Параметры в математических задачах и моделях

Для любой математической задачи желание упростить ее постановку вполне естественно. Однако в ряде случаев для получения решения задачи ее постановку приходится обобщать или даже усложнять.

Обобщение постановки обычно бывает необходимо для задач, содержащих некоторую неоднозначность в условии. Усложнение может потребоваться, если для решения задачи необходима дополнительная степень свободы, отсутствующая в условии.

Достаточно часто, как для описания неоднозначности, так и создания дополнительной степени свободы в постановке задачи используются *параметры*.

Целью данного пособия является исследование специальных классов параметрических задач и методов их решение. Поэтому предварительно уточним понятие параметра и поясним его примерами.

Определение 0.1.1	*Параметром* в условии задачи будем называть математический объект, не определенный *a priori* однозначно, но являющийся допустимым элементом из некоторого заданного множества.

Краткий обзор разных типов параметрических задач и методов начнем с очевидного примера. Допустим, что требуется

найти вещественные решения уравнения $\quad x^2 - 6x - 5 = 0\,.$

Эта задача может быть решена разложением левой части уравнения на множители или, просто, подбором корней по формуле Виета.

Однако, если вид уравнения обобщить до $x^2 + px + q = 0$, где $p, q \in \mathbb{R}$, то можно воспользоваться формулой для его корней при произвольных значениях параметров p и q.

Понятно, что, если нас интересуют корни только исходного уравнения, то такое усложнение условия вряд ли будет всегда полезным.

Рассмотрим другой пример: *из чисел $x_1 = 4$, $x_2 = -5$, $x_3 = 1$ требуется выбрать максимальное.*

Полный перебор очевидно даст ее решение $x_{\max} = 4$. Но, если мы почему-то не хотим применять этот метод, то решение может быть найдено,[1] например, так:

$$x_{\max} = \lim_{\tau\to+0} \tau \ln\left(e^{\frac{x_1}{\tau}} + e^{\frac{x_2}{\tau}} + e^{\frac{x_3}{\tau}}\right) = \lim_{\tau\to+0} \tau \ln\left(e^{\frac{4}{\tau}} + e^{\frac{-5}{\tau}} + e^{\frac{1}{\tau}}\right) = 4. \tag{0.1.1}$$

В формуле (0.1.1) используется вспомогательный положительный параметр τ. По этому параметру выполняется предельный переход к нулю, и, следовательно, значения τ никак не влияют на ответ задачи.

Формула (0.1.1), конечно, сложнее, чем перебор вариантов, но зато она не требует выполнения логической операции «*если…, то…, иначе…*»

Эти два примера позволяют дать уточняющее

Определение 0.1.2	Существуют параметры: — *экзогенные*, являющиеся для решаемой задачи носителем «внешней» информации, и — *инструментальные*, не влияющие на результат, но необходимые для реализации метода решения.

Понятно, что оба вида параметров могут одновременно использоваться в одной задаче. Так, во втором примере экзогенными параметрами являются $x_k \quad k = \overline{1,3}$, а параметр τ здесь очевидно инструментальный.

Рассмотрим задачу посложнее:

найти значение интеграла Дирихле

$$I(\alpha) = \int\limits_0^{+\infty} \frac{\sin \alpha x}{x}\, dx, \tag{0.1.2}$$

где α — произвольное вещественное число.

[1] Оцените на калькуляторе погрешность величины $x_{\max}$, возникающую, если в формуле (0.1.1) предел заменить значением функции $\tau \ln\left(e^{\frac{4}{\tau}} + e^{\frac{-5}{\tau}} + e^{\frac{1}{\tau}}\right)$, скажем, при $\tau = 0.1$.

Вычислить этот интеграл с помощью формулы Ньютона-Лейбница невозможно, поскольку неопределенный интеграл $\int \frac{\sin \alpha x}{x}\,dx$ не выражается через элементарные функции. Однако, по признаку Дирихле несобственный интеграл (0.1.2) сходится, то есть функция $I(\alpha)$ имеет конечные значения $\forall \alpha \in R$.

Найти эти значения можно, использовав вспомогательный интеграл

$$\Phi(\alpha,\beta) = \int\limits_0^{+\infty} e^{-\beta x}\,\frac{\sin \alpha x}{x}\,dx$$

с вещественным инструментальным параметром $\beta \in [0,1]$.

Построенный интеграл тождественно равен нулю при $\alpha = 0$. Для $\alpha \neq 0$ он также сходится по признаку Дирихле при любом фиксированном $\beta \in [0,1]$. Заметим также, что интеграл от *производной подынтегральной функции по* α, имеющий вид

$$\int\limits_0^{+\infty} e^{-\beta x}\cos \alpha x\,dx\,, \tag{0.1.3}$$

будет сходиться по признаку Вейерштрасса *равномерно* на множестве $\beta \in (0,1]$ к $\Phi'_\alpha(\alpha,\beta)$.

Неопределенный интеграл $\int e^{-\beta x}\cos \alpha x\,dx$ находится двукратным интегрированием по частям. Следовательно, несобственный интеграл (0.1.3) можно вычислить по формуле Ньютона-Лейбница. Он будет равен (проверьте это самостоятельно) $\dfrac{\beta}{\alpha^2+\beta^2}$.

Теперь найдем $\Phi(\alpha,\beta)$, интегрируя уравнение $\dfrac{\partial \Phi}{\partial \alpha} = \dfrac{\beta}{\alpha^2+\beta^2}$ по переменной α при постоянном значении β. В итоге мы получим $\Phi(\alpha,\beta) = \operatorname{arctg}\dfrac{\alpha}{\beta} + C(\beta)$, где $C(\beta) \equiv 0$ в силу $\Phi(0,\beta) = 0\ \ \forall \beta \in [0,1]$.

Перейдя в последней формуле к пределу $\beta \to +0$, для фиксированного $\alpha > 0$ мы найдем $I(\alpha) = \lim\limits_{\beta\to+0}\Phi(\alpha,\beta) = \dfrac{\pi}{2}$. Наконец, в силу нечетности синуса, для интеграла Дирихле при любом α имеем $I(\alpha) = \dfrac{\pi}{2}\operatorname{sgn}\alpha$.

Отметим, что в рассмотренной задаче параметр α является экзогенным, в то время как параметр β, очевидно, инструментальный.

Параметрический подход может также оказаться полезным при решении геометрических задач. Примером этого является параметризация описания семейств (пучков) линий второго порядка на плоскости.

Действительно, известно, что любая линейная комбинация уравнений линий второго порядка есть уравнение линии порядка *не выше, чем* 2.

Тогда можно ожидать, что при некоторых значениях коэффициентов (параметров) эта комбинация окажется линейной, то есть более простым уравнением, чем исходное. Соответственно, упростится поиск линии второго порядка, удовлетворяющей заданным условиям. Скажем, поиск линии, проходящей через фиксированный набор точек (см., например, [2x]).

К параметрическим также можно отнести алгоритмы, в которых некоторые переменные исходной задачи превращаются в параметры. Есть шанс, что в результате такого превращения задача упростится, например, станет линейной параметрической.

Примером такого метода служит классический алгоритм поиска собственных значений и собственных векторов линейного преобразования.

Согласно определению собственного значения λ для линейного преобразования $\widehat{A}$ в конечномерном линейном пространстве Λ^n должно выполняться равенство

$$\|\widehat{A} - \lambda\widehat{E}\|\|f\| = \|o\| \,. \qquad (0.1.4)$$

Здесь преобразование $\widehat{E}$ единичное, а $\|f\| = \|\,\varphi_1\,\varphi_2\,\ldots\,\varphi_n\,\|^{\mathrm{T}} \in \Lambda^n$ есть столбец собственного вектора преобразования $\widehat{A}$.

Равенство (0.1.4) представляет собой *нелинейную* систему n уравнений с неизвестными $\{\,\varphi_1,\,\varphi_2,\,\ldots,\,\varphi_n,\,\lambda\,\}$. Однако, если неизвестное собственное значение λ превратить в параметр, то равенство (0.1.4) окажется *линейной однородной* системой n уравнений с n неизвестными $\{\,\varphi_1,\ \varphi_2,\,\ldots,\,\varphi_n\,\}$.

Условие существования *ненулевого* решения для такой системы есть характеристическое уравнение вида $\det\|\widehat{A} - \lambda\widehat{E}\| = 0$. Решив его, мы получим искомые собственные значения λ. Затем из (0.1.4) найдем для каждого λ и собственные векторы $\|f\|$.

По мнению авторов, в обзоре параметрических алгоритмов нельзя не упомянуть о применяемом для решения задач оптимизации *методе штрафных функций.*

В этом методе, идея которого принадлежит Куранту [3], *инструментальный* параметр используется для преобразования исходной задачи поиска условного экстремума в параметрическую задачу без ограничений.

Поясним суть метода штрафных функций на примере задачи:

найти максимум функции $F(x) = 2x$
при условии $x \le 3$.

Построим вспомогательную функцию $A(\tau, x)$, значения которой мало отличаются от $F(x)$ при допустимых x и существенно меньше $F(x)$ для x недопустимых.

Например, такую функцию можно получить, прибавив к целевой функции $2x$ слагаемое вида $-\frac{1}{2\tau}\Big(\,|x-3|+x-3\,\Big)^2$, где τ — малый положительный параметр. Тогда функция $A(\tau, x)$ приобретает вид

$$A(\tau, x) = 2x - \frac{1}{2\tau}\Big(\,|x-3|+x-3\,\Big)^2 .$$

Добавленное слагаемое можно рассматривать как «штраф» за нарушение условия $x \le 3$.

Функция $A(\tau, x)$ определена и непрерывно дифференцируема на всей вещественной оси. Ее графики для различных значений τ показаны на рис. 0.0.1.

Нетрудно убедиться, что $A(\tau, x)$ при любом фиксированном положительном значении инструментально параметра τ имеет максимум в точке $\tilde{x}(\tau) = 3 + 2\tau$, равный $A(\tau, \tilde{x}(\tau)) = 6 + 2\tau$.

Этот экстремум является приближенным решением исходной задачи. Можно показать, что ее точное решение задачи определяется равенствами

$$x^* = \lim_{\tau \to +0} \tilde{x}(\tau) = 3 \qquad \text{и} \qquad F^* = \lim_{\tau \to +0} A(\tau, \tilde{x}(\tau)) = 6 .$$

Метод штрафных функций может быть использован для решения экстремальных задач разного вида. Например, обратимся еще раз к задаче выбора максимального числа в наборе $x_1 = 4$, $x_2 = -5$, $x_3 = 1$.

Эта задача равносильна задаче линейного программирования:

максимизировать по $v \in E^1$ *функцию* $F(v) = -v$
при условиях $4 - v \le 0, \quad -5 - v \le 0, \quad 1 - v \le 0.$

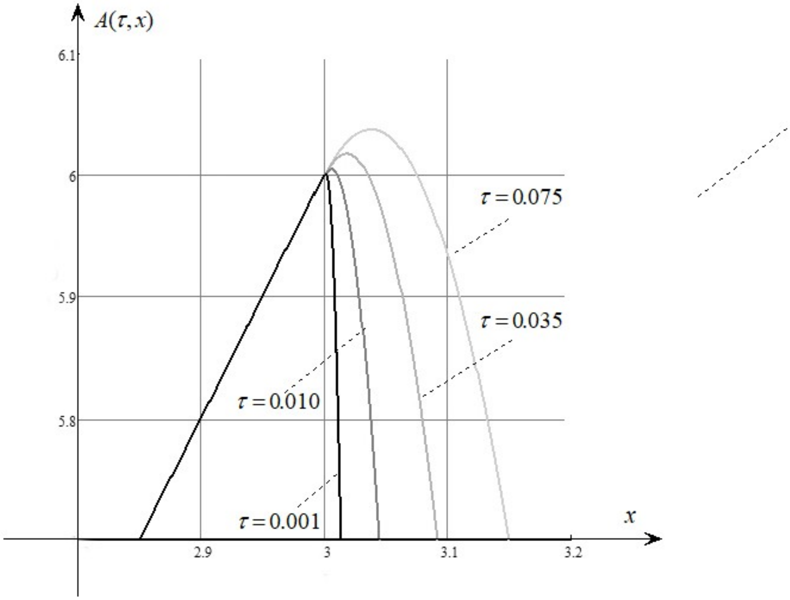

Рис. 0.0.1. Графики функции $A(\tau, x)$

Решим эту задачу методом штрафных функций, использовав $-\tau e^{\frac{s}{\tau}}$ в качестве штрафа за нарушение условия $s \leq 0$. Вспомогательная функция $A(\tau, v)$ в этом случае будет

$$A(\tau, v) = -v - \tau e^{\frac{4-v}{\tau}} - \tau e^{\frac{-5-v}{\tau}} - \tau e^{\frac{1-v}{\tau}} .$$

Нетрудно убедиться, что из равенства нулю производной $A(\tau, v)$ по переменной v

$$-1 + e^{\frac{4-\overline{v}}{\tau}} + e^{\frac{-5-\overline{v}}{\tau}} + e^{\frac{1-\overline{v}}{\tau}} = 0$$

следует

$$\overline{v}(\tau) \ = \tau \ln \left(e^{\frac{4}{\tau}} + e^{\frac{-5}{\tau}} + e^{\frac{1}{\tau}} \right) ,$$

откуда следует формула (0.1.1).

Для полноты картины отметим, что существуют также и более сложные схемы инструментальной параметризации. Скажем, алгоритмы, использующие несколько параметров одновременно. К ним, например, относится, рассматриваемая далее в § 2.4 пособия, процедура линейной экстраполяции с *векторными* инструментальными параметрами.

Рассмотренные примеры, позволяют констатировать, что, хотя использование параметров приводит к усложнению постановки задачи, возникающую при этом дополнительную степень свободы можно в ряде случаев использовать:

— для анализа свойств решений и поиска решений со специальными свойствами,

— для построения алгоритмов поиска самих решений.

В этом пособии мы рассмотрим альтернативный инструментально-параметрический алгоритм, называемый далее для краткости *методом функций обратных связей*. Этот подход позволяет находить и анализировать решения различных оптимизационных задач, содержащих экзогенные параметры.

Основной из них является параметрическая задача[2] поиска экстремума функции многих переменных с ограничениями-неравенствами

$$\begin{array}{ll} \textit{максимизировать функцию} \qquad F(X,V) \quad \textit{по } X \in E^n \\ \textit{при условиях:} \quad x \in \Omega \equiv \left\{ \; x \;\middle|\; f_i(X,V) \le 0 \quad i=\overline{1,m} \right\}, \end{array} \tag{0.1.5}$$

где вектор *экзогенных* параметров V фиксирован и принадлежит некоторой компактной области $\Upsilon \subseteq E^K$.

Пусть функции $F(X,V)$ и $f_i(X,V)$ $i=\overline{1,m}$ дважды непрерывно дифференцируемы по всем аргументам в их области определения $\Theta \subseteq E^n \otimes E^K$. Решением задачи (0.1.5) мы будем считать зависимость $\operatorname*{argmax}_{x\in\Omega} F(X,V)$ от V, обозначаемую далее как X_v^*.

Отметим, что ограничения типа *неравенство* в формулировке задачи (0.1.5) есть причина следующих свойств зависимости X_v^*:

1°. Область определения этой зависимости может оказаться *ýже*, чем Θ.

2°. Зависимость X_v^* может быть *неоднозначной* и, следовательно, не являться функций.

3°. Если зависимость X_v^* все-таки является функцией, она может оказаться *недифференцируемой* по компонентам V.

[2]Эту задачу иногда называют задачей *параметрического программирования*.

Понятно, что отмеченные особенности ограничивают применимость классических методов дифференциального исчисления для поиска и исследования свойств X_v^*.

Обычно для этих целей рекомендуются алгоритмы *недифференцируемой оптимизации* [2, 6, 13]. В данном пособии будет рассмотрен альтернативный подход, позволяющий решать и исследовать, как задачи (0.1.5), так и сводящиеся к ним, используя лишь стандартные формулы Тейлора.

Идея предлагаемого подхода состоит в построении специальной *гладкой аппроксимации* зависимости X_v^*, которая не только сохраняет основные свойства зависимости X_v^*, но и является в E^K *функцией* с областью определения *не ýже*, чем Θ.

Инструментальной основой предлагаемого метода служит совместное использование вспомогательного параметра и функций реализующих в процессе вычислений *обратную связь* между *переменными* задачи (0.1.5) и ее *множителями Лагранжа*.