№ 499

МИСиС

Г.М. Островский
Ю.М. Волин

Методы глобальной оптимизации сложных систем

Учебное пособие

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

№ 499

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
ИНСТИТУТ СТАЛИ и СПЛАВОВ
Технологический университет
МИСиС

Кафедра инженерной кибернетики

Г.М. Островский
Ю.М. Волин

Методы глобальной оптимизации сложных систем

Учебное пособие

Рекомендовано редакционно-издательским
советом института

Москва Издательство «УЧЕБА» 2005

УДК 517.9
О-77

Р е ц е н з е н т
д-р техн. наук, проф., заслуженный деятель
науки РФ *Л.С. Гордеев* (РХТУ им. Д.И. Менделеева)

Островский Г.М., Волин Ю.М.
О-77 Методы глобальной оптимизации сложных систем: Учеб. пособие. – М.: МИСиС, 2005. – 105 с.

Дается элементарное введение некоторых понятий выпуклого анализа как теоретической основы методов глобальной оптимизации. Рассматривается проблема поиска глобального решения в трех классах задач математического программирования: задачах дифференцируемой оптимизации, задачах дискретно-непрерывного программирования и задачах полубесконечного программирования. Описываются детерминированные методы решения этих задач, основанные на идеях метода ветвей и границ. Поскольку эффективность алгоритмов, основанных на методе ветвей и границ, в основном зависит от эффективности процедуры получения нижней оценки (ее точности и трудоёмкости), то большое внимание уделено алгоритмам её получения.

Соответствует государственному образовательному стандарту дисциплины «Современные методы оптимизации сложных систем».

Предназначено для студентов четвёртого курса специальности «Прикладная математика».

ОГЛАВЛЕНИЕ

Введение 4
1. Элементы выпуклого анализа.. 5
 1.1. Выпуклые области, выпуклые функции и их свойства............ 5
 1.2. Многогранник и его свойства... 14
2. Решение задач глобальной оптимизации 28
 2.1. Формулировка задачи... 28
 2.2. Метод ветвей и границ 31
 2.3. Построение выпуклых нижних оценочных функций для некоторого класса функций 42
 2.4. Конструирование выпуклых нижних оценочных функций для произвольных функций 59
 2.5. Использование метода ветвей и границ для решения специальных задач.. 72
 2.6. Метод ветвей и границ уменьшенной размерности 74
 2.7. Метод линеаризации 78
 2.8. Использование методов интервальной математики 80
 2.9. Дискретно-непрерывное нелинейное програмирование......... 86
 2.10. Полубесконечное программирование...................................... 94
Заключение.. 98
Библиографический список 99
Приложение.. 101

ВВЕДЕНИЕ

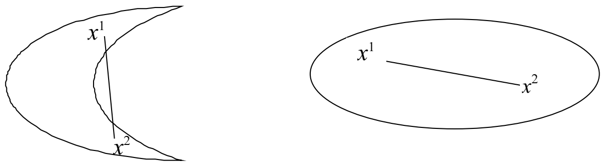
Широкое использование методов математического программирования для решения технических задач началось в 60-е годы прошлого столетия. В настоящее время они стали важной составной частью проектирования и планирования технологических процессов, поскольку целью любого проектирования или планирования является получение оптимальной конструкции, режима, плана. Использование нелинейных математических моделей создает возможность появления многоэкстремальности в задачах оптимизации. Например, проблема поиска глобального оптимума может возникнуть при оптимизации технологического процесса с множественностью стационарных состояний, в задачах проектной оптимизации с невыпуклыми функциями капитальных затрат. Важность этой проблемы для приложений и большая трудоемкость ее решения обусловили интенсивное развитие методов решения задач глобальной оптимизации в последнее время. Если вначале развивались в основном эвристические и стохастические методы глобальной оптимизации, часто не гарантировавшие получения глобального оптимума, то в последнее время большой интерес вызывают детерминированные методы глобальной оптимизации.

Мы рассмотрим здесь проблему глобальной оптимизации для трех классов задач математического программирования: задач дифференцируемой оптимизации, задач дискретно-непрерывного программирования и задач полубесконечного программирования. Основное внимание будет уделено первому классу задач. В настоящее время получили большое развитие методы решения задач глобальной оптимизации, основанные на идеях метода ветвей и границ (branch and bound method). Поэтому здесь будут рассмотрены детерминированные методы решения этих задач, основанные на этих идеях. Основная задача при использовании метода ветвей и границ состоит в разработке эффективного алгоритма получения нижней оценки (при минимизации) целевой функции задачи оптимизации, от которого в значительной степени зависит эффективность всего алгоритма метода ветвей и границ. Эффективность алгоритма получения нижней оценки (границы) характеризуется, прежде всего, точностью получаемой нижней оценки и трудоемкостью ее получения. Здесь будут рассмотрен ряд подходов к получению нижней оценки.

1. ЭЛЕМЕНТЫ ВЫПУКЛОГО АНАЛИЗА*

1.1. Выпуклые области, выпуклые функции и их свойства

Рассмотрим некоторую область D. Если для любой пары точек x^1, $x^2 \in D$ (рис. 1.1) отрезок $[x^1, x^2]$, включающий конечные точки x^1, x^2, принадлежит области D, то область называется выпуклой, в противном случае она называется невыпуклой. На рис. 1.1 область D_1 – невыпуклая, а область D_2 – выпуклая.



D_1 D_2

Рис. 1.1. Невыпуклые и выпуклые области

Дадим теперь определение выпуклой функции. Рассмотрим некоторую одномерную функцию $f(x)$ в интервале $[x^L, x^U]$. Выберем любые две точки $x^1, x^2 \in [x^L, x^U]$. Проведем прямую линию через точки $A_1 = (x^1, f_1)$ и $A_2 = (x^2, f_2)$ такие, что $f_1 = f(x^1)$ и $f_2 = f(x^2)$ (рис. 1.2). Уравнение этой прямой имеет вид

$$\overline{f} = f_1 + \frac{f_2 - f_1}{x^2 - x^1}\left(x - x^1\right).$$

Представим точку внутри отрезка $[x^1, x^2]$ в виде

$$x = \left(1-\alpha\right)x^1 + \alpha x^2, \quad 0 \le \alpha \le 1. \tag{1.1}$$

Это соотношение эквивалентно соотношению

$$x = \lambda_1 x^1 + \lambda_2 x^2, \quad \lambda_1 + \lambda_2 \le 1. \tag{1.2}$$

---

* Здесь рассматриваются некоторые элементы выпуклого анализа, которые будут использоваться при изложении методов глобальной минимизации. Детальное описание этого вопроса может быть найдено в [1].

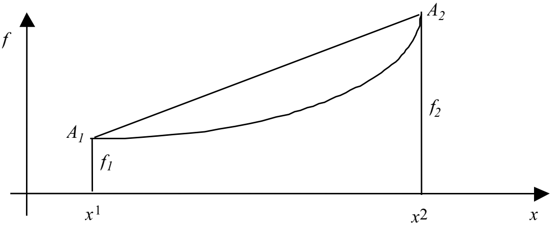


Рис. 1.2. Выпуклая функция

С использованием выражения (1.1) уравнение прямой A_1A_2 может быть преобразовано следующим образом:

$$\overline{f}(\alpha) = f_1 + \frac{f_2 - f_1}{x^2 - x^1}\left[\left(1-\alpha\right)x^1 + \alpha x^2 - x^1\right] =$$
$$= f_1 + \frac{f_2 - f_1}{\left(x^2 - x^1\right)}\alpha\left(x^2 - x^1\right) = \qquad (1.3)$$
$$= f_1 + \left(f_2 - f_1\right)\alpha = \left(1-\alpha\right)f_1 + \alpha f_2, \qquad 0 \le \alpha \le 1.$$

Это параметрическая форма прямой. *Функция, определенная на отрезке [x^L, x^U], называется выпуклой, если для любой пары x^1, x^2 $\in$[x^L, x^U], выполняется следующее условие:*

$$f\left(x\right) \le \overline{f}\left(x\right); \qquad \forall x \in [x^1, x^2], \qquad (1.4)$$

Другими словами, функция $f(x)$ должна лежать ниже прямой, соединяющей точки x^1 и x^2.

Подставив в (1.4) выражения для x и $\overline{f}$ из (1.2) и (1.3), получим другую форму условия выпуклости:

$$f\left[\left(1-\alpha\right)x^1 + \alpha x^2\right] \le \left(1-\alpha\right)f\left(x^1\right) + \alpha f\left(x^2\right),\ 0 \le \alpha \le 1. \qquad (1.5)$$

Функция $f(x)$ называется строго выпуклой, если в (1.5) выполняется строгое неравенство (за исключением концевых точек)

$$f\left[\left(1-\alpha\right)x^1 + \alpha x^2\right] < \left(1-\alpha\right)f\left(x^1\right) + \alpha f\left(x^2\right),\ 0 < \alpha < 1. \qquad (1.6)$$

Вогнутая функция удовлетворяет условию

$$f[(1-\alpha)x^1 + \alpha x^2] \ge (1-\alpha)f\left(x^1\right) + \alpha f\left(x^2\right), 0 \le \alpha \le 1,\ \forall x^1, x^2 \in [x^L, x^U]. \qquad (1.7)$$

Функция $f(x)$ называется квазивыпуклой, если выполняется неравенство

$$f[(1-\alpha)x^1+\alpha x^2]\le\max[f(x^1),\ f(x^2)],\ 0\le\alpha\le1,\ \forall x^1,x^2\in[x^L,x^U].\ (1.8)$$

Функция $f(x)$ называется квазивогнутой, если выполняется неравенство

$$f[(1-\alpha)x^1+\alpha x^2]\ge\min[f(x^1),\ f(x^2)],\ 0\le\alpha\le1,\ \forall x^1,x^2\in[x^L,x^U].\ (1.9)$$

Функция $f(x)$ называется строго квазивыпуклой, если в (1.8) выполняется строгое неравенство (за исключением концевых точек). Аналогично, функция $f(x)$ называется строго квазивогнутой, если в (1.9) выполняется строгое неравенство (за исключением концевых точек). Соотношения (1.5) – (1.9) верны и в n-мерном пространстве. В этом случае x^1, x^2 – точки n-мерного пространства.

Покажем, что класс выпуклых функций включает класс квазивыпуклых функций (однако обратное неверно). Пусть функция $f(x)$ является выпуклой. Предположим, что

$$f(x^2)\ge f(x^1).$$

Для любых двух точек x^1, x^2 условие (1.4) выполняется, поэтому имеем

$$f\left[(1-\alpha)x^1+\alpha x^2\right]\le(1-\alpha)f(x^1)+\alpha f(x^2)\le(1-\alpha)f(x^2)+\alpha f(x^2)=$$
$$=f(x^2)=\max\left[f(x^1),f(x^2)\right].$$

Но это условие есть условие того, что функция $f(x)$ является квазивыпуклой. Вогнутая функция может быть квазивыпуклой функцией. Капитальные затраты технического оборудования часто выражаются с помощью функции x^{α} (где $\alpha < 1$). Легко проверить, что эта функция есть вогнутая функция.

Данные определения выпуклости, вогнутости, квазивыпуклости, квазивогнутости, строгой выпуклости и вогнутости применимы и в многомерном случае. Ясно, что свойство выпуклости (вогнутости) функции зависит от области, в которой эта функция рассматривается.

Рассмотрим теперь второе определение выпуклой функции: *Дифференцируемая функция $f(x)$ является выпуклой, если для любой точки $\bar{x}$ отрезка $[x^L, x^U]$ выполняется условие*

$$f(x) \geq f(\bar{x}) + \left[\operatorname{grad} f(\bar{x})^T (x - \bar{x})\right], \quad \forall x, \bar{x} \in [x^L, x^U]. \qquad (1.10)$$

На рис. 1.3 показана геометрическая интерпретация этого условия. На этом рисунке прямая $y = f(\bar{x}) + (\nabla f(\bar{x}))^T (x - \bar{x})$ есть касательная к функции $f(x)$ в точке $\bar{x}$. Условие (1.10) означает, что в точке $\bar{x}$ функция $f(x)$ должна лежать выше касательной к ней в этой точке. Докажем, что, если выполняется условие (1.10), функция $f(x)$ является выпуклой.

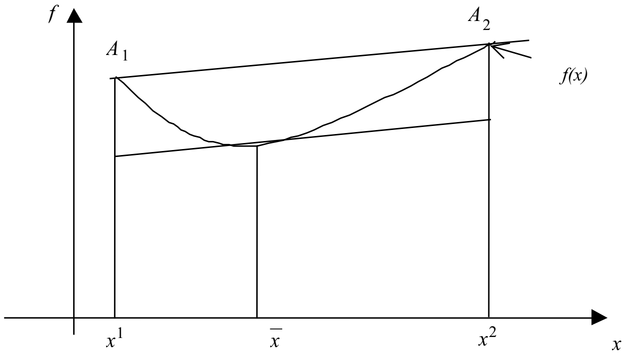


Рис. 1.3. Геометрическая интерпретация условия (1.10)

Предположим противное, т.е. что функция $f(x)$ не является выпуклой. Это значит, что найдутся точки в интервале $[x^1, x^2]$, в которых значение функции $f(x)$ будет выше прямой A_1A_2. Выберем среди них точку $\bar{x}$, в которой значение функции $f(x)$ находится на наибольшем расстоянии от прямой A_1A_2. В этой точке касательная к функции $f(x)$ будет параллельна прямой A_1A_2. Уравнение этой касательной имеет вид

$$f(\bar{x}) + \left[\operatorname{grad} f(\bar{x})\right]^T (x - \bar{x}) = y.$$

В некоторой окрестности D точки $\bar{x}$ все значения функции $f(x)$ находятся ниже касательной, т.е. выполняется условие

$$f(\bar{x}) + \left[\operatorname{grad} f(\bar{x})\right]^T (x - \bar{x}) \leq f(x), \qquad \forall x \in D.$$

Но это условие противоречит условию (1.10). Следовательно, не может существовать точка выше прямой A_1A_2.

Рассмотрим теперь третье определение выпуклой функции: *Дважды дифференцируемая функция f(x) является выпуклой в области D, если гессиан* $\nabla^2 f$ *является положительной полуопределенной матрицей во всех точках этой области.*

Разложим в ряд Тейлора функции $f(x)$ в некоторой точке $\bar{x}$, сохраняя члены не выше второго порядка малости.

$$f(x) = f(\bar{x}) + \nabla f^T (x - \bar{x}) + 0{,}5(x - \bar{x})^T \nabla^2 f(\bar{x})(x - \bar{x}).$$

В соответствии с определением положительной полуопределенной матрицы имеем

$$(x - \bar{x})^T \nabla^2 f(\bar{x})(x - \bar{x}) \geq 0.$$

Таким образом,

$$f(x) - \left[f(\bar{x}) + \nabla f^T (x - \bar{x}) \right] = 0{,}5\ (x - \bar{x})^T \nabla^2 f(\bar{x})(x - \bar{x}) \geq 0.$$

Это условие совпадает с условием (1.10). Поскольку это условие выполняется во всех точках области D, то функции $f(x)$ является выпуклой в этой области.

Теорема 1. Пусть функция $\varphi(x)$ является квазивыпуклой в области D. Тогда область $D = \{x : \varphi(x) \leq 0\}$ является выпуклой.

Доказательство. Возьмем две точки $x^1, x^2 \in D$ такие, что $\varphi(x^1) \leq 0$ и $\varphi(x^2) \leq 0$. В соответствии с определением квазивыпуклой функции для любой точки отрезка $[x^1, x^2]$ удовлетворяется условие

$$\varphi(x) \leq \max\left[\varphi\left(x^1\right), \varphi\left(x^2\right)\right] \leq 0.$$

Отсюда любая точка отрезка $[x^1, x^2]$ принадлежит области D. Таким образом, область D является выпуклой.

Теорема 2. Локальный минимум выпуклой или строго квазивыпуклой функции $f(x)$ в выпуклой, ограниченной области R является глобальным минимумом в этой области.

Доказательство. Рассмотрим вначале случай, когда область R совпадает со всем пространством X, а функция $f(x)$ является выпуклой (случай выпуклой, безусловной минимизации). Докажем