

## Лекция. Общие сведения об оптимизации решений

**Цель:** Раскрыть сущность задач оптимизации и охарактеризовать их основные классы.

Время - 2 часа

**Учебные вопросы:**

1. Классическая задача оптимизации
2. Скалярная оптимизация
3. Векторная оптимизация

**Введение.**

### 1. Классическая задача оптимизации

Обычная постановка задачи оптимизации состоит в следующем. И некотором пространстве  $S$  тем или иным способом выделяется некоторое непустое множество  $M$  точек этого пространства, называемое *допустимым множеством*. Далее фиксируется некоторая вещественная функция  $f(x)$ , заданная во всех точках  $x$  допустимого множества  $M$ . Она называется *целевой функцией*. Задача оптимизации состоит в том, чтобы найти точку  $x_0$  в множестве  $M$ , для которой функция  $f(x)$  принимает экстремальное (Максимальное или минимальное) значение. В первом случае для всех точек  $x$  множества  $M$  удовлетворяется неравенство  $f(x_0) \geq f(x)$ , во втором случае - неравенство  $f(x_0) \leq f(x)$ .

Заметим, что как при *максимизации*, так и при *минимизации* задача **оптимизации** может не иметь решения. Например, если допустимое множество  $M$  состоит из всех положительных вещественных чисел ( $x > 0$ ), то для целевой функции  $f(x) = x^2$  в множестве  $M$  не существует ни точки максимума, ни точки минимума. Действительно, для любого  $x_0 > 0$  всегда найдутся положительные числа  $x_1$  и  $x_2$ , удовлетворяющие неравенству  $x_1 < x_0$  и  $x_2 > x_0$ . Тогда, очевидно,  $f(x_1) = x_1^2 < x_0^2 = f(x_0)$  и  $f(x_2) = x_2^2 > x_0^2 = f(x_0)$ .

Наоборот, может оказаться, что экстремум достигается не в одной, а в целом множестве точек. Например, в области, заданной неравенствами  $-1 \leq x \leq 1$  и  $-1 \leq y \leq 1$ , функция  $f(x, y) = x^2 + y^2$  достигает максимума (равного двум) в четырех точках с координатами  $(-1, -1)$ ,  $(-1, 1)$ ,  $(1, -1)$ ,  $(1, 1)$ . Функция  $f(x)$ , заданная соотношениями  $f(x) = 1$  при  $x \leq 1/2$  и  $f(x) = 2x$  при  $x > 1/2$ , в области, определяемой неравенством  $x \geq 0$ , принимает минимальное значение (равное единице) во всех точках отрезка  $[0, 1/2]$ .

Следует различать *абсолютный (глобальный)* и *локальный* экстремумы. В случае достижения функцией  $f(x)$  в точке  $x_0$  абсолютного экстремума (допустим максимума) неравенство  $f(x_0) \geq f(x)$  имеет место для всех точек допустимой области  $M$ . В случае же локального экстремума (в данном случае максимума) это неравенство выполняется для всех достаточно близких к точке  $x_0$  точек области  $M$ . Например  $f(x_0) \geq f(x)$  для всех  $x$  таких, что расстояние  $r(x, x_0)$  от точки  $x$  до точки  $x_0$  меньше некоторого (обычно довольно малого) положительного числа  $\varepsilon$ .

При перемене знака целевой функции  $f(x)$  все точки ее максимума превращаются в точки минимума и наоборот. Поэтому в теории достаточно рассматривать лишь какой-нибудь один из видов оптимума (максимум или минимум).

**Оптимизация функций и оптимизация функционалов.** В практических оптимизационных задачах приходится иметь дело с двумя основными видами их постановок. В первом случае задача ставится в обычном (евклидовом) пространстве конечной размерности. Точками  $x$  допустимого множества будут кортежи  $x = \langle x_1, x_2, \dots, x_n \rangle$  вещественных чисел, целевой же функцией  $f(x) = F(x_1, x_2, \dots, x_n)$  будет обычная вещественная

функция от  $n$  вещественных аргументов ( $n$  - размерность пространства). Такую задачу называют *задачей оптимизации функций*. Во втором случае постановки оптимизационной задачи в качестве допустимого множества выступает некоторое множество  $M$  функций вещественных переменных  $y(x_1, x_2, \dots, x_m)$ , а целевая функция есть некоторый функционал  $F$ , сопоставляющий каждой функции  $y = y(x_1, x_2, \dots, x_m)$  некоторое вещественное число  $F(y)$ . Такую задачу называют *задачей оптимизации функционалов* или *вариационной задачей*.

С абстрактной точки зрения обе указанные постановки идентичны: в обоих случаях речь идет о выборе точки в некотором пространстве, оптимизирующей значение целевой функции. Различие состоит в природе самого пространства, однако этого достаточно, чтобы вызвать существенные различия в практических методах оптимизации, применяемых в этих двух случаях. В принципе вторая постановка может быть с определенными допущениями сведена к первой. Это возможно, если каждая функция  $y(x_1, x_2, \dots, x_m)$ , составляющая допустимое множество, может быть с достаточной для практики степенью точности аппроксимирована множеством своих значений на некотором конечном множестве точек. Впрочем, в большинстве случаев число таких точек должно быть настолько велико, что сведение вариационной задачи к задаче оптимизации функций становится малопрактичным.

**Стандартные формы задач оптимизации функций.** В стандартных формах объектом оптимизации является непрерывная функция  $f(x)$  вещественных переменных  $\langle x_1, x_2, \dots, x_n \rangle$ , допустимая область  $M$  задается конечной системой равенств и неравенств вида  $p(x) = 0$ ,  $q(x) \geq 0$ ,  $r(x) \leq 0$  с непрерывными левыми частями. Если при этом область  $M$  ограничена, то в ней обязательно существует по крайней мере одна точка абсолютного максимума и по крайней мере одна точка абсолютного минимума.

Поскольку перемена знака у левых частей неравенств  $q(x) \geq 0$  и  $r(x) \leq 0$  меняет знаки этих неравенств на противоположные, можно ограничиться одним из этих типов неравенств. Обычно при максимизации пользуются неравенствами вида  $q(x) \geq 0$ , а при минимизации -  $r(x) \leq 0$ . Таким образом возникают две стандартные постановки задачи оптимизации функций.

В задаче максимизации требуется найти максимум функции  $f(x)$  в области  $M$ , заданной системой равенств  $p_i(x) = 0$  ( $i = 1, \dots, l$ ) и неравенств  $p_i(x) \geq 0$  ( $i = l+1, \dots, m$ ).

В задаче минимизации требуется найти минимум функции  $f(x)$  в области  $M$ , заданной системой равенств  $p_i(x) = 0$  ( $i = 1, \dots, l$ ) и неравенств  $p_i(x) \leq 0$  ( $i = l+1, \dots, m$ ).

Легко видеть, что при одновременной перемене знаков у функций  $f(x)$  и  $p_i(x)$  ( $i = 1, \dots, m$ ) задача максимизации переходит в задачу минимизации и наоборот. Поэтому разработку методов оптимизации достаточно вести применительно лишь к одной из этих постановок.

Следует заметить, что *ограничения типа неравенств* можно заменить *ограничениями типа равенств* и простыми координатными неравенствами, введя дополнительные (вещественные) переменные  $z$ . При этом ограничения вида  $p_i(x) \geq 0$  заменятся парой ограничений  $p_i(x) - z = 0$ ,  $z \geq 0$ , а ограничение  $p_i(x) \leq 0$  - парой ограничений  $p_i(x) + z = 0$ ,  $z \geq 0$ . Этот прием называется *элиминацией нетривиальных неравенств*. Его особенно удобно применять, когда по смыслу задачи все точки допустимой области имеют неотрицательные координаты. В результате его применения при этих условиях возникает третья стандартная форма постановки задачи оптимизации функций:

Найти оптимум (максимум или минимум) функции  $f(x)$  в области  $M$ , заданной системой равенств  $p_i(x) = 0$  ( $i = 1, \dots, m$ ) и простых (координатных) неравенств  $x_j \geq 0$  ( $j = 1, \dots, n$ ).

Во всех перечисленных постановках может присутствовать дополнительное требование о том, чтобы все координаты точки оптимума были целыми числами. Это требование превращает задачу *непрерывной оптимизации* в задачу *целочисленной оптимизации*.

В зависимости от того, является целевая функция (критерий эффективности) скалярной или векторной, различают две разновидности оптимизационных задач - задачи **скалярной** и задачи **векторной оптимизации**. Для каждого из выделенных классов задач характерны и свои *методы оптимизации*.

## 2. Скалярная оптимизация

Универсальным методом оптимизации является **метод простого перебора**. При его реализации последовательно просматриваются все допустимые решения и для каждого подсчитывается значение целевой функции с выбором того, для которого это значение экстремально. Недостаток метода очевиден - число допустимых решений велико даже при небольшом числе переменных. На практике обычно используют методы направленного перебора - математические и эвристические.

*По точности получаемых решений* методы делятся на **точные** (обеспечивают нахождение оптимального решения за конечное число шагов) и **приближенные** (после любого конечного числа шагов приводят к решению, в какой-то степени отличному от оптимального).

*По виду получаемых решений* выделяют две группы методов: аналитические и численные.

**Аналитические методы** обеспечивают отыскание решений в виде функции параметров, известных непосредственно из постановки задачи, причем безотносительно к конкретным значениям. К данной группе методов относятся методы классические методы дифференциального и вариационного исчисления.

**Численные методы** позволяют найти решение только для конкретных значений параметров. К ним относятся методы линейного, нелинейного, дискретного (целочисленного), стохастического и динамического программирования, методы регулярного, случайного и эвристического поиска.

Если целевая функция и ограничения линейны, а операция одноэтапная, то можно использовать один из известных методов *линейного программирования*. Данные методы используют одну идею - оптимальное решение находится путем последовательных приближений, для чего задаются некоторым начальным решением (планом), которое затем улучшается в направлении приближения к оптимальному. Линейное программирование является в настоящее время наиболее разработанной и применяемой ветвью математического программирования.

На практике встречается много задач, в которых целевая функция и/или система ограничений содержат выражения, не линейны относительно переменных. Для решения таких задач применяются методы *нелинейного программирования*. Общих методов решения этого типа задач пока не существует, за исключением случаев квадратичной зависимости между критерием и параметрами при линейных ограничениях.

Некоторые задачи могут содержать условие дискретности значений параметров (например, по своей физической сущности параметры должны быть только целыми числами). Решение таких задач осуществляется специальными методами *дискретного (целочисленного) программирования*.

Отыскание решений значительно усложняется, если приходится иметь дело со случайными величинами и функциями. Для подобных задач разрабатываются методы *стохастического программирования*.

Если операция носит многоэтапный характер, то для ее исследования может применяться метод *динамического программирования*. Его сущность состоит в том, оптимальное решение отыскивается не за все этапы одновременно, а последовательно от этапа к этапу. Оптимизация каждого этапа проводится с учетом всех последующих. В основе выбора решений на каждом этапе лежит принцип оптимальности, утверждающий, что если не выбирать наилучшее решение сейчас, то последствий этого нельзя исправить в будущем.

Перечисленные методы относятся к числу точных. Их использование не всегда возможно или целесообразно. Поэтому приходится прибегать к услугам приближенных методов - регулярного, случайного и эвристического поиска.

Методы **регулярного поиска** - это численные методы поиска экстремума функции одной или многих переменных при наличии ограничений на область их изменения. В этих методах используются строго определенные процедуры. В зависимости от того, учитываются результаты предыдущих шагов поиска или нет, различают *активный* и *пассивный* поиск. Теория и алгоритмы поисковых методов хорошо разработаны лишь применительно к унимодальным функциям с одной переменной. Универсальных эффективных методов, гарантирующих отыскание глобального экстремума для других функций, пока не разработано. Наибольшее распространение в этих случаях находят градиентные методы.

В методах **случайного поиска** для поисковых процедур характерно наличие случайности. Случайный поиск может быть *ненаправленным* и *направленным*. При ненаправленном случайном поиске строится последовательность независимых случайных решений, равномерно распределенных в области существования, и находится наилучшее из них. Для получения решения, достаточно близкого к оптимальному, ненаправленный поиск требует большой последовательности решений. В методах направленного поиска при получении каждого последующего решения используются результаты анализа предыдущих решений. Методы случайного поиска отличаются универсальностью и незаменимы при решении сложных оптимизационных задач приближенного характера.

**Эвристические** методы представляют собой набор правил конструирования, сравнения, анализа и отбора вариантов возможных решений, значительно упрощающих и улучшающих процедуру поиска по сравнению с полным перебором. Эти правила называют эвристиками. Эвристики строят по результатам изучения опыта специалистов в той или иной сфере деятельности. Характерная область применения эвристических методов - задачи с нечеткой формальной постановкой и большой размерностью. Недостаток эвристических методов - невысокая точность и невозможность построения оценок близости получаемых решений к оптимальным.

### 3. Векторная оптимизация

В практике оптимизации часто приходится встречаться со случаем, когда вместо одной целевой функции  $f(x)$  задано несколько целевых функций (критериев)  $f_1(x), \dots, f_k(x)$ . Возникающая при этом задача *многокритериальной (векторной) оптимизации* имеет несколько **постановок**.

В одной из них требуется оптимизировать один из критериев, например  $f_1(x)$ , удерживая остальные критерии в заданных границах:  $a \leq f_i(x) \leq b$  ( $i=2,3,\dots,k$ ). В этом случае фактически речь идет об обычной однокритериальной оптимизации. Что же касается неравенств, ограничивающих другие критерии, то их можно рассматривать как дополнительные ограничения на допустимую область  $M$ . Правда на практике в ряде случаев подобная задача решается несколько иными методами, чем обычная однокритериальная задача.

Вторая постановка состоит в упорядочении заданного множества критериев и в последовательной оптимизации по каждому из них. Иными словами, проведя оптимизацию по первому критерию  $f_1(x)$ , получаем некоторое множество  $M_1 \subset M$ , на котором функция  $f_1(x)$  принимает оптимальное (экстремальное) значение. Принимая его за новое допустимое множество, проводим оптимизацию по второму критерию, получая в результате новое допустимое множество  $M_2 \subset M_1$ . Продолжая этот процесс, получим после оптимизации по последнему критерию  $f_k(x)$  множество  $M_k$ , которое и будет конечным результатом многокритериальной оптимизации. Ясно, что если на некотором шаге  $i$  ( $i < k$ ) множество  $M_i$  сведется к одной точке, то процесс оптимизации можно за-

кончить, поскольку  $M_i = M_{i+1} = \dots = M_k$ . Не исключено, что, как и в случае обычной однокритериальной оптимизации, задача может вообще не иметь решения.

Третья постановка применяет процесс сведения многих критериев к одному за счет введения весовых коэффициентов  $\alpha_i$  для каждого из критериев  $f_i(\mathbf{x})$ . В качестве таких коэффициентов в общем случае могут быть выбраны любые вещественные числа. Их значения выбираются из интуитивного представления о степени важности различных критериев: более важные критерии получают веса с большими значениями. После установления весов и нормирования значений целевых функций многокритериальная задача сводится к однокритериальной с целевой функцией  $\alpha_1 f_1(\mathbf{x}) + \dots + \alpha_k f_k(\mathbf{x})$ .

Вместо простой линейной комбинации исходных критериев могут использоваться и более сложные способы формирования из них нового критерия для однокритериальной оптимизации.

Методы решения задач векторной оптимизации будут рассмотрены в одной из следующих лекций.