T\to\infty}(\dim L^h_T)/2T = h^{-1}$. Это число и называется средней размерностью (или размерностью на единицу времени) пространства L^h (заметим, что мы, очевидно, получим тот же результат, если будем рассматривать, например, только ограниченные ломаные). Эта процедура (с заменой в последней формуле предела на нижний предел) позволяет определить среднюю размерность любого пространства, сужение которого на $[-T,T]$ конечномерно для всех $T>0$.

Ситуация несколько сложнее, если сужение подпространства на $[-T,T]$ не конечномерно. В этом случае поступаем следующим образом. Пусть L — подпространство, скажем, пространства $C^b(\mathbb{R})$ ограниченных непрерывных функций $x(\cdot)$ на прямой с нормой $\|x(\cdot)\|_{C^b(\mathbb{R})} = \sup_{t\in\mathbb{R}}|x(t)|$ и $BC^b(\mathbb{R})$ — единичный шар в $C^b(\mathbb{R})$. Обозначим через $(L\cap BC^b(\mathbb{R}))_T$ —

сужение множества $L \cap BC^b(\mathbb{R})$ на отрезок $[-T,T]$ и предположим, что это сужение компактно в $C([-T,T])$ (пространство непрерывных функций на $[-T,T]$ с обычной нормой) для каждого $T > 0$. Тогда его можно с любой точностью приблизить конечномерными подпространствами и значит, для любого $\varepsilon > 0$ конечна величина

$$K_\varepsilon(T, L, C^b(\mathbb{R})) = \min\{n \in \mathbb{Z}_+ \mid d_n((L \cap BC^b(\mathbb{R}))_T, C([-T,T])) < \varepsilon\},$$

где $d_n((L \cap BC^b(\mathbb{R}))_T, C([-T,T]))$ — n-поперечник по Колмогорову множества $(L \cap BC^b(\mathbb{R}))_T$ в метрике $C([-T,T])$. Ясно, что $K_\varepsilon(T, L, C^b(\mathbb{R}))$ — минимальная размерность подпространства, аппроксимирующего множество $(L \cap BC^b(\mathbb{R}))_T$ с точностью до ε.

Легко проверить, что функция $\varepsilon \to K_\varepsilon(T, L, C^b(\mathbb{R}))$ не возрастает при каждом $T > 0$.

Средней размерностью пространства L в $C^b(\mathbb{R})$ назовем величину

$$\overline{\dim}(L, C^b(\mathbb{R})) = \lim_{\varepsilon \to 0} \liminf_{T \to \infty} \frac{K_\varepsilon(T, L, C^b(\mathbb{R}))}{2T}.$$

Легко проверить, что если сужение L на $[-T,T]$ конечномерно, то данное определение равносильно определению, данному выше.

В терминах средней размерности результат В. М. Тихомирова, о котором говорилось выше, звучит так

$$\overline{\dim}(B_\sigma(\mathbb{R}) \cap C^b(\mathbb{R}), C^b(\mathbb{R})) = \frac{\sigma}{\pi}.$$

Пусть $\nu > 0$. Тогда отсюда и из предыдущего следует, что если h и σ таковы, что $1/h = \sigma/\pi = \nu$, то средние размерности пространства ограниченных ломаных и пространства ограниченных функций из $B_\sigma(\mathbb{R})$ равны ν. Рассматривая все пространства функций на прямой, у которых средняя размерность не превосходит ν, можно говорить о выборе наилучшего среди них для приближении того или иного класса функций на прямой. Точнее говоря, можно определить средний ν-поперечник по Колмогорову и другие средние поперечники и поставить те же вопросы, что и в компактной ситуации: точные и асимптотически точные значения этих поперечников, экстремальные пространства и т. п. Это определило новое направление в теории приближений — наилучшие приближения некомпактных классов функций. Здесь получено много интересных результатов. В настоящее время наиболее активно эта тематика развивается в Китае.

3. Принцип Лагранжа. В 1897 году Лагранж высказал следующий принцип: «Если ищется максимум или минимум некоторой функции многих переменных при условии, что между этими переменными имеется связь, задаваемая одним или несколькими уравнениями, нужно прибавить

к функции, о которой говорилось, функции, задающие уравнения связи, умноженные на неопределенные множители, и искать затем максимум и минимум построенной суммы, как если бы переменные были независимы. Полученные уравнения, присоединенные к уравнениям связи, послужат для определения всех неизвестных».

Обдумывая необходимые условия экстремума в различных задачах, Владимир Михайлович пришел к выводу, что этот принцип (если придать ему чуть более расширенное толкование) имеет универсальный характер: практически все необходимые условия экстремума могут быть формально выведены, пользуясь рецептом Лагранжа.

Слова Лагранжа относились к задачам вида

$$f_0(x) \to \mathrm{extr}, \quad f_i(x) = 0, \quad i = 1, \ldots, m,$$

где extr означает либо максимум, либо минимум. Согласно рекомендациям надо составить функцию (функцию Лагранжа) данной задачи: $\mathcal{L}(x, \lambda) = \sum_{i=0}^m \lambda_i f_i(x)$ (разумно ставить множитель и у f_0; это некоторое продвижение со времен Лагранжа, но, правда, единственное), где $\lambda = (\lambda_0, \ldots, \lambda_m)$ — набор «неопределенных множителей» (множителей Лагранжа). Далее надо искать максимум и минимум функции $\mathcal{L}$ по x «как если бы переменные были независимы», для чего сначала надо выписать необходимые условия экстремума: равенство нулю частных производных. Если f_i — функции n переменных, то мы получим n уравнений. Относительно множителей λ_i эти уравнения однородны и поэтому один из них можно считать, скажем, равным единице. В результате имеем $n+m$ уравнений (m уравнений связи) для нахождения $n+m$ неизвестных. Точный результат (правило множителей Лагранжа) звучит так: *если функции f_i, $i = 0, 1, \ldots, m$, непрерывно дифференцируемы в некоторой окрестности точки $\widehat{x}$ и в этой точке достигается локальный экстремум, то найдется такой ненулевой набор множителей Лагранжа $\lambda = (\lambda_0, \ldots, \lambda_m)$, что $\widehat{x}$ удовлетворяет соотношению $\mathcal{L}_x(\widehat{x}, \lambda) = 0$, т. е. $\sum_{i=0}^m \lambda_i f_i'(\widehat{x}) = 0$, где $f_i'(\widehat{x})$ — производные (градиенты) функций f_i, $0 \leqslant i \leqslant m$, в точке $\widehat{x}$.*

Рассмотрим теперь так называемую задачу оптимального управления

$$\int_{t_0}^{t_1} f(t, x(t), u(t))\, dt \to \min, \quad \dot{x} = \varphi(t, x, u), \quad u(t) \in U,$$
$$x(t_0) = x_1, \quad x(t_1) = x_1, \quad (3.1)$$

и попробуем понять вид необходимых условий минимума в этой задаче, следуя формально рекомендациям Лагранжа.

В этой задаче f и φ — функции трех переменных, U — некоторое множество на прямой. Переменная u называются управлением, а x — фазовой переменной. Интерес к такого сорта задачам возник в пятидесятые годы прошлого века в ответ на запросы практики: требовалось оптимально (в том или ином смысле) управлять различными процессами, учитывая естественную ограниченность ресурсов (материальных, энергетических и т. п.). В нашем случае процесс описывается дифференциальным уравнением $\dot{x} = \varphi(t, x, u)$ (на «вход» подается управление $u(\cdot)$, на «выходе» получаем $x(\cdot)$). Мы хотим найти такое управление $\widehat{u}(\cdot)$, чтобы соответствующая фазовая переменная $\widehat{x}(\cdot)$ в начальный и конечный моменты времени t_0 и t_1 принимала значения x_0 и x_1, чтобы $\widehat{u}(\cdot)$ для каждого $t \in [t_0, t_1]$ не выходило за пределы множества U (отражающего ограниченность наших возможностей) и, наконец, чтобы это управление было оптимально в том смысле, что интеграл на паре $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ принимал минимальное значение.

Необходимые условия минимума в подобной задаче были получены в пятидесятые годы и получили название принципа максимума Понтрягина. Это одно из наиболее ярких достижений теории экстремальных задач.

Будем смотреть на задачу (3.1) как на задачу минимизации функции двух переменных $x(\cdot)$ и $u(\cdot)$, удовлетворяющих соответствующим ограничениям. Ограничение $\dot{x}(t) - \varphi(t, x(t), u(t)) = 0$, $t \in [t_0, t_1]$, можно воспринимать как континуум равенств, каждое из которых надо умножить на «неопределенный множитель», который обозначим (следуя традиции) $p(t)$ и «сложить», т. е. проинтегрировать. Таким образом, функция Лагранжа задачи (3.1) имеет вид

$$\mathcal{L}(x(\cdot), u(\cdot), \lambda) = \int_{t_0}^{t_1} (\lambda_0 f(t, x(t), u(t)) + p(t)(\dot{x}(t) - \varphi(t, x(t), u(t))))dt + \\ + \mu_0(x(t_0) - x_0) + \mu_1(x(t_1) - x_1),$$

где $\lambda = (\lambda_0, p(\cdot), \mu_0, \mu_1)$ — набор множителей Лагранжа. Пусть $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — решение задачи. Выпишем необходимые условия минимума функции Лагранжа в этой точке отдельно по $x(\cdot)$ и по $u(\cdot)$. По $x(\cdot)$ — это так называемая задача Больца — стандартная задача классического вариационного исчисления. Если обозначить через L подынтегральную функцию, то необходимые условия в этой задаче имеют вид

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \Leftrightarrow \quad \dot{p}(t) = -p(t)\widehat{\varphi}_x(t) - \lambda_0 \widehat{f}_x(t) \tag{3.2}$$

(где $\widehat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t), \widehat{u}(t), \lambda_0, p(t))$ и аналогично для остальных функций с крышкой) и

$$\widehat{L}_{\dot{x}}(t_0) = \mu_0, \ \widehat{L}_{\dot{x}}(t_1) = -\mu_1 \quad \Leftrightarrow \quad p(t_0) = \mu_0, \ p(t_1) = -\mu_1. \tag{3.3}$$

Нетрудно проверить, что если $\widehat{u}(\cdot)$ доставляет минимум по $u(\cdot)$ функции Лагранжа, то необходимо (и достаточно), чтобы в каждой точке t, где функция $\widehat{u}(\cdot)$ непрерывна, функция L достигала минимума по $u \in U$ в точке $\widehat{u}(t)$, т. е.

$$\min_{u\in U}(\lambda_0 f(t,\widehat{x}(t),u) - p(t)\varphi(t,\widehat{x}(t),u)) = \\ = \lambda_0 f(t,\widehat{x}(t),\widehat{u}(t)) - p(t)\varphi(t,\widehat{x}(t),\widehat{u}(t)). \quad (3.4)$$

Можно записать это выражение в форме максимума, поменяв знаки перед $\lambda_0 f$ и $p\varphi$ на противоположные.

Соотношения (3.2), (3.3) и (3.4) (в форме максимума) и есть принцип максимума Понтрягина для данной задачи. Точную формулировку не будем приводить, но суть ее (как и в правиле множителей Лагранжа) в том, что существует такой ненулевой набор множителей Лагранжа λ, что выполняются условия (3.2)–(3.4).

Решение конкретных задач на основе принципа Лагранжа происходит, как правило, по следующей схеме: выписываются необходимые условия экстремума, потом они анализируются и в результате находят «подозреваемый» на экстремум объект (скажем, функцию). После этого проверяют, что найденный объект действительно есть решение данной задачи (в следующем сюжете будет рассмотрен пример). Важно при этом отметить, что зачастую нам нужна лишь структура необходимых условий, для чего вполне достаточно владения принципом Лагранжа, а не знания соответствующего точного результата. Более того, иногда такого результата просто нет, а принцип Лагранжа дает правильный ориентир для решения задачи.

В течении многих лет Владимир Михайлович пропагандирует принцип Лагранжа как эффективный прием для решения самых различных экстремальных задач. К сожалению, пока его идея: «Почти все экстремальные задачи устроены одинаково и решаются стандартно» еще не «овладела массами». До сих пор появляются работы, посвященные решению той или иной экстремальной задачи, которые суть упражнения на применение принципа Лагранжа. Особенно важно понимание принципа Лагранжа для «потребителей» теории экстремума (инженеров, экономистов, управленцев), которым эта теория экстремума преподносится как набор отдельных рецептов для решения задач того или иного типа.

Различные вопросы теории экстремума, в частности, принцип Лагранжа Владимир Михайлович продумывал со многими своими учениками и коллегами, и в первую очередь здесь следует назвать Александра Давидовича Иоффе и Владимира Михайловича Алексеева.

Систематическое изложение необходимых условий экстремума в различных экстремальных задачах с точки зрения принципа Лагранжа можно найти в монографии [3].

4. Неравенства для производных. Неравенства вида

$$\|x^{(k)}(\cdot)\|_{L_q(T)} \leqslant K\|x(\cdot)\|^{\alpha}_{L_p(T)}\|x^{(n)}(\cdot)\|^{\beta}_{L_r(T)}\,, \tag{4.1}$$

где $0 \leqslant k < n$, $1 \leqslant p,q,r \leqslant \infty$, $T = \mathbb{R}_+$ или R и $\alpha \geqslant 0$, $\beta \geqslant 0$ (при этом α и β однозначно определяются параметрами k,n,p,q,r), называют *неравенствами колмогоровского типа* или *неравенствами Ландау – Колмогорова*. Такое название связано с тем, что в 1913 году Э. Ландау доказал неравенство (4.1) для случая $k = 1$, $n = 2$, $q = p = r = \infty$ и $T = \mathbb{R}_+$ и нашел наилучшую (т.е. наименьшую из возможных) констант K в этом неравенстве, которая оказалась равной 2, а в 1938 году А. Н. Колмогоров нашел точную константу в неравенстве (4.1) для любых целых k и n, $1 \leqslant k < n$, $q = p = r = \infty$ и $T = \mathbb{R}$.

Неравенства вида (4.1) привлекали внимание многих известных математиков (Адамар, Надь, Харди, Литтльвуд, Полиа и др.). Всплеск интереса к этим неравенствам произошел в шестидесятые годы прошлого века в связи с задачей о наилучшем приближении неограниченного оператора (обычно оператора дифференцирования) ограниченными операторами, поставленной С. Б. Стечкиным. Был получен ряд новых точных результатов. Неравенства (4.1) играют важную роль в различных вопросах анализа и теории приближений.

Нахождение точной константы в таком неравенстве равносильно решению следующей экстремальной задачи

$$\|x^{(k)}(\cdot)\|_{L_q(T)} \to \max, \quad \|x(\cdot)\|_{L_p(T)} \leqslant \gamma_1,\ \|x^{(n)}(\cdot)\|_{L_r(T)} \leqslant \gamma_2 \tag{4.2}$$

для любых $\gamma_1 > 0$ и $\gamma_2 > 0$.

В начале семидесятых годов В. М. Тихомиров поставил общую задачу

$$\|D^{\alpha_0}x(\cdot)\|_{L_{p_0}(T)} \to \max, \quad \|D^{\alpha_i}x(\cdot)\|_{L_{p_i}(T)} \leqslant \gamma_i,\ i = 1,\dots,m$$

(где D^{α} — оператор (вообще говоря) дробного дифференцирования, T — не только прямая или полупрямая, но, скажем, отрезок или все пространство $\mathbb{R}^n$) и предложил исследовать ее, опираясь на общие принципы теории экстремума. При этом основная цель, которую преследовал Владимир Михайлович, заключалась в том, чтобы испытать возможности этой теории, так как при различных p_i моделируется практически весь спектр задач, которые в ней изучаются (задачи вариационного исчисления, оптимального управления, выпуклые задачи, задачи с фазовыми ограничениями и т. д.). Не последнюю роль играла, конечно, и эстетическая сторона дела: точное решение красивой экстремальной задачи — всегда немалое удовольствие.

Через некоторое время выяснилось, что фактически все решенные к тому времени задачи вида (4.2) представляют собой, по сути дела, упражнения на применение принципа Лагранжа. Более того, такой естественный подход (решение экстремальной задачи методами теории экстремума) позволил сразу увидеть возможные обобщения полученных результатов, продвинуться в решении новых задач и увидеть связи их с другими задачами теории приближений.

В качестве примера найдем точную константу в одной достаточно простой ситуации, но при этом важно заметить, что схема рассуждений будет такой же и при доказательстве любого другого неравенства, где можно получить точный ответ аналитически.

Докажем точное неравенство

$$\|x(\cdot)\|_{C^b(\mathbb{R}_+)} \leqslant \sqrt{2}\,\|x(\cdot)\|^{1/2}_{L_2(\mathbb{R}_+)}\|\dot{x}(\cdot)\|^{1/2}_{L_2(\mathbb{R}_+)}\,, \tag{4.3}$$

справедливое для всех локально абсолютно непрерывных (т.е. абсолютно непрерывных на каждом отрезке) функций $x(\cdot) \in L_2(\mathbb{R}_+)$, у которых первая производная $\dot{x}(\cdot)$ также принадлежит $L_2(\mathbb{R}_+)$. Обозначим это пространство через $\mathcal{W}^1_2(\mathbb{R}_+)$.

Рассмотрим на $\mathcal{W}^1_2(\mathbb{R}_+)$ экстремальную задачу

$$x(0) \to \max, \quad \int_{\mathbb{R}_+} x^2(t)\,dt \leqslant 1, \ \int_{\mathbb{R}_+} \dot{x}^2(t)\,dt \leqslant 1. \tag{4.4}$$

Если мы найдем решение $\widehat{x}(\cdot)$ этой задачи, на котором ограничения обращаются в равенства, то докажем соответствующее точное неравенство. Действительно, пусть $x(\cdot) \in \mathcal{W}^1_2(\mathbb{R}_+)$ и $x(\cdot) \neq 0$. Тогда легко проверить, что функция $y(t) = ax(bt)$, где $a = (\|x(\cdot)\|_{L_2(\mathbb{R}_+)}\|\dot{x}(\cdot)\|_{L_2(\mathbb{R}_+)})^{-1/2}$ и $b = \|x(\cdot)\|_{L_2(\mathbb{R}_+)}/\|\dot{x}(\cdot)\|_{L_2(\mathbb{R}_+)}$ является допустимой в (4.4) и поэтому $y(0) = ax(0) \leqslant \widehat{x}(0)$, или

$$x(0) \leqslant \widehat{x}(0)\|x(\cdot)\|^{1/2}_{L_2(\mathbb{R}_+)}\|\dot{x}(\cdot)\|^{1/2}_{L_2(\mathbb{R}_+)}\,. \tag{4.5}$$

Отсюда следует, что справедливо такое неравенство

$$\|x(\cdot)\|_{C^b(\mathbb{R}_+)} \leqslant \widehat{x}(0)\|x(\cdot)\|^{1/2}_{L_2(\mathbb{R}_+)}\|\dot{x}(\cdot)\|^{1/2}_{L_2(\mathbb{R}_+)}\,. \tag{4.6}$$

В самом деле, если для некоторого $\overline{x}(\cdot) \in \mathcal{W}^1_2(\mathbb{R}_+)$ это неравенство не выполняется, то найдется такое $\tau \in \mathbb{R}_+$, что $\overline{x}(\tau)$ будет больше правой части в (4.6) при $x(\cdot) = \overline{x}(\cdot)$. Но на функции $x(t) = \overline{x}(t-\tau)$ значение правой части в (4.6), очевидно, не больше, чем на $\overline{x}(\cdot)$, а так как $x(0) = \overline{x}(\tau)$, то приходим к противоречию с неравенством (4.5).

Итак, для доказательства (4.3) осталось найти решение задачи (4.4) и убедиться, что его значение в нуле равно $\sqrt{2}$. Функция Лагранжа задачи (4.4) имеет вид

$$\mathcal{L}(x(\cdot),\lambda) = \lambda_0 x(0) + \lambda_1\biggl(\int_{\mathbb{R}_+} x^2(t)\,dt - 1\biggr) + \lambda_2\biggl(\int_{\mathbb{R}_+} \dot{x}^2(t)\,dt - 1\biggr),$$

где $\lambda = (\lambda_0, \lambda_1, \lambda_2)$. Согласно принципу Лагранжа, если $\widehat{x}(\cdot)$ — решение (4.4), то найдется такой ненулевой вектор λ, что функция Лагранжа удовлетворяет необходимым условиям минимума в точке $\widehat{x}(\cdot)$. Задача минимизации функции Лагранжа — это задача Больца (о которой говорилось уже в предыдущем сюжете), правда, на бесконечном промежутке. Не обращая на это внимания, выпишем формально необходимые условия минимума. Они заключаются в том, что должно выполняться уравнение Эйлера:

$$-2\lambda_2\ddot{\widehat{x}}(t) + 2\lambda_1\widehat{x}(t) = 0 \tag{4.7}$$

и условие трансверсальности:

$$2\lambda_2\dot{\widehat{x}}(0) = \lambda_0. \tag{4.8}$$

Проанализируем полученные соотношения. Во-первых, λ_1 и λ_2 не равны нулю, ибо, если $\lambda_1 = 0$, то из (4.7) следует, что либо $\lambda_2 = 0$, либо $\ddot{\widehat{x}}(\cdot) = 0$. В первом случае из (4.8) вытекает, что $\lambda_0 = 0$, т. е. все множители Лагранжа равны нулю, что невозможно. Во втором случае получаем, что $\widehat{x}(t) = at + b$. Но эта функция не принадлежит $L_2(\mathbb{R}_+)$, если $ab \neq 0$, а если $a = b = 0$, то $\widehat{x}(\cdot) = 0$, что, очевидно, тоже не так по смыслу задачи. Аналогичные рассуждения показывают, что и $\lambda_2 \neq 0$.

Далее, λ_1 и λ_2 должны быть одного знака, так как в противном случае никакое нетривиальное решение (4.7) не принадлежит $L_2(\mathbb{R}_+)$. Поскольку в соотношениях (4.7) и (4.8) множители Лагранжа определены с точностью до ненулевого сомножителя, то можно считать, что $\lambda_1 > 0$ и $\lambda_2 > 0$. Тогда из (4.7) следует, что

$$\widehat{x}(t) = C_1 e^{-\sqrt{\lambda_1/\lambda_2}\,t} + C_2 e^{\sqrt{\lambda_1/\lambda_2}\,t}.$$

Но $\widehat{x}(\cdot) \in L_2(\mathbb{R}_+)$, поэтому $C_2 = 0$ и тем самым $\widehat{x}(t) = C_1 e^{-\sqrt{\lambda_1/\lambda_2}\,t}$. Так как $\widehat{x}(\cdot) \neq 0$, то отсюда вытекает, что $\dot{\widehat{x}}(0) \neq 0$ и по смыслу задачи ясно, что $\dot{\widehat{x}}(0) < 0$. Тогда из (4.8) получаем, что $\lambda_0 < 0$ и можем считать, что $\lambda_0 = -1$.

Константы C_1, λ_1 и λ_2 найдем из (4.8) и соотношений

$$\int_{\mathbb{R}_+} \widehat{x}^2(t)\,dt = 1, \qquad \int_{\mathbb{R}_+} \dot{\widehat{x}}^2(t)\,dt = 1.$$

Элементарный подсчет показывает, что

$$\widehat{x}(t) = \sqrt{2}e^{-t}, \quad \lambda_1 = \lambda_2 = \frac{1}{2\sqrt{2}}\,.$$

Ясно, что $\widehat{x}(\cdot) \in \mathcal{W}_2^1(\mathbb{R}_+)$.

Докажем теперь, что функция $\widehat{x}(\cdot)$ действительно является решением задачи (4.4). Она удовлетворяет уравнению (4.7). Умножим обе его части на $x(\cdot) \in \mathcal{W}_2^1(\mathbb{R}_+)$, проинтегрируем по $\mathbb{R}_+$, затем проинтегрируем в первом слагаемом по частям[1)] и подставим вместо $\widehat{x}(\cdot)$, λ_1, λ_2 их выражения. Тогда получим, что для всех $x(\cdot) \in \mathcal{W}_2^1(\mathbb{R}_+)$ имеет место тождество

$$x(0) = \int_{\mathbb{R}_+} e^{-t} x(t)\, dt - \int_{\mathbb{R}_+} e^{-t} \dot{x}(t)\, dt.$$

(справедливость которого легко проверить и непосредственно).

Пусть $x(\cdot)$ — допустимая функция в задаче (4.4). По неравенству Коши – Буняковского

$$x(0) \leqslant \left(\int_{\mathbb{R}_+} e^{-2t}\, dt \right)^{1/2} \|x(\cdot)\|_{L_2(\mathbb{R}_+)} + \\ + \left(\int_{\mathbb{R}_+} e^{-2t}\, dt \right)^{1/2} \|\dot{x}(\cdot)\|_{L_2(\mathbb{R}_+)} \leqslant \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \sqrt{2}.$$

Но $\widehat{x}(0) = \sqrt{2}$ и следовательно, $\widehat{x}(\cdot)$ — решение задачи (4.4). Неравенство (4.3) доказано.

Относительно подхода, основанного на принципе Лагранжа, к задачам о неравенствах для производных и их взаимосвязям с другими задачами теории приближений см. в [4].

5. Выпуклый анализ. Выпуклый анализ — раздел математики, где изучают выпуклые множества, выпуклые функции и выпуклые экстремальные задачи.

Работы Владимира Михайловича по выпуклому анализу очень разнообразны и связаны как собственно с развитием этой дисциплины, так и с приложениями ее к задачам теории приближений и теории экстремума. Но важно и то, что под влиянием этих работ сформировался определенный взгляд на предмет выпуклого анализа, который оказался весьма плодотворным, особенно для приложений. Суть этого взгляда состоит в том, что основное содержание выпуклого анализа — это соотношения двойственности для некоторого набора операторов и порожденное ими выпуклое исчисление. Поясним сказанное на примере оператора сопряжения для конусов и затем применим это к вопросу о критериях существования решений систем линейных неравенств.

[1)] Учитывая, что функции из $\mathcal{W}_2^1(\mathbb{R}_+)$ ограничены. Действительно, $\int_0^t x(\tau)\,\dot{x}(\tau)\,d\tau = x^2(\tau)|_0^t - \int_0^t x(\tau)\,\dot{x}(\tau)\,d\tau$. Следовательно, $2\int_0^t x(\tau)\,\dot{x}(\tau)\,d\tau = x^2(t) - x^2(0)$. Отсюда, применяя неравенство Коши – Буняковского, получим неравенство $x^2(t) \leqslant x^2(0) + 2\left(\int_{\mathbb{R}_+} x^2(\tau)\,d\tau\right)^{1/2}\left(\int_{\mathbb{R}_+} \dot{x}^2(\tau)\,d\tau\right)^{1/2}$ для всех $t \in \mathbb{R}_+$.

Обозначим через $\mathbb{R}^n$ обычное евклидово пространство всех вектор-столбцов $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, а через $(\mathbb{R}^n)'$ — евклидово пространство вектор-строк $y = (y_1, \dots, y_n)$. Если $y \in (\mathbb{R}^n)'$, то отображение $x \mapsto y \cdot x = \sum_{i=1}^n y_i x_i$ есть, очевидно, линейный функционал и любой линейный функционал на $\mathbb{R}^n$ может быть представлен в таком виде. Таким образом, мы отождествляем $(\mathbb{R}^n)'$ с пространством всех линейных функционалов на $\mathbb{R}^n$ и говорим, что $(\mathbb{R}^n)'$ — двойственное пространство к $\mathbb{R}^n$.

Сопоставим каждому выпуклому конусу $C \in \mathbb{R}^n$ конус $C^* \in (\mathbb{R}^n)'$ (который называется сопряженным конусом к C) по формуле: $C^* = \{\, y \in (\mathbb{R}^n)' \mid y \cdot x \geqslant 0,\ \forall x \in C \,\}$. Легко видеть, что это выпуклый конус (и даже замкнутый). Соответствующий оператор обозначим «$*$». Повторное применение этого оператора приводит к конусу $C^{**} = (C^*)^* = \{\, x \in \mathbb{R}^n \mid y \cdot x \geqslant 0,\ \forall y \in C^* \,\}$. Элементарно проверяется, что $C \subset C^{**}$.

Соотношение двойственности для данного оператора состоит в том, что *если C — выпуклый замкнутый конус, то*

$$C^{**} = C. \tag{5.1}$$

Доказательство есть простое следствие теоремы отделимости точки от замкнутого выпуклого множества.

Если C_1 и C_2 — выпуклые конусы, то, очевидно, $C_1 + C_2$ и $C_1 \cap C_2$ — также выпуклые конусы. Далее, если C — выпуклый конус в $\mathbb{R}^n$ и $A\colon \mathbb{R}^n \to \mathbb{R}^m$ — линейный оператор (который будем отождествлять с его матрицей размера $m \times n$ в стандартных базисах $\mathbb{R}^n$ и $\mathbb{R}^m$ и обозначать то же буквой), то образ (прообраз) C при действии A, который обозначим AC (CA), есть снова выпуклый конус в $\mathbb{R}^m$ $(\mathbb{R}^n)$.

Выпуклое исчисление для данного случая — это формулы для сопряженных конусов только что определенных операций. Для того, чтобы их выписать определим еще оператор $A^*\colon (\mathbb{R}^m)' \to (\mathbb{R}^n)'$ — сопряженный к A, действующий по правилу $A^* y = yA$. Итак, формулы выглядят так:

$$\begin{array}{llll} (a) & (C_1 \cap C_2)^* = C_1^* + C_2^* & (b) & (C_1 + C_2)^* = C_1^* \cap C_2^* \\ (c) & (AC)^* = C^* A^* & (d) & (CA)^* = A^* C^*. \end{array}$$

Формулы (b) и (c) справедливы без каких-либо предположений (и проверяются без труда). Для справедливости формул (a) и (d) требуются дополнительные предположения, но они заведомо выполняются, если считать, что конусы полиэдральны, т. е. являются пересечениями конечного числа замкнутых полупространств. Можно проверить, что определенные выше операции переводят полиэдральные конусы в полиэдральные.

Перейдем к приложениям. Если A — матрица размера $m \times n$ и $b \in \mathbb{R}^m$, то системы m линейных уравнений и неравенств с n неизвестными могут быть записаны так: $Ax = b$ и $Ax \leqslant b$, где неравенство понимается покоординатно.

Первое, что мы выясним — это условия существования неотрицательного решения у системы $Ax = b$ или, другими словами, когда $b \in A\mathbb{R}^n_+$, где $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x \geqslant 0\}$. Ясно, что $\mathbb{R}^n_+$ — полиэдральный конус. В силу соотношения (5.1), формулы (c) и очевидного равенства $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$, будем иметь

$$A\mathbb{R}^n_+ = ((A\mathbb{R}^n_+)^*)^* = (\mathbb{R}^n_+ A^*)^*,$$

т. е. $b \in A\mathbb{R}^n_+$ в том и только в том случае, когда $y \cdot b \geqslant 0$ для тех $y \in (\mathbb{R}^m)'$, для которых $A^*y \geqslant 0$.

Это известная теорема Минковского – Фаркаша.

Следующий вопрос: какие условия существования решения у системы неравенств $Ax \leqslant b$? Другими словами, когда $b \in A\mathbb{R}^n + \mathbb{R}^m_+$? В силу (5.1), (b), (c) и того, что $(\mathbb{R}^n)^* = \{0\}$, будем иметь

$$A\mathbb{R}^n + \mathbb{R}^m_+ = ((A\mathbb{R}^n + \mathbb{R}^m_+)^*)^* = ((\mathbb{R}^n)^* A^* \cap (\mathbb{R}^m_+)^*)^* = (0A^* \cap \mathbb{R}^m_+)^*.$$

Но $0A^*$ — это ядро A^* и поэтому $b \in A\mathbb{R}^n + \mathbb{R}^m_+$ в том и только в том случае, когда $y \cdot b \geqslant 0$ для тех $y \in (\mathbb{R}^m)'$, для которых $y \geqslant 0$ и $A^*y = 0$.

Это также известная теорема Ки Фаня (Фань Цзи).

Наконец естественно поставить вопрос о существовании неотрицательного решения у системы неравенств $Ax \leqslant b$, т. е. когда $b \in A\mathbb{R}^n_+ + \mathbb{R}^m_+$? В силу (5.1), (b) и (c) получим

$$A\mathbb{R}^n_+ + \mathbb{R}^m_+ = ((A\mathbb{R}^n_+ + \mathbb{R}^m_+)^*)^* = ((\mathbb{R}^n_+)^* A^* \cap (\mathbb{R}^m_+)^*)^* = (\mathbb{R}^n_+ A^* \cap \mathbb{R}^m_+)^*$$

и таким образом, $b \in A\mathbb{R}^n_+ + \mathbb{R}^m_+$ в том и только в том случае, когда $y \cdot b \geqslant 0$ для тех $y \in (\mathbb{R}^m)'$, для которых $y \geqslant 0$ и $A^*y \geqslant 0$.

Это утверждение известно как теорема Гейла.

Если выпуклый конус — подпространство, то сопряженный конус — это аннулятор данного подпространства и соответствующее выпуклое исчисление позволяет получить критерии существования решений систем линейных уравнений.

Многие факты теории приближений и теории экстремума могут быть получены как следствия соотношений двойственности и формул выпуклого исчисления для тех или иных операторов, переводящих выпуклые объекты (выпуклые множества и выпуклые функции) в себя.

В монографиях [5]–[8] отражены основные воззрения Владимира Михайловича на предмет выпуклого анализа и его взаимосвязи с анализом, геометрией, теорией экстремума и теорией приближений.

Список литературы

[1] Тихомиров В. М. *Некоторые вопросы теории приближений*. Изд-во МГУ, 1976. 304 с.

[2] Тихомиров В. М. *Об ε-энтропии некоторых классов аналитических функций* // ДАН СССР, т. 117, №2, 1957, с. 191–194.

[3] Алексеев В. М., Тихомиров В. М., Фомин С. В. *Оптимальное управление*. Наука, Москва, 1979. 429 с.

[4] Магарил-Ильяев Г. Г., Тихомиров В. М. *О неравенствах для производных колмогоровского типа* // Мат. сборник, т. 188, №12, 1997, с. 73–106.

[5] Иоффе А. Д., Тихомиров В. М. *Теория экстремальных задач*. Наука, Москва, 1974. 479 с.

[6] Тихомиров В. М. *Выпуклый анализ*. Итоги науки и техники. Серия Современные проблемы математики. Фундаментальнык направления. ВИНИТИ, Москва, 1987, с. 5–101.

[7] Магарил-Ильяев Г. Г., Тихомиров В. М. *Выпуклый анализ и его приложения*. УРСС, Москва, 2003 (2-ое изд.), 176 с.

[8] Magaril-Il'yaev G. G., Tikhomirov V. M. *Convex Analysis: Theory and Applications*. AMS, Translations of Mathematical Monographs, vol. 222, 2003, 183 p.

Г. Г. Магарил-Ильяев, Московский государственный институт радиотехники, электроники и автоматики (технический университет)