Министерство образования Республики Беларусь

Учреждение образования
«Гомельский государственный технический
университет имени П. О. Сухого»

Кафедра «Высшая математика»

МАТЕМАТИЧЕСКИЕ МЕТОДЫ ПОИСКА ОПТИМАЛЬНЫХ РЕШЕНИЙ

ПРАКТИКУМ
по одноименной дисциплине для студентов специальности 1-43 01 02 «Электроэнергетические системы и сети» дневной формы обучения

Гомель 2017

Библиотека ГГТУ им.П.О.Сухого

УДК 519.85(075.8)
ББК 22.18я73
М34

*Рекомендовано научно-методическим советом
факультета автоматизированных и информационных систем
ГГТУ им. П. О. Сухого
(протокол № 7 от 29.02.2016 г.)*

Составители: *М. В. Задорожнюк*, *Н. Н. Бородин*

Рецензент: доц. каф. «Физика» ГГТУ им. П. О. Сухого канд. физ.-мат. наук *А. И. Кравченко*

М34 **Математические** методы поиска оптимальных решений : практикум по одноим. дисциплине для студентов специальности 1-43 01 02 «Электроэнергетические системы и сети» днев. формы обучения / сост.: М. В. Задорожнюк, Н. Н. Бородин. – Гомель : ГГТУ им. П. О. Сухого, 2017. – 66 с. – Систем. требования: PC не ниже Intel Celeron 300 МГц ; 32 Mb RAM ; свободное место на HDD 16 Mb ; Windows 98 и выше ; Adobe Acrobat Reader. – Режим доступа: https://elib.gstu.by. – Загл. с титул. экрана.

Приведены основные теоретические сведения, касающиеся условной и безусловной оптимизации функций одной и многих переменных, примеры, а также вопросы и задачи для самостоятельного решения.

Для студентов специальности 1-43 01 02 «Электроэнергетические системы и сети» дневной формы обучения.

УДК 519.85(075.8)
ББК 22.18я73

© Учреждение образования «Гомельский государственный технический университет имени П. О. Сухого», 2017

Библиотека ГГТУ им. П.О.Сухого

СОДЕРЖАНИЕ

Библиотека ГГТУ им. П.О. Сухого

ВВЕДЕНИЕ

На практике часто встречается ситуация, когда из нескольких возможных вариантов поведения необходимо выбрать один, в каком-то смысле наилучший. Такой выбор можно осуществлять по-разному. Один из подходов состоит в количественной оценке каждого из возможных вариантов поведения и выборе того из них, у которого оценка наилучшая (минимальная или максимальная, в зависимости от смысла конкретной рассматриваемой задачи). Так мы приходим к задаче оптимизации, которая формулируется следующим образом:

Есть некоторое множество решений, называемых *альтернативами*. Каждой альтернативе можно дать некоторую количественную оценку на основе некого *критерия оптимальности*. *Решение задачи* состоит в определении той альтернативы, для которой критерий оптимальности дает наибольшую (или наименьшую) оценку.

Для решения конкретной инженерной задачи оптимизации необходимо иметь *математическую модель* объекта оптимизации, которая описывает объект с помощью связи между величинами, характеризующими его свойства. Эти величины можно варьировать в некоторых пределах, что и порождает множество альтернатив, из которых предстоит выбрать наилучшую. Изменяемые при оптимизации величины, входящие в математическую модель, называют *параметрами оптимизации*, а соотношения, устанавливающие границы возможного изменения этих параметров, – *ограничениями*.

Хотя прикладные задачи оптимизации относятся к различным областям инженерной практики и представляют различные системы, они имеют общую форму. Все эти задачи можно классифицировать как задачи минимизации вещественнозначной (целевой) функции $f(x)$ n-мерного вектора $x=(x_1,x_2,...,x_n)$, компоненты которого удовлетворяют системе уравнений $h_k(x)=0$, набору неравенств $g_j(x)\geq 0$, а также ограничены сверху и снизу: $x_i^{m_i} \leq x_i \leq x_i^{L_i}$.

Библиотека ГГТУ им. П.О. Сухого

БЕЗУСЛОВНАЯ ОПТИМИЗАЦИЯ

1. Методы одномерной минимизации

Задачи оптимизации, в которой целевая функция выражена функцией одной переменной, относятся к наиболее простому типу оптимизационных задач и занимают центральное место в исследованиях теоретической и практической направленности, т.к. используются для решения подзадач в более сложных многомерных методах. Классификация методов решения одномерных задач основывается на различных предположениях и допущениях относительно природы и свойств целевой функции $f(x)$.

В общем виде задача выглядит следующим образом:

$$f(x) \to \min,$$

$$x \in R \ \ (\text{или } x \in S \subset R)$$

1.1. Критерии оптимальности

Напомним некоторые определения.

Определение 1.1. Функция $f(x)$, определенная на множестве S, достигает своего ***глобального минимума*** в точке $x^{**} \in S$ в том и только в том случае, когда $f(x^{**}) \leq f(x)$ для всех $x \in S$.

Определение 1.2. Функция $f(x)$, определенная на множестве S, достигает своего ***локального минимума*** в точке $x^{*} \in S$ в том и только в том случае, когда $f(x^{*}) \leq f(x)$ для всех $x \in S$, удаленных от x^{*} на расстояние, меньшее е.

Аналогично определяются глобальный и локальный максимумы.

При анализе оптимизационных задач возникают, как правило, два общих вопроса:

1. Вопрос *анализа «в статике»*: как определить, представляет ли данная точка x^{*} оптимальное решение задачи?

2. Вопрос *анализа «в динамике»*: если x^{*} не является точкой оптимума, то какая последовательность действий приведет к получению оптимального решения?

Рассмотрим сначала первый вопрос, т.е. определим *критерии*, позволяющий установить, является ли данное решение оптимальным.

Необходимое условие оптимума функции в точке. Если точка x^{*} является точкой локального максимума (минимума) дифференцируемой функции $f(x)$, то $f'(x^{*}) = 0$.

Библиотека ГГТУ им. П.О.Сухого

Замечание 1.1. Точки, в которых производная равна нулю, называют ***стационарными*** точками функции. Если стационарная точка не соответствует локальному оптимуму, то ее называют точкой перегиба (***седловой*** точкой). ***Критическими*** называют стационарные точки и точки, в которых производная не существует.

Первое достаточное условие оптимума. Если при переходе дифференцируемой функции через стационарную точку x^* производная $f'(x)$ меняет знак с плюса на минус, то x^* является точкой максимума, если с минуса на плюс, то x^* является точкой минимума.

Второе достаточное условие оптимума. Пусть в точке x^* первые $(n-1)$ производные равны нулю, а производная n-го порядка отлична от нуля. Тогда:

1) если n – нечетное, то x^* – седловая точка (не точка оптимума);

2) если n – четное, то x^* – точка локального оптимума, причем если производная $f^{(n)}(x^*) > 0$, то x^* – точка локального минимума, а если производная отрицательна, то x^* – точка локального максимума.

Пример 1.1. Найти точки оптимума функции $y = \frac{x}{3} - \sqrt[3]{x^2}$.

Решение

Вычислим производную:

$$y' = \left(\frac{x}{3} - x^{2/3}\right)' = \frac{1}{3} - \frac{2}{3}x^{-1/3} = \frac{1}{3} - \frac{2}{3\sqrt[3]{x}} = \frac{\sqrt[3]{x} - 2}{3\sqrt[3]{x}}.$$

Найдем критические точки, приравняв к нулю числитель и знаменатель.

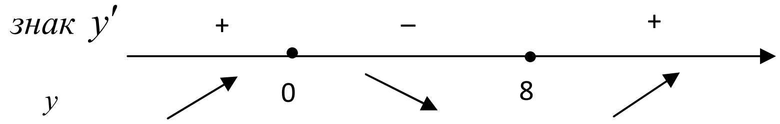
$\sqrt[3]{x} - 2 = 0 \Rightarrow x_1 = 8,$

$\sqrt[3]{x} = 0 \Rightarrow x_2 = 0.$

Обе эти точки являются критическими, т.к. являются внутренними точками области определения функции.

Определим знак производной на каждом из интервалов $(-\infty;0)$, $(0;8)$ и $(8;+\infty)$.

Библиотека ГГТУ им. П.О.Сухого



При переходе через точку $x = 0$ производная меняет знак с плюса на минус, следовательно, в этой точке достигается локальный максимум, $y_{\max} = 0$. При переходе через точку $x = 8$ производная меняет знак с минуса на плюс, значит, $x = 8$ – точка минимума, $y_{\min} = \frac{8}{3} - \sqrt[3]{8^2} = \frac{8}{3} - 4 = -\frac{4}{3}$.

Пример 1.2. Определить, является ли точка $x = 0$ точкой оптимума функции $f(x) = \cos x + \mathrm{ch} x$.

Решение

Найдем значения первой, второй, третьей производных функции в нуле.

$$f'(x) = -\sin x + shx\big|_{x=0} = 0,$$

$$f''(x) = -\cos x + \mathrm{ch}x\big|_{x=0} = 0,$$

$$f'''(x) = \sin x + shx\big|_{x=0} = 0,$$

$$f^{(4)}(x) = \cos x + \mathrm{ch}x\big|_{x=0} = 2 > 0.$$

Т.к. три производные равны нулю, то $n = 4$, следовательно, согласно второму достаточному условию оптимума, $x = 0$ является точкой оптимума. Т.к. четвертая производная положительна, то $x = 0$ – точка минимума.

Следующим шагом в решении оптимизационных задач является отыскание оптимума на отрезке (отыскание глобального минимума (максимума) функции в рассматриваемой области).

Алгоритм отыскания ***глобальных оптимумов*** на отрезке основан на теореме Вейерштрасса.

Теорема 1.1. Непрерывная на отрезке функция достигает своего наибольшего и наименьшего значения либо в стационарных точках внутри отрезка, либо на концах этого отрезка.

Итак, для отыскания глобальных оптимумов необходимо:

1) найти стационарные точки функции;

2) выбрать те из них, которые принадлежат интервалу $(a;b)$;

3) вычислить значения функции в выбранных точках и на концах отрезка;

Библиотека БГТУ им. П.О. Сухого

4) выбрать наименьшее значение (глобальный минимум) и наибольшее значение (глобальный максимум).

Пример 1.3. Найти наибольшее и наименьшее значения функции $y = x^5 - 5x^4 + 5x^3 + 1$ на отрезке $[-1;2]$.

Решение

Вычислим производную функции и найдем стационарные точки.

$y' = 5x^4 - 20x^3 + 15x^2 = 5x^2(x^2 - 4x + 3)$.

Производная обращается в ноль в точках $x_1 = 0$, $x_2 = 1$ и $x_3 = 3$, но $x_3 \notin (-1;2)$. Следовательно, вычисляем значение функции в точках $x_1 = 0$, $x_2 = 1$, а также на концах отрезка. Получим:

$f(0) = 1$,

$f(1) = 1 - 5 + 5 + 1 = 2$,

$f(-1) = -1 - 5 - 5 + 1 = -10$,

$f(2) = 32 - 5 \cdot 16 + 5 \cdot 8 + 1 = -7$.

Таким образом, в точке $x = 1$ функция принимает наибольшее значение $f_{наиб} = f(1) = 2$, а в точке $x = -1$ она принимает наименьшее значение $f_{наим} = f(-1) = -10$.

1.2. Методы поиска оптимальных решений

В предыдущем пункте рассматривался вопрос анализа «в статике» – определить, является ли данное решение оптимальным. Теперь изучим вопрос анализа «в динамике», связанный непосредственно с нахождением оптимального решения. Для этого рассмотрим ряд *методов поиска*, ориентированных на нахождение точки оптимума функции внутри заданного интервала. Необходимость в применении таких методов возникает, например, если уравнение $f'(x) = 0$ не решается простым способом, если функция не имеет аналитического выражения, а задана таблицей значений, полученных в ходе эксперимента, и т.п.

В таких случаях используются методы *прямого поиска* (в них не используются значения производной функции, а используются только значения самой функции). Если же в методе используется первая производная, то его относят к методам первого порядка, если $f''(x)$ – то к методам второго порядка.

Для описания методов поиска нам потребуется понятие унимодальной функции.

Библиотека ГГТУ им. П.О.Сухого

Определение 1.3. Функция $f(x)$ называется ***унимодальной*** на отрезке $[a;b]$ в том и только в том случае, если она монотонна по обе стороны от единственной рассматриваемой на этом отрезке оптимальной точки x^*.

Все одномерные методы поиска основаны на предположении, что исследуемая функция в допустимой области унимодальна.

Можно выделить две группы методов прямого поиска, исходя из двух принципиально различных подходов к решению задачи:

1) методы *пассивного* поиска – *точки*, в которых вычисляются значения функции для определения ее наименьшего значения, *выбираются заранее*;

2) методы *последовательного* поиска – *точки*, в которых вычисляются значения функции, *выбираются последовательно*, исходя из уже выбранных точек и значений функции в них.

Последовательный поиск, как правило, более эффективен (ответ получается за меньшее количество шагов), но пассивный поиск гораздо проще в реализации.

Во всех методах поиска используют один из двух подходов:

1) строят последовательность точек, сходящихся к искомой точке минимума (это общий подход, характерный и для многомерной минимизации);

2) строят последовательность вложенных отрезков, каждый из которых содержит точку минимума функции (этот подход характерен только для одномерной минимизации). Каждый такой отрезок называют ***интервалом неопределенности***.

Построение вложенных отрезков опирается на ***принцип исключения интервала***:

Теорема 1.2. Пусть $f(x)$ унимодальна на отрезке $[a;b]$ и достигает минимума в точке $x^* \in [a;b]$ и пусть $c, d \in [a;b]$, причем $c < d$. Тогда если $f(c) \le f(d)$, то $x^* \in [a;d]$, а если $f(c) \ge f(d)$, то $x^* \in [c;b]$.

Примечание. Если $f(c) = f(d)$, то можно исключить оба крайних интервала, т.к. точка минимума будет располагаться на отрезке $[c;d]$.

Так как все методы поиска относятся к приближенным методам, то для их реализации необходимо:

– определить алгоритм выбора точек x_k;

Библиотека БГТУ им. П.О. Сухого

– определить условие прекращения поиска, т.е. условие, при выполнении которого значение f^* считается найденным с заданной точностью.

Существует множество алгоритмов выбора точек, по сути, этим и различаются методы между собой. Что касается условия прекращения вычислений, то обычно используют следующее: длина последнего интервала неопределенности меньше наперед заданного положительного числа ε.

1.2.1. Метод оптимального пассивного поиска

Пусть функция $f(x)$ унимодальна на отрезке $[0;1]$ и достигает наименьшего значения $f(x^*)$. Требуется найти точку x^* методом оптимального *пассивного поиска*.

Напомним, что при методе пассивного поиска точки (и их количество N) выбирают заранее. Оптимальный пассивный поиск состоит в выборе точек, равномерно расположенных на отрезке, при этом ***скорость сходимости*** (длина последнего N-го интервала неопределенности) определяется формулой:

$$l_N^* = \frac{2}{N+1}.$$

Пример 1.4. Минимизировать функцию $f(x) = (100 - x)^2$ на отрезке $[60;150]$ методом пассивного поиска.

Решение

Возьмем шесть точек (N=6). Тогда отрезок разобьется на 7 интервалов длины $h = \frac{150-60}{7} \approx 12{,}86$. Вычислим значения функции на концах интервала и в семи точках.

$f(60) = 1600$, $f(72{,}86) \approx 736{,}58$, $f(85{,}72) = 203{,}92$, $f(98{,}58) \approx 2{,}02$, $f(111{,}44) \approx 127{,}92$, $f(124{,}30) \approx 590{,}49$, $f(137{,}16) \approx 1380{,}87$, $f(150) = 2500$.

Очевидно, что наименьшее значение функции заключено на отрезке $[85{,}72;111{,}44]$. Длина интервала неопределенности в этом случае равна $l_6^* = \frac{2 \cdot 90}{7} \approx 25{,}7$.

Библиотека ГГТУ им. П.О. Сухого

1.2.2. Методы последовательного поиска

I. Метод дихотомии

Этот метод основан на теореме 1.2 позволяет на каждой итерации исключать ровно половину интервала.

Пусть функция $f(x)$ унимодальна на отрезке $[a;b]$ и достигает наименьшего значения $f(x^*)$. Приведем ***алгоритм*** метода.

Шаг 1. Положим $x_m = \dfrac{a+b}{2}$, $L = b - a$; вычислим $f(x_m)$.

Шаг 2. Положим $x_1 = a + \dfrac{L}{4}$, $x_2 = b - \dfrac{L}{4}$.

Вычислим $f(x_1), f(x_2)$.

Шаг 3. Сравним $f(x_1)$ и $f(x_m)$:

а) если $f(x_1) < f(x_m)$, то x^* находится на промежутке $[a;x_m]$, т.е. положим $b = x_m$. Тогда средней точкой отрезка становится точка x_1, следовательно, положим $x_m = x_1$ и к шагу 5.

б) если $f(x_1) \ge f(x_m)$, то

Шаг 4. Сравним $f(x_2)$ и $f(x_m)$:

а) если $f(x_2) < f(x_m)$, то x^* находится на промежутке $[x_m;b]$, т.е. положим $a = x_m$, $x_m = x_2$ и к шагу 5.

б) если $f(x_2) \ge f(x_m)$, то минимум находится на промежутке $[x_1;x_2]$. Присвоим $a = x_1$, $b = x_2$. Заметим, что x_m продолжает оставаться средней точкой нового интервала.

Шаг 5. Вычислим $L = b - a$. Если величина $|L|$ достаточно мала, то закончим поиск, иначе – к шагу 2.

Замечание 1.2. В этом методе на каждой итерации требуется вычислить ровно два значения функции.

Замечание 1.3. Если проведено N вычислений значений функции, то длина интервала неопределенности равна $\left(\dfrac{1}{2}\right)^{N/2}$ величины исходного интервала, т.е. $(b-a)\left(\dfrac{1}{2}\right)^{N/2}$.

Пример 1.5. Минимизировать функцию $f(x) = (100 - x)^2$ на отрезке $[60;150]$ методом дихотомии.

Решение

Библиотека ГГТУ им. П.О. Сухого

Положим $a = 60, b = 150$. Тогда $L = 90$, $x_m = \frac{60+150}{2} = 105$, $f(x_m) = 25$.

Итерация 1

$x_1 = a + \frac{L}{4} = 60 + \frac{90}{4} = 82{,}5$, $x_2 = b - \frac{L}{4} = 150 - \frac{90}{4} = 121{,}5$.

$f(82{,}5) = 306{,}25 > 25 = f(105)$, $f(127{,}5) = 756{,}25 > 25 = f(105)$.

Следовательно, $a := 82{,}5$, $b := 127{,}5$, $x_m := 105$.

Итерация 2

$L = b - a = 45$, $x_1 = 82{,}5 + \frac{45}{4} = 93{,}75$, $x_2 = 127{,}5 - \frac{45}{4} = 116{,}25$.

$f(93{,}75) = 39{,}06 > 25 = f(105)$, $f(116{,}25) = 264{,}06 > 25 = f(105)$.

Следовательно, $a := 93{,}75$, $b := 116{,}25$, $x_m := 105$.

Итерация 3

$L = b - a = 22{,}5$,

$x_1 = 93{,}75 + \frac{22{,}5}{4} = 99{,}375$, $x_2 = 116{,}25 - \frac{22{,}5}{4} = 110{,}625$.

$f(99{,}375) = 0{,}39 < 25 = f(105)$,

следовательно, $a := 93{,}75$, $b := 105$, $x_m := 99{,}375$.

За три итерации вычислено 6 значений функции, интервал неопределенности уменьшился с 90 до $90 \cdot \left(\frac{1}{2}\right)^3 = 11{,}25$.

II. Метод «золотого сечения»

В предыдущем методе на каждой итерации вычислялось два значения функции. Неплохо было бы иметь такой метод, чтобы приходилось вычислять только одно значение. Можно заметить, что процедура исключения отрезка дает новый отрезок, который содержит внутри себя одну из точек x_1, x_2, x_m, в связи с чем возникает идея использовать эту точку повторно, чтобы сократить количество вычислений функции.

Рассмотрим для удобства функцию $f(x)$, унимодальную на отрезке $[0;1]$. Выберем начальные точки x_1 и x_2 симметрично на расстоянии τ от граничных точек.

Библиотека ГГТУ им.П.О.Сухого

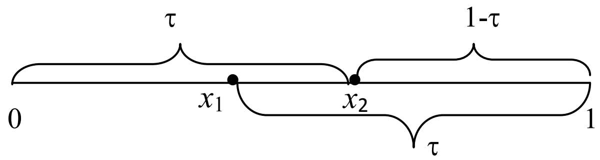


Рис. 1.1

Тогда после процедуры исключения длина оставшегося отрезка в любом случае будет равна τ.

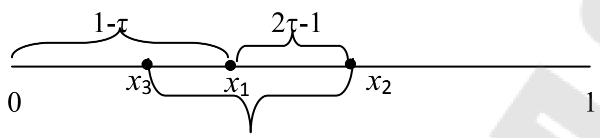


Рис. 1.2

Предположим, исключается интервал $(x_2;1)$. Оставшийся интервал содержит одну пробную точку x_1 на расстоянии $1-\tau$ от левой границы. Чтобы сохранилась симметрия, должно выполняться соотношение

$$\frac{\tau}{1-\tau}=\frac{1-\tau}{2\tau-1},$$

откуда

$$\tau^2+\tau-1=0,$$

$$\tau=\frac{-1\pm\sqrt{5}}{2}\approx 0{,}61803 \text{ – положительное решение.}$$

Если руководствоваться такой стратегией при выборе точек, то величина интервала неопределенности после N вычислений значений функции станет равна

$$l_N^*=\tau^{N-1}.$$

Пример 1.6. Минимизировать функцию $f(x)=(100-x)^2$ на отрезке $[60;150]$ методом «золотого сечения».

Решение

Перейдем сначала к интервалу единичной длины, введя замену переменной

$$w=\frac{x-60}{150-60}=\frac{x-60}{90}.$$

Тогда $x=90w+60$, $f(w)=(40-90w)^2$, $0\le w\le 1$.

$f(0)=1600$, $f(1)=2500$.

Библиотека ГГТУ им. П.О.Сухого

Итерация 1
$I = [0;1]$, $L_1 = 1$.
$w_1 = \tau = 0{,}618$, $w_2 = 1 - \tau = 0{,}382 = \tau^2$.

0 0,382 0,618 1 $f(w_1) = 244$, $f(w_2) = 31{,}6$,

следовательно, исключаем интервал $(0{,}618;1)$.
Итерация 2
$I = [0;0{,}618]$, $L_2 = 0{,}618 = \tau$.
$w_3 = 0{,}618 - \tau^2 0{,}618 - 0{,}382 = \tau^3 = 0{,}236$.

0 0,236 0,382 0,618 $f(w_3) = 352$.

Следовательно, исключаем интервал $(0;0{,}236)$.
Итерация 3
$I = [0{,}236;0{,}618]$, $L_3 = 0{,}382 = \tau^2$.
$w_4 = \tau - (\tau^2 - \tau^3) = 0{,}618 - (0{,}382 - 0{,}236) = 0{,}472$.

0,236 0,382 0,472 0,618 $f(w_4) = 6{,}15$.

Следовательно, исключаем интервал $(0{,}236;0{,}382)$.
Итерация 4
$I = [0{,}382;0{,}618]$, $L_4 = 0{,}236 = \tau^3$.
$w_5 = 0{,}382 + (0{,}618 - 0{,}472) = 0{,}528$

0,382 0,472 0,528 0,618 $f(w_5) = 56{,}55$.

Следовательно, исключаем интервал $(0{,}528;0{,}618)$.
Итерация 5
$I = [0{,}382;0{,}528]$, $L_4 = 0{,}146 = \tau^4$.
$w_6 = 0{,}528 - (0{,}472 - 0{,}382) = 0{,}438$.

0,382 0,438 0,472 0,528 $f(w_6) = 0{,}34$.

Следовательно, исключаем интервал $(0{,}472;0{,}528)$, остается интервал $(0{,}382;0{,}472)$.

Библиотека ГГТУ им.П.О.Сухого

Т.к. $x = 90w + 60$, то $(94{,}38; 102{,}48)$.

После вычисления шести значений функции длина интервала неопределенности стала равна $\tau^5 = 0{,}618^5 = 0{,}09$ (возвращаясь к первоначальной переменной x: $90 \cdot 0{,}09 = 8{,}1$).

1.3. Полиномиальная аппроксимация и методы точечного оценивания

В предыдущем пункте была сделана попытка найти малый *интервал*, в котором находится минимальное значение функции. Попробуем применить другой подход: используем несколько значений функции в определенных точках для аппроксимации функции обычным полиномом по крайней мере на небольшой области значений. Затем положение минимума функции аппроксимируется положением минимума полинома, т.к. последний находить проще. Качество оценки точки оптимума, полученной таким способом, можно повысить за счет использования полинома более высокой степени (однако, при этом сильно возрастает сложность вычислений) или за счет уменьшения интервала аппроксимации. Необходимыми условиями применения методов аппроксимации является непрерывность функции и ее унимодальность.

1.3.1. Квадратичная аппроксимация

Метод основан на том, что если функция принимает минимальное значение во внутренней точке отрезка, то она, по крайней мере, квадратична.

Итак, если известны значения f_1, f_2, f_3 функции $f(x)$ в трех различных точках x_1, x_2, x_3, то функцию $f(x)$ можно аппроксимировать квадратичной функцией вида

$$\varphi(x) = Ax^2 + Bx + \mathrm{C},$$

так что $\varphi(x_i) = f_i$.

На практике этот многочлен удобнее записывать в виде:

$$\varphi(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2).$$

Нетрудно вычислить коэффициенты a_i, $i = 1,2,3$.

$$a_0 = f_1, \tag{1.1}$$

$$a_1 = \frac{f_2 - f_1}{x_2 - x_1}. \tag{1.2}$$

$$a_2 = \frac{1}{x_3 - x_2}\left[\frac{f_3 - f_1}{x_3 - x_1} - \frac{f_2 - f_1}{x_2 - x_1}\right]. \tag{1.3}$$

Библиотека ГГТУ им. П.О.Сухого

Если точность аппроксимации достаточно высока, то построенный полином можно использовать для нахождения точки оптимума (найти производную $\varphi'(x)$, приравнять ее к нулю).

$$\varphi'(x) = a_1 + 2a_2 x - a_2(x_1 + x_2) = 0,$$

$$\bar{x} = \frac{x_1 + x_2}{2} - \frac{a_1}{2a_2} \qquad (1.4)$$

Пример 1.7. Оценить координаты точки минимума функции $f(x) = 2x^2 + \dfrac{16}{x}$ в интервале [1;5].

Решение

Пусть $x_1 = 1,\ x_2 = 3,\ x_3 = 5$.

Тогда $f_1 = 18,\ f_2 = 23{,}33,\ f_3 = 53{,}2$.

Найдем a_1 и a_2 по формулам (1.2) и (1.3).

$$a_1 = \frac{f_2 - f_1}{x_2 - x_1} = \frac{23{,}33 - 18}{3 - 1} = \frac{8}{3},\ a_2 = \frac{1}{5-3}\left[\frac{53{,}2 - 18}{5 - 1} - \frac{8}{3}\right] = \frac{46}{15}.$$

Следовательно, по формуле (1.4)

$$\bar{x} = \frac{3+1}{2} - \frac{8/3}{2 \cdot 46/15} \approx 1{,}565 \text{ – точка минимума,}$$

$$f_{\min} = f(\bar{x}) \approx 15{,}122.$$

(точный минимум достигается в точке $x^* \approx 1{,}5874$).

1.3.2. Кубическая аппроксимация

Можно аппроксимировать функцию $f(x)$ полиномом третьего порядка. При этом используется меньшее количество точек, т.к. можно вычислять не только значения функции, но и ее производной.

Работа алгоритма начинается в произвольной точке x_1. Находим точку x_2, такую, что $f'(x_1)$ и $f'(x_2)$ имеют разные знаки (наша оценка $\bar{x} \in (x_1; x_2)$). Аппроксимирующая кубическая функция выглядит следующим образом:

$$\varphi(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2) + a_3(x - x_1)^2(x - x_2). \qquad (1.5)$$

Параметры a_i подбираются таким образом, чтобы значения $\varphi(x_i) = f_i,\ \varphi'(x_i) = f_i'\ (i = 1,2)$.

Т.к. $\varphi'(x) = a_1 + a_2(x - x_1) + a_2(x - x_2) + a_3(x - x_1)^2 + 2a_3(x - x_1)(x - x_2)$, то приходим к системе уравнений:

Библиотека ГГТУ им. П.О.Сухого

$$\begin{cases} f_1 = \varphi(x_1) = a_0, \\ f_2 = \varphi(x_2) = a_0 + a_1(x_2 - x_1), \\ f_1' = \varphi'(x_1) = a_1 + a_2(x_1 - x_2), \\ f_2' = \varphi'(x_2) = a_1 + a_2(x_2 - x_1) + a_3(x_2 - x_1)^2. \end{cases}$$

Когда найдем a_i, можно с помощью производной оценить координату $\bar{x}$ стационарной точки функции $f(x)$ с помощью многочлена (1.5):

$$\bar{x} = \begin{cases} x_2, \text{если } \mu < 0, \\ x_2 - \mu(x_2 - x_1), \text{если } 0 \le \mu \le 1, \\ x_1, \text{если } \mu > 1, \end{cases} \tag{1.6}$$

где

$$\mu = \frac{f_2' + w - z}{f_2' - f_1' + 2w}, \tag{1.7}$$

$$z = \frac{3(f_1 - f_2)}{x_2 - x_1} - f_1' + f_2', \tag{1.8}$$

$$w = \begin{cases} \sqrt{z^2 - f_1' f_2'}, \text{если } x_1 < x_2, \\ -\sqrt{z^2 - f_1' f_2'}, \text{если } x_2 < x_1. \end{cases} \tag{1.9}$$

Затем снова выбираются две точки – $\bar{x}$ и одна из точек x_1 или x_2 (чтобы производные были противоположны по знаку) и процедура кубической аппроксимации продолжается.

Пример 1.8. Минимизировать функцию $f(x) = 2x^2 + \dfrac{16}{x}$, используя кубическую аппроксимацию с начальной точкой $x_0 = 1$, шагом $\Delta = 1$, параметрами $\varepsilon_1 = 10^{-2}$, $\varepsilon_2 = 3 \cdot 10^{-2}$.

Решение

Итерация 1

Шаг 1: $f'(1) = -12 < 0$, следовательно, $x_1 = 1 + 1 = 2$.

$f'(2) = 4 > 0$, следовательно, $\bar{x} \in (1;2)$. Положим $x_1 = 1$, $x_2 = 2$.

Шаг 2: Вычислим значения функции и производной в точках x_1 и x_2:

$$f_1 = 18,\ f_2 = 16,\ f_1' = -12,\ f_2' = 4.$$

Шаг 3: По формулам (1.6)-(1.9):

Библиотека ГГТУ им. П.О. Сухого

$$z = -2,\ w = 7{,}211,\ \mu = 0{,}4343,$$
$$\bar{x} = 2 - 0{,}4343(2-1) = 1{,}5657.$$

Шаг 4: Т.к. $f(1{,}5657) = 15{,}1219 < 18 = f(x_1)$, то продолжаем поиск.

Шаг 5: Проверка окончания поиска.

Т.к. $f'(1{,}5657) = -0{,}2640 > \varepsilon_1$, то продолжаем итерации.

Т.к. $f'(\bar{x})f'(x_2) < 0$, то положим $x_1 = \bar{x} = 1{,}5657$.

Итерация 2

Шаг 3: По формулам (1.6)-(1.9) имеем:

$$z = -2{,}3296,\ w = 2{,}5462,\ \ \mu = 0{,}9486,$$
$$\bar{x} = 2 - 0{,}4343(2-1) = 1{,}5880.$$

Шаг 4: $f(1{,}5880) = 15{,}1191 < 18 = f(x_1)$, следовательно, продолжаем поиск.

Шаг 5: проверка окончания поиска. Т.к.

$$f'(1{,}5880) = -0{,}0072 < \varepsilon_1,$$
$$\left|\frac{1{,}5880 - 1{,}5657}{1{,}5880}\right| = 0{,}014 < \varepsilon_2,$$

то окончание поиска.

Оценка точки минимума

$$\bar{x} = 1{,}5880,\ f_{\min} = f(\bar{x}) \approx 15{,}1191.$$

Заметим, что за меньшее число итераций получено более точное значение по сравнению с квадратичной интерполяцией.

Напомним, что необходимым условием существования экстремума в точке является $f'(x) = 0$. Если это уравнение достаточно сложно, то нельзя найти точное его решение, но существует множество приближенных методов. Например, *метод Ньютона.*

1.3.3. Метод Ньютона (метод Ньютона-Рафсона)

Предположим, что функция $f(x)$ дважды дифференцируема. Выберем начальное приближение x_1 корня уравнения $f'(x) = 0$. Затем построим линейную аппроксимацию функции $f'(x)$ (по сути, касательную) и точку, в которой эта линейная функция обращается в ноль, примем за следующее приближение x_2.

Библиотека ГГТУ им. П.О. Сухого

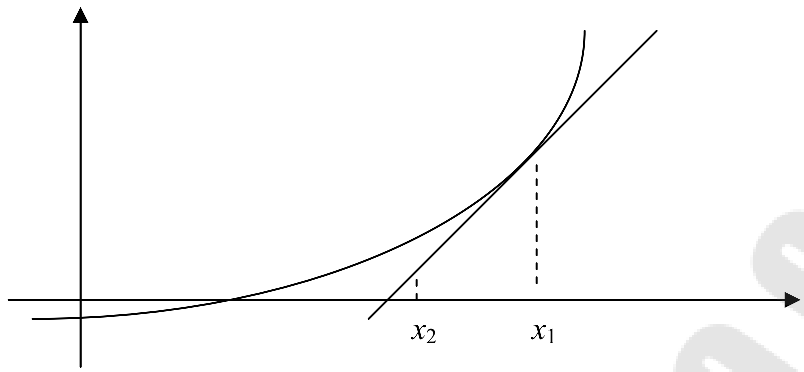


Рис. 1.3

Если x_k – текущая оценка стационарной точки функции $f'(x)$, то аппроксимирующая линейная функция имеет вид:

$$\varphi(x, x_k) = f'(x_k) + f''(x_k)(x - x_k).$$

Приравнивая правую часть к нулю, получим:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}. \qquad (1.10)$$

Итерации продолжаются до тех пор, пока $|f'(x_k)|$ не станет меньше наперед заданной точности ε.

Пример 1.9. Минимизировать функцию $f(x) = 2x^2 + \dfrac{16}{x}$ методом Ньютона.

Решение

Найдем $f'(x) = 4x - \dfrac{16}{x^2}$, $f''(x) = 4 + \dfrac{32}{x^3}$.

В качестве начального приближения возьмем $x_1 = 1$, точность $\varepsilon = 0{,}05$.

Итерация 1

$x_1 = 1$, $f'(1) = -12$, $f''(1) = 36$.

Согласно (1.10) $x_2 = 1 - \dfrac{-12}{36} \approx 1{,}33$. Т.к. $|f'(2)| = |-3{,}73| > \varepsilon$, продолжаем итерации.

Итерация 2

$x_2 = 1{,}33$, $f'(x_2) = -3{,}73$, $f''(x_2) = 17{,}6$.

$x_3 = 1{,}33 - \dfrac{-3{,}73}{17{,}6} \approx 1{,}54$, $|f'(x_2)| \approx |-0{,}59| > \varepsilon$.

Библиотека ГГТУ им. П.О. Сухого

Итерация 3

$x_3 = 1{,}54$, $f'(x_3) \approx -0{,}59$, $f''(x_3) \approx 12{,}76$.

$$x_4 = 1{,}54 - \frac{-0{,}59}{12{,}76} \approx 1{,}59, \ |f'(x_4)| \approx |0{,}03| < \varepsilon .$$

Следовательно, примем за точку минимума $x^* \approx 1{,}59$. При этом $f_{\min} = f(x^*) \approx 15{,}119$.

Вывод: методы *точечного* оценивания, конечно, намного эффективнее методов *интервального* оценивания (даже метод, например, квадратичной аппроксимации сходится быстрее, чем метод золотого сечения). Что касается преимуществ того или иного метода точечного оценивания, то все зависит от вида функции.

Контрольные вопросы и задания

1. Какая функция называется монотонной?
2. Какая точка, называется точкой глобального минимума? локального минимума?
3. Какие точки называются стационарными? критическими?
4. Сформулировать необходимое условие оптимума в точке.
5. Сформулировать первое достаточное условие локального оптимума.
6. Сформулировать второе достаточное условие локального оптимума.
7. Пусть данная точка удовлетворяет достаточным условиям локального минимума. Как установить, является ли этот оптимум глобальным?
8. Заданы следующие функции одной переменной:

а) $f(x) = x^5 + x^4 - \dfrac{x^3}{3} + 2$,

б) $f(x) = (2x+1)^2(x-4)$.

Для каждой из заданных функций найти

1) интервалы возрастания (убывания);
2) критические точки;
3) локальные и глобальный максимум (если таковые имеются);
4) локальные и глобальный минимум (если таковые имеются).

9. Рассмотрите задачу безусловной оптимизации с целевой функцией одной переменной $f(x)$. Пользуясь приведенными в таблице данными о значениях производных порядка 1, 2, 3, 4 в точках $x_i (i = 1,2,\ldots 10)$, идентифицируйте каждую из точек (установите,

Библиотека ГГТУ им. П.О. Сухого

оказывается ли она точкой максимума, минимума, или не является точкой оптимума; укажите случаи, когда нельзя сделать определенный вывод и т.д.).

x_i	$f'(x_i)$	$f''(x_i)$	$f'''(x_i)$	$f''''(x_i)$
x_1	0	+	Нет данных	Нет данных
x_2	0	0	+	Нет данных
x_3	0	–	Нет данных	Нет данных
x_4	–	–	Нет данных	Нет данных
x_5	0	0	–	Нет данных
x_6	0	0	0	–
x_7	0	0	0	0
x_8	0	0	0	+
x_9	+	9	Нет данных	Нет данных
x_{10}	0	–	+	–

10. Исследуйте заданную функцию в указанном интервале:

а) $f(x) = x^3 - 12x + 3,\ -4 \le x \le 4$;

б) $f(x) = \sqrt[3]{2(x+2)^2(x-4)} + 3,\ -4 \le x \le 2$

Найдите локальные минимумы, локальные максимумы, глобальный минимум и глобальный максимум функции f в указанном интервале.

11. Исследовать поведение функции в окрестности заданной точки, используя второе достаточное условие оптимальности:

а) $f(x) = x^2 - 4x - (x-2)\ln(x-1)$, $x_0 = 2$;

б) $f(x) = 6e^{x-2} - x^3 - 3x^2 - 6x$, $x_0 = 2$.

12. Дан шар радиуса R. Указать параметры (высоту и радиус) вписанного в него цилиндра наибольшего объема.

13. Дан цилиндр радиуса R и высоты H. Описать вокруг него конус, имеющий наименьший объем (плоскость и центр их круговых оснований совпадают).

Библиотека ГГТУ им. П.О. Сухого

14. Дан эллипс $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, в который вписан прямоугольник со сторонами, параллельными осям координат. Каковы должны быть стороны прямоугольника, чтобы он имел наибольшую площадь?

15. Из круглого бревна радиуса R вырезана балка прямоугольного сечения, оказывающая наибольшее сопротивление на изгиб. Найти ширину и высоту сечения (сопротивление на изгиб пропорционально произведению ширины сечения на квадрат высоты).

16. Человек стоит напротив картины, закрепленной на вертикальной стене. Нижний край картины расположен выше уровня глаз человека на a, верхний – на b. На каком расстоянии от стены должен стоять человек, чтобы видеть картину под наибольшим углом?

17. Найти точку оптимума (минимума или максимума) функции $y=\mathrm{arctg}x-x^2$, используя метод пассивного поиска, метод дихотомии и метод золотого сечения, проделав пять итераций. Вычислить длины интервалов неопределенности. Сравнить полученные результаты.

18. Найти точку оптимума (минимума или максимума) функции $y=x^4-e^x$, тремя интервальными методами с точностью $\varepsilon=0{,}001$.

19. Решить задачи 17, 18, используя методы точечного оценивания (квадратичную и кубическую аппроксимацию, метод Ньютона).

2. Многомерная безусловная оптимизация

В общем случае задача оптимизации функции нескольких переменных имеет вид:

минимизировать функцию $f(x)$, где $x=(x_1,x_2,...,x_n)$ – вектор переменных.

$$f(x)\to\min,$$
$$x\in R^n.$$

Рассмотрим точные и некоторые приближенные методы решения этой задачи.

2.1. Критерий оптимальности

Рассмотрим вопрос *«анализа в статике»*: является ли данная точка точкой минимума, максимума или седловой точкой. Для определенности рассмотрим задачу минимизации:

Библиотека БГТУ им. П.О. Сухого

$$f(x_1, x_2, ..., x_n) \to \min .$$

Предположим, что функция и все ее частные производные существуют и непрерывны всюду. Обозначим через

$$\nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix} \qquad (2.1)$$

градиент функции $f(x)$. ***Матрица Гессе*** (гессиан) функции $f(x)$ обозначается как $G(x)$ и является симметрической матрицей n-го порядка с элементами вида:

$$G_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}. \qquad (2.2)$$

Необходимое условие оптимума: Если x^* – точка оптимума, то $\nabla f(x^*) = 0$.

Для понимания достаточного условия экстремума, нам понадобятся некоторые факты из теории квадратичных форм.

Выражение вида

$$Q(x_1, x_2, ..., x_n) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{1n}x_n^2 + 2a_{12}x_1x_2 + \\ + 2a_{13}x_1x_3 + \cdots + 2a_{n(n-1)}x_nx_{n-1},$$

где a_{ij} – некоторые числа, называется ***квадратичной формой***. Сокращенно ее можно записать в следующем виде, обозначив $x = (x_1, x_2, ..., x_n)$:

$$Q(x) = x^T Ax .$$

Матрица A называется ***матрицей квадратичной формы***. Говорят, что квадратичная форма $Q(x)$ ***положительно определена,*** если $Q(x) > 0$ для любых x (соответственно, ***отрицательно определена***, если $Q(x) < 0$ для любых x).

Критерий Сильвестра. Квадратичная форма положительно определена тогда и только тогда, когда все угловые миноры ее матрицы A положительны. Квадратичная форма отрицательно определена тогда и только тогда, когда все угловые миноры

Библиотека ГГТУ им. П.О. Сухого

нечетного порядка отрицательны, а угловые миноры четного порядка – положительны.

Отсюда получаем необходимые и достаточные условия оптимума.

Необходимое и достаточное условие минимума: $\nabla f(x^*) = 0$ и матрица Гессе $G(x^*)$ положительно определена.

Необходимое и достаточное условие максимума: $\nabla f(x^*) = 0$ и матрица Гессе $G(x^*)$ отрицательно определена.

Пример 2.1. Исследовать экстремальную точку функции $f(x) = x_1^2 + x_2^2 + x_3^2 - 4x_1 - 8x_2 - 12x_3 + 100$.

Решение

Найдем градиент по формуле (2.1) и приравняем его к нулю:

$$\nabla f = \begin{pmatrix} 2x_1 - 4 \\ 2x_2 - 8 \\ 2x_3 - 12 \end{pmatrix} = 0 \text{ при } x_1 = 2, x_2 = 4, x_3 = 6.$$

Матрицу Гессе найдем по формуле (2.2),

$$G(x) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Она положительно определена, следовательно, в точке (2,4,6) функция достигает минимума.

2.2. Методы прямого поиска

Теперь рассмотрим вопрос «*анализа в динамике*»: если точка $\bar{x}$ не является точкой оптимума, то какие шаги надо предпринять, чтобы найти ее (точку оптимума), хотя бы приближенно?

Так же, как и в случае с функцией одной переменной, существует несколько групп методов:

– методы прямого поиска;

– градиентные методы (методы первого порядка);

– методы второго порядка (используют первые и вторые частные производные).

I. Поиск по симплексу (S^2-метод)

В основе метода лежит идея замены допустимой области дискретным множеством (решеткой) точек пространства с

Библиотека ГГТУ им. П.О.Сухого

последующим ее уменьшением. В качестве экспериментального образца, содержащего наименьшее количество точек, можно использовать так называемый регулярный симплекс. ***Регулярный симплекс*** в n-мерном пространстве – это многогранник, образованный $n+1$ равноотстоящими друг от друга точками-вершинами. Например, в R^2 это равносторонний треугольник, в R^3 – тетраэдр и т.д. В алгоритме используют важное свойство симплексов:

Новый симплекс можно построить на любой грани начального симплекса путем переноса выбранной вершины на надлежащее расстояние вдоль прямой, проведенной через центр тяжести начального симплекса. Полученная таким образом точка является вершиной нового симплекса, а выбранная при построении вершина начального симплекса исключается. При переходе к новому симплексу требуется одно вычисление значения целевой функции. Целевая функция вычисляется во всех точках симплекса, затем выбирается вершина с наибольшим значением $f(x)$. Эта вершина через центр тяжести остальных вершин симплекса проецируется в новую точку, которая используется в качестве вершины нового симплекса.

Если функция убывает достаточно плавно, то итерации продолжаются до тех пор, пока не будет накрыта точка минимума или не начнется циклическое движение по симплексам. В этой ситуации пользуются следующими правилами:

Правило 1. «Накрытие» точки минимума. Если вершина, которой соответствует наибольшее значение целевой функции, построена на предыдущей итерации, то вместо нее берется вершина, которой соответствует следующее по величине значение целевой функции.

Правило 2. Циклическое движение. Если некоторая вершина симплекса не исключается на протяжении более M итераций, то необходимо уменьшить размер симплекса с помощью коэффициента редукции, выбрав в качестве базовой точку, которой соответствует наименьшее значение. Значение M вычисляется по формуле: $M = 1{,}65n + 0{,}05n^2$, где n – размерность задачи (M округляется до целого числа). При применении этого правила требуется задать величину коэффициента редукции.

Правило 3. Критерий окончания поиска. Поиск завершается, когда размеры симплекса или разности между значениями функции в

Библиотека БГУИР

вершинах становятся достаточно малыми. Для применения этого правила необходимо задать величину параметра окончания поиска.

При реализации алгоритма используются вычисления двух типов:

1) построение регулярного симплекса при заданной базовой точке и масштабном множителе;

2) расчет координат отраженной точки.

1) При заданных (базовой) точке $x^{(0)}$ и масштабном множителе α (который подбирается исходя из характеристик задачи) координаты остальных n вершин симплекса вычисляются по формуле:

$$x^{(i)} = \begin{cases} x_j^{(0)} + \delta_1, \text{если } i \neq j, \\ x_j^{(0)} + \delta_2, \text{если } i = j \end{cases} \tag{2.3}$$

где $i, j = \overline{1, n}$.

Приращения δ_1 и δ_2 определяются по формулам:

$$\delta_1 = \left(\frac{\sqrt{n+1} + n - 1}{n\sqrt{2}} \right) \alpha, \ \delta_2 = \left(\frac{\sqrt{n+1} - 1}{n\sqrt{2}} \right) \alpha. \tag{2.4}$$

2) Пусть $x^{(j)}$ – точка, подлежащая отражению. Тогда новая точка $x^{(j)}_{\text{нов.}}$ вычисляется согласно формуле

$$x^{(j)}_{\text{нов.}} = 2x_c - x^{(j)}_{\text{предыдущ.}}, \tag{2.5}$$

где $x_c = \frac{1}{n} \sum\limits_{\substack{i=0, \\ i \neq j}}^{n} x^{(i)}$ – центр тяжести остальных n вершин.

Пример 2.2. Минимизировать функцию

$$f(x) = (1 - x_1)^2 + (2 - x_2)^2.$$

Решение

Выберем базовую точку и масштабный множитель:

$x^{(0)} = (0;0), \alpha = 2$.

Тогда, согласно (2.4),

$$\delta_1 = \left(\frac{\sqrt{2+1} + 2 - 1}{2\sqrt{2}} \right) \cdot 2 \approx 1{,}93, \ \delta_2 = \left(\frac{\sqrt{2+1} - 1}{2\sqrt{2}} \right) \cdot 2 \approx 0{,}52.$$

Вычислим $x^{(1)}$ и $x^{(2)}$ по формулам (2.3):

$x^{(1)} = (x_1^{(0)} + \delta_2; x_2^{(0)} + \delta_1) \approx (0 + 0{,}52; 0 + 1{,}93) = (0{,}52; 1{,}93)$,

Библиотека ГГТУ им. П.О.Сухого

$x^{(2)} = (x_1^{(0)} + \delta_1; x_2^{(0)} + \delta_2) \approx (0 + 1{,}93; 0 + 0{,}52) = (1{,}93; 0{,}52)$.

Итерация 1

$f\left(x^{(1)}\right) \approx 0{,}2353$, $f\left(x^{(2)}\right) \approx 3{,}0553$, $f\left(x^{(0)}\right) = 1 + 4 = 5$ – наибольшее значение, следовательно, «избавляемся» от точки $x^{(0)}$.

Находим центр тяжести вершин $x^{(1)}$ и $x^{(2)}$, а затем новую вершину $x^{(3)}$ по формуле (2.5):

$$x_c = \frac{1}{2}(x^{(1)} + x^{(2)}),$$

$$x^{(3)} = 2x_c - x^{(0)} = x^{(1)} + x^{(2)} - x^{(0)} = (2{,}45; 2{,}45).$$

Т. к. значение функции в новой точке $f\left(x^{(3)}\right) \approx 2{,}31 < f\left(x^{(0)}\right)$, следовательно, точки $x^{(1)}, x^{(2)}, x^{(3)}$ образуют новый симплекс.

Снова находим наибольшее из значений

$f\left(x^{(1)}\right) \approx 0{,}2353$, $f\left(x^{(2)}\right) \approx 3{,}0553$, $f\left(x^{(3)}\right) \approx 2{,}31$.

Т.к. $f\left(x^{(2)}\right) \approx 3{,}0553$ – наибольшее, то на следующей итерации отражаем точку $x^{(2)}$ относительно прямой, проходящей через точки $x^{(1)}$ и $x^{(3)}$.

Итерация 2

$$x^{(4)} = x^{(1)} + x^{(3)} - x^{(2)} = (0{,}52; 1{,}93) + (2{,}45; 2{,}45) - (1{,}93; 0{,}52) = (1{,}04; 3{,}86).$$

$f\left(x^{(4)}\right) \approx 3{,}46$.

Т.к. новое значение функции $f\left(x^{(4)}\right)$ больше, чем предыдущее значение $f\left(x^{(2)}\right)$, то воспользуемся ***правилом 1*** (правилом «накрытия» точки минимума), т.е. на следующей итерации будем отражать не точку $x^{(2)}$, давшую наибольшее значение функции, а точку $x^{(3)}$, давшую следующее по величине значение. Тогда мы снова получим точку $x^{(0)}$, а значит, потребуется применение ***правила 2*** – уменьшение коэффициента редукции.

Итерация 3

Согласно правилу 2, возьмем в качестве базовой точку с наименьшим значением функции, т.е. $x^{(1)} = (0{,}52; 1{,}93)$ (назовем ее $x_н^{(0)}$, чтобы было удобнее пользоваться формулами), а коэффициент α примем равным 1. Вычислим δ_1 и δ_2 по формулам (2.4):

Библиотека ГГТУ им. П.О.Сухого

$$\delta_1 = \left(\frac{\sqrt{2+1}+2-1}{2\sqrt{2}} \right) \cdot 1 \approx 0{,}97,\ \delta_2 = \left(\frac{\sqrt{2+1}-1}{2\sqrt{2}} \right) \cdot 1 \approx 0{,}26.$$

Тогда согласно (2.3)

$$x^{(1)} = (x_1^{(0)} + \delta_2; x_2^{(0)} + \delta_1) \approx (0{,}52 + 0{,}26; 1{,}93 + 0{,}97) = (0{,}78; 2{,}90),$$

$$x^{(2)} = (x_1^{(0)} + \delta_1; x_2^{(0)} + \delta_2) \approx (0{,}52 + 0{,}97; 1{,}93 + 0{,}26) = (1{,}49; 2{,}19).$$

Вычисляем значения функции в вершинах симплекса:

$$f\left(x^{(1)}\right) = 0{,}8584,\ f\left(x^{(2)}\right) = 0{,}2762,\ f\left(x_н^{(0)}\right) = 0{,}2353.$$

Наибольшее значение функции достигается в точке $x^{(1)} = (0{,}78; 2{,}90)$, следовательно, на следующей итерации отражаем ее относительно прямой, проходящей через точки $x_н^{(0)}$ и $x^{(2)}$:

$$x^{(3)} = 2x_c - x^{(0)} = x_н^{(0)} + x^{(2)} - x^{(1)} =$$
$$= (0{,}52; 1{,}93) + (1{,}49; 2{,}19) - (0{,}78; 2{,}90) = (1{,}23; 1{,}22).$$

Т.к. $f\left(x^{(3)}\right) = 0{,}6613 < f\left(x^{(1)}\right)$, то новый симплекс будет образован точками $x_н^{(0)}$, $x^{(2)}$ и $x^{(3)}$. Т.к. наибольшее значение достигается в точке $x^{(3)}$, а ее отражение приведет к зацикливанию, то нужно снова воспользоваться ***правилом 2*** и уменьшить коэффициент редукции. После этого выбрать новую базовую точку (в этом примере это снова будет точка $x_н^{(0)}$, т.к. она дает наименьшее значение функции), пересчитать новые значения δ_1 и δ_2 по формулам (2.4) при коэффициенте α, равном, например, $0{,}5$.

Продолжаем процесс до тех пор, пока не потребуется применение ***правила 3*** (размеры симплекса станут слишком малы или станет мала разность значений функции в вершинах симплекса).

II. Метод Хука-Дживса

Изложенный выше метод поиска основан на операциях над образцами, составленными из пробных точек. Но очевидно, что эти точки нужны лишь для определения направления, в котором дальше вести поиск, само расположение точек не существенно. Следовательно, можно усовершенствовать поиск путем введения системы векторов, задающих направление поиска. Простейший подход заключается в простом переборе n возможных направлений (например, вдоль осей координат). Тогда на каждом этапе можно менять только одну переменную, и таким образом, свести задачу к

Библиотека ГГТУ им. П.О. Сухого

минимизации функции одной переменной. Но если функция «неудобная» (слишком искривленная или, наоборот, слишком сильно растянутая), то такой метод неэффективен, а может вообще расходиться.

Поэтому Хук и Дживс предложили в определении нового направления поиска учитывать информацию, полученную на предыдущих этапах. Метод представляет собой комбинацию «***исследующего поиска***» и «***поиска по образцу***».

«***Исследующий поиск***» начинается в некоторой исходной точке. Задается шаг Δ (он может меняться в процессе решения задачи) и вычисляется значение целевой функции в некоторой пробной точке. Если оно меньше предыдущего значения функции в исходной точке, то шаг рассматривается как успешный, иначе делаем шаг в противоположном направлении и вычисляем значение целевой функции. Перебираем так все n координатных направлений, полученную в результате «наилучшую» точку назовем ***базовой***.

«***Поиск по образцу***» заключается в реализации единственного шага из полученной базовой точки $x^{(k)}$ вдоль прямой, соединяющей эту точку с предыдущей базовой точкой $x^{(k-1)}$. Новая точка определяется по формуле

$$x_p^{(k+1)} = x^{(k)} + (x^{(k)} - x^{(k-1)}) . \tag{2.6}$$

Как только движение по образцу не приводит к уменьшению целевой функции, точка $x_p^{(k+1)}$ фиксируется в качестве новой временной базовой и вновь проводится «исследующий поиск». Если в результате находим новую точку с меньшим значением целевой функции, то получаем новую базовую точку и снова осуществляем «поиск по образцу». Если не находим, то следует вернуться в предыдущую точку $x^{(k)}$ и провести исследование с целью выявления нового направления минимизации. В конце концов, возникает ситуация, когда такой поиск не приводит к успеху. Тогда уменьшаем шаг $\left(\Delta := \dfrac{\Delta}{\alpha}\right)$ и возобновляем поиск. Процесс завершается, когда шаг становится достаточно малым.

Пример 2.3. Минимизировать функцию

$$f(x) = 8x_1^2 + 4x_1x_2 + 5x_2^2 ,$$

используя начальную точку $x^{(0)} = (-4;-4)$.

Библиотека ГГТУ им. П.О. Сухого

Решение

Зададим начальный шаг $\Delta = (1;1)$,точность $\varepsilon = 10^{-4}$, параметр уменьшения шага $\alpha = 2$.

Итерация 1

Вычислим значение $f\left(x^{(0)}\right) = 272$.

Делаем шаг вдоль оси x_1:

$x_1 := x_1 + \Delta_1 = -4 + 1 = -3$, $x_2 := -4$, $f(-3;-4) = 200 < f\left(x^{(0)}\right)$.

Делаем шаг из полученной точки вдоль оси x_2:

$x_1 := x_1 = -3$, $x_2 := x_2 + \Delta_2 = -4 + 1 = -3$, $f(-3;-3) = 153 < 200$, следовательно, шаг считается успешным, в результате исследующего поиска найдена точка $x^{(1)} = (-3;-3)$.

Переходим к поиску по образцу согласно формуле (2.6):

$x_p^{(2)} = x^{(1)} + (x^{(1)} - x^{(0)}) = (-2;-2)$. Т.к. $f\left(x_p^{(2)}\right) = 68 < 153$, то проводим исследующий поиск вокруг точки $x^{(2)}$.

Снова делаем шаг вдоль оси x_1:

$x_1 := x_1 + \Delta_1 = -2 + 1 = -1$, $x_2 := -2$, $f(-1;-2) = 20 < 68$, шаг расцениваем как успешный.

Делаем шаг из полученной точки вдоль оси x_2:

$x_1 := x_1 = -1$, $x_2 := x_2 + \Delta_2 = -2 + 1 = -1$, $f(-1;-1) = 17 < 20$, следовательно, шаг считается успешным, в результате исследующего поиска найдена новая базовая точка $x^{(2)} = (-1;-1)$. Переобозначив $x^{(0)} := x^{(1)} = (-2;-2)$, $x^{(1)} := x^{(2)} = (-1;-1)$ и используя поиск по образцу, найдем новое значение $x_p^{(2)} = x^{(1)} + (x^{(1)} - x^{(0)}) = (0;0)$.

2.3. Методы первого и второго порядка

В прошлом пункте были рассмотрены методы прямого поиска. Достоинством этих методов является то, что их применение позволяет получить решение задачи на основе использования только значений целевой функции. Теперь рассмотрим методы с использованием частных производных. В отличие от прямых методов, где мы перебирали все направления осей, чтобы выбрать «наилучшее», в градиентных методах мы имеем возможность выбирать «наилучшее из приемлемых».

В описанных ниже методах предполагается, что функция $f(x)$, градиент $\nabla f(x)$ и матрица Гессе $G(x)$ существуют и непрерывны и что компоненты градиента могут быть записаны в аналитическом

Библиотека ГГТУ им. П.О.Сухого

виде либо с достаточной степенью точности вычислены при помощи численных методов.

Все градиентные методы основаны на итерационной процедуре, реализуемой формулой

$$x^{(k+1)} = x^{(k)} - \text{б}^{(k)} s(x^{(k)}), \tag{2.7}$$

где $x^{(k)}$ – текущее приближение к решению x^*, $\text{б}^{(k)}$ – параметр, характеризующий длину шага, $s\left(x^{(k)}\right) = s^{(k)}$ – направление поиска в n-мерном пространстве управляемых переменных. Способ определения $s(x)$ и б на каждой итерации связан с особенностями применяемого метода.

I. Метод наискорейшего спуска (метод Коши)

Как известно, градиент указывает направление наибольшего роста функции, следовательно, наименьшее значение соответствует противоположному направлению, т.е. $s(x) = -\nabla f(x)$, поэтому в основе простейшего градиентного метода лежит формула

$$x^{(k+1)} = x^{(k)} - \text{б}\nabla f(x^{(k)}), \tag{2.7}$$

где α – заданный положительный параметр. Но недостатками такого метода являются медленная сходимость и трудности при подборе подходящего α. Поэтому лучше подбирать α на каждой итерации. Тогда формула (3.16) примет вид:

$$x^{(k+1)} = x^{(k)} - \text{б}^{(k)}\nabla f(x^{(k)}). \tag{2.8}$$

Значения $\alpha^{(k)}$ вычисляются путем решения задачи одномерной минимизации функции $f\left(x^{(k+1)}\right)$ вдоль направления $\nabla f\left(x^{(k)}\right)$ с помощью какого-либо метода одномерного поиска. Такой метод носит название метода наискорейшего спуска (или метода Коши). К «плюсам» этого метода можно отнести его устойчивость, однако он по-прежнему очень медленно сходится. Обычно метод наискорейшего спуска используется в качестве начальной процедуры перед применением других методов, т.к. позволяет существенно уменьшить значение целевой функции.

Пример 2.4. Минимизировать функцию

$$f(x) = 8x_1^2 + 4x_1x_2 + 5x_2^2.$$

Решение

Библиотека ГГТУ им. П.О.Сухого

Найдем градиент $\nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 16x_1 + 4x_2 \\ 4x_1 + 10x_2 \end{pmatrix}$.

Зададим начальное приближение $(-4;-4) = x^{(0)}$.

Итерация 1

Построим новое значение $x^{(1)}$ по формуле (2.8):

$$x^{(1)} = x^{(0)} - б^{(0)}\nabla f(x^{(0)}) = (-4;-4) - \alpha^{(0)}(-80;-56) =$$
$$= (-4 + 80\alpha^{(0)}; -4 + 56\alpha^{(0)}).$$

Подставим координаты точки $x^{(1)}$ в функцию $f(x) = 8x_1^2 + 4x_1x_2 + 5x_2^2$. Получим

$$f(\alpha) = 8(-4 + 80\alpha^{(0)})^2 + 4(-4 + 80\alpha^{(0)})(-4 + 56\alpha^{(0)}) + 5(-4 + 56\alpha^{(0)}).$$

Это функция одной переменной $\alpha^{(0)}$, и мы можем найти ее точку минимума, вычислив производную и приравняв ее к нулю:

$$f'(\alpha) = 16 \cdot 80(-4 + 80\alpha^{(0)}) + 4(2 \cdot 80 \cdot 56\alpha^{(0)} - 4 \cdot 56 - 4 \cdot 80) +$$
$$+ 10 \cdot 56(-4 + 56\alpha^{(0)}) = 0.$$

После раскрытия скобок и преобразования выражения получим уравнение

$$2650\alpha^{(0)} = 149, \text{ откуда } \alpha^{(0)} \approx 0{,}056.$$

Подставив в (2.8), найдем новое значение

$$x^{(1)} \approx (0{,}498;-0{,}864),\ f\left(x^{(1)}\right) \approx 3.87738$$

Итерация 2

$$x^{(2)} = x^{(1)} - б^{(1)}\nabla f(x^{(1)}) = (0{,}498;-0{,}864) - \alpha^{(1)}(4{,}512;-6{,}648) =$$
$$= (0{,}498 - 4{,}512\alpha^{(1)}; -0{,}864 + 6{,}648\alpha^{(1)}).$$

Подставив $x^{(2)}$ в функцию, найдя производную и приравняв ее к нулю, получим значение

$б^{(1)} \approx 0{,}13$, откуда $x^{(2)} \approx (-0{,}089;0{,}000)$, $f\left(x^{(2)}\right) \approx 0.063368$.

Замечание 2.1. В качестве условия окончания поиска часто используется оценка относительной погрешности:

$$\frac{\left\| x^{(k+1)} - x^{(k)} \right\|}{\left\| x^{(k)} \right\|} < \varepsilon.$$

Библиотека ГГТУ им. П.О. Сухого

II. Метод Ньютона

Метод Коши основывается на последовательной линейной аппроксимации целевой функции и требует вычисления значений функции и ее первых производных на каждой итерации. Чтобы построить более общую стратегию, нужно привлечь информацию о вторых производных целевой функции. С этой задачей справляется *метод Ньютона*, который реализуется по формуле:

$$x^{(k+1)} = x^{(k)} - G^{-1}\left(x^{(k)}\right) \cdot \nabla f(x^{(k)}) . \qquad (2.9)$$

Пример 2.5. Минимизировать функцию

$$f(x) = 8x_1^2 + 4x_1x_2 + 5x_2^2 .$$

Решение

Вычислим градиент ∇f и матрицу Гессе $G(x)$ целевой функции:

$$\nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 16x_1 + 4x_2 \\ 4x_1 + 10x_2 \end{pmatrix}, \; G(x) = \begin{pmatrix} 16 & 4 \\ 4 & 10 \end{pmatrix}.$$

При $x^{(0)} = (-4;-4)$ имеем

$$\nabla f(x^{(0)}) = \begin{pmatrix} -80 \\ -56 \end{pmatrix},$$

$$G^{-1}(x) = \frac{1}{\Delta}\begin{pmatrix} G_{11} & G_{21} \\ G_{12} & G_{22} \end{pmatrix} = \frac{1}{144}\begin{pmatrix} 10 & -4 \\ -4 & 16 \end{pmatrix}.$$

Тогда из формулы (2.9) получаем:

$$x^{(1)} = (-4;-4) - \frac{1}{144}\begin{pmatrix} 10 & -4 \\ -4 & 16 \end{pmatrix} \cdot \begin{pmatrix} -80 \\ -56 \end{pmatrix} =$$

$$(-4;-4) - \frac{1}{144}(-576;-576) = (0;0).$$

Таким образом, задача минимизации квадратичной функции решается методом Ньютона с помощью одной итерации при любой начальной точке.

Контрольные вопросы и задания

1. Как определяется градиент функции?
2. Как записывается матрица Гессе?
3. Сформулировать необходимое условие существования оптимума функции многих переменных в точке.

Библиотека БНТУ им. П.О. Сухого

4. Сформулировать достаточные условия существования максимума и минимума функции многих переменных в точке.

5. Сформулировать критерий Сильвестра.

6. Дать определение квадратичной формы.

7. Какая квадратичная форма называется положительно определенной? отрицательно определенной?

8. Найти точки экстремума функции (либо доказать, что их нет):

а) $f(x, y) = x^3 + y^3 - 3xy$;

б) $f(x, y) = x\sqrt{y} - x^2 - y + 6x + 3$;

в) $f(x, y) = x^3 + y^2 - 6xy - 39x + 18y + 20$;

г) $f(x, y) = xy(12 - x - y)$;

д) $f(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - x_1x_3 + 4x_2x_3$.

9. В результате поиска минимума функции

$f(x_1, x_2, x_3) = \left[x_1^2 + (x_2 + 1)^2\right]\left[x_1^2 + (x_2 - 1)\right]^2$

Найдены следующие точки:

а) $x^{(1)} = (0;0)$; б) $x^{(2)} = (0;1)$; в) $x^{(3)} = (0;-1)$; $x^{(4)} = (0;1)$.

Классифицировать полученные точки.

10. Описать алгоритм поиска по симплексу.

11. Привести примеры градиентных методов оптимизации. В чем суть этих методов? Каковы достоинства и недостатки методов первого порядка по сравнению с прямыми методами поиска?

12. Описать алгоритм метода Хука-Дживса.

13. В чем состоит суть метода Ньютона?

14. Рассматривается функция Розенброка $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$ и начальная точка $x^{(0)} = (-1{,}2;0)$. Найти точку x^*, которой соответствует минимальное значение $f(x^*)$, пользуясь:

а) методом поиска по симплексу (провести четыре итерации);

б) методом Хука-Дживса;

в) методом наискорейшего спуска.

15. Найти точку максимума функции $f(x, y) = 4(x - y) - x^2 - y^2$ методом поиска по симплексу и методом наискорейшего спуска, взяв в качестве начального приближения точку $(1;1)$. Точность принять равной 0,1.

Библиотека ГГТУ им. П.О.Сухого

16. Найти точку максимума функции $f(x,y) = 2xy - 5x^2 - 3y^2 + 2$ методами Ньютона и Хука-Дживса, взяв в качестве начального приближения точку $(2;2)$. Точность принять равной 0,1. Какой метод для данной задачи сходится быстрее? Почему?

ОПТИМИЗАЦИЯ ПРИ НАЛИЧИИ ОГРАНИЧЕНИЙ

Многие инженерные задачи связаны с оптимизацией при наличии ограничений. Такие ограничения существенно уменьшают размер области, в которой проводится поиск оптимума, однако, процесс оптимизации становится более сложным, т.к. критерии оптимальности для решения задач безусловной минимизации могут не работать.

3. Аналитические методы нелинейного программирования. Функция Лагранжа

3.1. Задачи с ограничениями в виде равенств

Рассмотрим общую задачу оптимизации, содержащую ограничения в виде равенств.

$$f(x_1, x_2, \ldots, x_n) \to \min, \tag{3.1}$$

$$\begin{cases} h_1(x_1, x_2, \ldots, x_n) = 0, \\ h_2(x_1, x_2, \ldots, x_n) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ h_k(x_1, x_2, \ldots, x_n) = 0. \end{cases} \tag{3.2}$$

Уравнения (3.2) носят название ***уравнений связи.***

Для решения задачи (3.1)-(3.2) можно воспользоваться ***методом исключений***: выразить k переменных из равенств-ограничений и подставить в целевую функцию. Таким образом, задачу условной минимизации функции n переменных можно свести к задаче безусловной оптимизации функции $n-k$ переменных, для решения которой можно воспользоваться критерием оптимальности, рассмотренным в §2.

Пример 3.1. Минимизировать функцию $f(x) = x_1 x_2 x_3$ при наличии ограничения $x_1 + x_2 + x_3 - 1 = 0$.

Решение

Библиотека БГУИР

Выразим x_3 из уравнения связи и подставим в целевую функцию.

$x_3 = 1 - x_1 - x_2$,

$f(x_1, x_2) = x_1 x_2 - x_1^2 x_2 - x_1 x_2^2$

Необходимое условие оптимальности:

$$\begin{cases} \dfrac{\partial f}{\partial x_1} = x_2 - 2x_1 x_2 - x_2^2 = 0, \\ \dfrac{\partial f}{\partial x_2} = x_1 - 2x_1 x_2 - x_1^2 = 0 \end{cases}$$

Решениями этой системы являются точки $(0;0), (0;1/2), (1/2;0), (1/3;1/3)$. Проверкой убеждаемся, что наименьшее значении функции $f_{\min} = 0$ и достигается оно в первых трех точках.

Однако метод исключения переменных применим в том случае, когда уравнения, представляющие ограничения, разрешимы относительно некоторого набора переменных. Если же ограничений много и они достаточно сложны, то лучше воспользоваться *методом множителей Лагранжа*.

В соответствии с ***методом множителей Лагранжа*** задача (3.1)-(3.2) преобразуется в эквивалентную задачу безусловной минимизации, в которой фигурируют неизвестные параметры, называемые ***множителями Лагранжа***.

Рассмотрим задачу минимизации целевой функции с одним ограничением.

$$\begin{aligned} &f(x_1, x_2, \ldots, x_n) \to \min, \\ &h(x_1, x_2, \ldots, x_n) = 0. \end{aligned} \tag{3.3}$$

Новая задача будет иметь вид:

$$F(x, \lambda) = f(x) + \lambda h(x) \to \min \tag{3.4}$$

Функция $F(x, \lambda)$ называется ***функцией Лагранжа***, параметр λ – ***множителем Лагранжа***.

Если точка x^0 является решением задачи (3.3), то, очевидно, минимум функции $f(x)$ совпадает с минимумом функции Лагранжа

Библиотека ГГТУ им. П.О. Сухого

$F(x, \lambda)$. Обратно, точки минимума функции Лагранжа удовлетворяют необходимому условию оптимальности, т.е. частные производные функции Лагранжа по всем переменным (включая λ) должны быть равны нулю. Это влечет за собой выполнение условия $h(x_1, x_2, \ldots, x_n) = 0$ и, таким образом, минимум функции Лагранжа совпадает с минимумом функции $f(x)$.

Проиллюстрируем это на конкретном примере.

Пример 3.2. Минимизировать функцию

$$f(x_1, x_2) = x_1 + 3x_2 \to \min$$

при ограничении $x_1^2 + x_2^2 = 10$.

Решение

Запишем функцию Лагранжа:

$$F(x_1, x_2, \lambda) = x_1 + 3x_2 + \lambda\left(x_1^2 + x_2^2 - 10\right).$$

Согласно необходимому условию оптимальности, найдем частные производные и приравняем их к нулю.

$$\begin{cases} F'_{x_1} = 1 + 2\lambda x_1 = 0, \\ F'_{x_2} = 3 + 2\lambda x_1 = 0, \\ F'_{\lambda} = x_1^2 + x_2^2 - 10 = 0. \end{cases}$$

Отсюда $x_1 = -\dfrac{1}{2\lambda}$, $x_2 = -\dfrac{3}{2\lambda}$.

Подставив в третье уравнение, получаем $\lambda^2 = \dfrac{1}{4}$, откуда $\lambda = \pm\dfrac{1}{2}$.

Для $\lambda_1 = \dfrac{1}{2}$ имеем $x_1 = -1$, $x_2 = -3$. Подставим $\lambda_1 = \dfrac{1}{2}$ в функцию Лагранжа. Получим функцию двух переменных x_1, x_2:

$$F(x_1, x_2) = x_1 + 3x_2 + \frac{1}{2}\left(x_1^2 + x_2^2 - 10\right).$$

Построим матрицу Гессе (матрицу вторых частных производных):

$$G = \begin{pmatrix} 2\lambda & 0 \\ 0 & 2\lambda \end{pmatrix}$$

Библиотека ГГТУ им. П.О. Сухого

В точке $M_1(-1;-3)$ матрица Гессе $G=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ положительно определена, а значит, точка $M_1(-1;-3)$ является точкой локального минимума, $f_{\min}=-10$.

Для $\lambda_2=-\dfrac{1}{2}$ имеем $x_1=1,\ x_2=3$,

$$F(x_1,x_2)=x_1+3x_2-\frac{1}{2}\left(x_1^2+x_2^2-10\right).$$

Матрица Гессе в точке $M_2(1;3)$ примет вид $G=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$. Она является отрицательно определенной, а значит, точка $M_2(1;3)$ является точкой локального максимума, $f_{\max}=10$.

Замечание 3.1. При построении матрицы Гессе нужно рассматривать функцию Лагранжа как функцию ***двух*** переменных_x_1 и x_2. Если же мы рассмотрим функцию Лагранжа как функцию ***трех*** переменных $F(x_1,x_2,\lambda)$, то точки $(1/2;-1;-3)$ и $(-1/2;1;3)$ не будут являться ее точками безусловного оптимума.

Замечание 3.2. Метод множителей Лагранжа можно обобщить на случай нескольких ограничений.

Для решения общей задачи (3.1)-(3.2) требуется:

1) Построить функцию Лагранжа:

$$F(x,\lambda)=f(x)+\sum_{j=1}^{k}\lambda_j h_j(x) \tag{3.5}$$

2) Записать и решить систему

$$\begin{cases}\dfrac{\partial F(x,\lambda)}{\partial x_i}=0,\ i=1,2\ldots n,\\ h_j(x)=0,\ j=1,2\ldots k.\end{cases}$$

3) Используя методы безусловной оптимизации, исследовать характер найденных точек оптимума функции Лагранжа как функции $F(x)$ переменой x при найденных значениях λ.

Библиотека ГГТУ им. П.О. Сухого

3.2. Задачи с ограничениями в виде неравенств. Задача Куна-Таккера

Кун и Таккер обобщили рассмотренный подход для решения общих задач нелинейного программирования с ограничениями как в виде равенств, так и в виде неравенств.

Рассмотрим общую задачу нелинейного программирования:

$$f(x_1, x_2, \ldots, x_n) \to \min, \tag{3.6}$$

$$g_l(x_1, x_2, \ldots, x_n) \geq 0, l = 1,2,\ldots m, \tag{3.7}$$

$$h_j(x_1, x_2, \ldots, x_n) = 0,\ j = 1,2,\ldots k. \tag{3.8}$$

Ограничение $g(x) \geq 0$ называется ***активным*** (связывающим) в точке $\bar{x}$, если $g(\bar{x}) = 0$, и ***неактивным*** (несвязывающим), если $g(\bar{x}) > 0$.

Если существует возможность обнаружить неактивные в точке оптимума ограничения, то их можно сразу исключить из модели и таким образом уменьшить размерность задачи. Кун и Таккер построили необходимые и достаточные условия оптимальности для задач нелинейного программирования для дифференцируемых функций $0\, f(x), g_i(x), h_j(x)$.

Рассмотрим пример.

Пример 3.3. Решить задачу

$$f(x) = (x_1 - 1)^2 + x_2^2 \to \min,$$

$$-x_1 + \frac{x_2^2}{5} \geq 0.$$

Решение

1) Превратим ограничение-неравенство в равенство, отняв от левой части заведомо неотрицательную величину u^2:

$$-x_1 + \frac{x_2^2}{5} - u^2 = 0.$$

2) Построим функцию Лагранжа для полученной задачи с ограничением-равенством:

$$F(x_1, x_2, \mu, u) = (x_1 - 1)^2 + x_2^2 + \mu\left(-x_1 + \frac{x_2^2}{5} - u^2\right).$$

3) Найдем все частные производные и приравняем их к нулю.

Библиотека ГГТУ им. П.О.Сухого

$$\begin{cases} F'_{x_1} = 2(x_1 - 1) - \mu = 0, \\ F'_{x_2} = 2x_2 + \dfrac{2\mu x_2}{5} = 0, \\ F'_{\mu} = -x_1 + \dfrac{x_2^2}{5} - u^2 = 0, \\ F'_{u} = -2u\mu = 0. \end{cases} \quad \ldots\ldots \qquad \ldots\ldots\ldots\ldots(3.9)$$

Умножим последнее уравнение на $u/2$ и подставим вместо u^2 выражение из третьего уравнения:

$$-u^2\mu = 0 = \mu\left(x_1 - \frac{x_2^2}{5}\right) = 0\,.$$

Кроме того, добавим в систему ограничение $-x_1 + \dfrac{x_2^2}{5} \geq 0$.

Тогда система (3.9) преобразуется к виду:

$$\begin{cases} 2(x_1 - 1) - \mu = 0, \\ 2x_2\left(1 + \dfrac{\mu}{5}\right) = 0, \\ \mu\left(x_1 - \dfrac{x_2^2}{5}\right) = 0, \\ -x_1 + \dfrac{x_2^2}{5} \geq 0. \end{cases}$$

Полученная система носит название *задачи Куна-Таккера* (или *условий Куна-Таккера*).

Если $x_2 = 0$, то $x_1 = 0$, $\mu = -2$. Функция Лагранжа примет вид

$$F(x_1, x_2, \mu, u) = (x_1 - 1)^2 + x_2^2 - 2\left(-x_1 + \frac{x_2^2}{5} - u^2\right).$$

Матрица Гессе в точке (0;0) примет вид $G = \begin{pmatrix} 2 & 0 \\ 0 & 6/5 \end{pmatrix}$. Т.к. она положительно определена, то точка (0;0) является точкой минимума.

Пусть теперь $\mu = -5$. Тогда $x_1 = -3/2$, $\dfrac{x_2^2}{5} = -3/2 < 0$, следовательно, система в этом случае не имеет решений.

Библиотека ГГТУ им. П.О.Сухого

Если $\mu = 0$, то $x_1 = 1$, $x_2 = 0$. Однако эта точка не удовлетворяет ограничению $-x_1 + \frac{x_2^2}{5} \geq 0$.

Случай, когда $x_1 - \frac{x_2^2}{5} = 0$ приводит к требованию $x_2 = 0$ или $\mu = -5$, а эти варианты мы уже рассмотрели.

Таким образом, точкой локального минимума функции при заданном ограничении является точка (0;0), $f_{\min} = 1$.

Описанный в примере метод можно обобщить на случай нескольких ограничений-неравенств, а также на случай общей задачи нелинейного программирования (3.6)-(3.8), содержащей как ограничения-равенства, так и ограничения-неравенства.

фДля задачи (3.6)-(3.8) Функция Лагранжа примет вид:

$$F(x, \lambda, \mu) = f(x) + \sum_{j=1}^{k} \lambda_j h_j(x) + \sum_{l=1}^{m} \mu_l \left(g_l(x) - u_l^2 \right), \quad (3.10)$$

а условия Куна-Таккера

$$F'_{x_i} = \frac{\partial f(x)}{\partial x_i} + \lambda_j \frac{\partial h_j(x)}{\partial x_i} + \mu_j \frac{\partial g_j(x)}{\partial x_i} = 0, \quad (3.11)$$

$$h_j(x) = 0,\ j = 1,2,\ldots,k, \quad (3.12)$$

$$\mu_l g_l(x) = 0, \quad (3.13)$$

$$g_l(x) \geq 0,\ l = 1,2,\ldots,m. \quad (14.14)$$

Условия (3.12) носят название ***уравнений связи***, а (3.13) – ***условий дополняющей нежесткости***.

Теорема 3.1 (необходимое условие оптимума). Пусть x^* – оптимальное решение задачи нелинейного программирования (3.6)-(3.8). Тогда существует такая пара векторов $(\mu^*; \lambda^*)$, что $(x^*; \mu^*; \lambda^*)$ является решением задачи (3.11)-(3.14).

Теорема 3.2 (достаточное условие оптимума). Рассмотрим задачу нелинейного программирования (3.6)-(3.8). Пусть функция $f(x)$ выпукла, все функции $g_l(x)$ – вогнутые, а $h_j(x) = 0,\ j = 1,2,\ldots,k$ – линейные. Тогда если существует решение $(x^*; \mu^*; \lambda^*)$ задачи (3.11)-(3.12), то x^* является оптимальным решением задачи (3.6)-(3.8).

Библиотека БГТУ имени П.О. Сухого

Контрольные вопросы и задания

1. В чем суть метода множителей Лагранжа при решении задач условной оптимизации?

2. Сформулировать условия Куна-Таккера.

3. С помощью метода множителей Лагранжа решить следующие задачи:

1) $f(x) = x_1^2 + x_2^2 + x_3^2 \to \min$

$x_1 + 2x_2 + 3x_3 = 7,$

$2x_1 + 2x_2 + x_3 = 9/2.$

2) $f(x) = x_1^2 + (x_2 - 2)^2 \to \min$

$2x_1 + x_2 \ge 7,$

$x_1 + 2x_2 \ge 5.$

4. Применить метод множителей Лагранжа для нахождения точек глобального минимума и глобального максимума функции $f(x) = x_2 - x_1^2$ при ограничении $x_1^2 + x_2^2 = 1$.

5. Найти кратчайшее расстояние отточки (1;0) до параболы $y = 4x^2$

а) путем исключения переменной y;

б) с помощью метода множителей Лагранжа.

Объяснить, почему процедура б) в отличие от а) приводит к получению решения задачи.

6. Дана следующая задача:

$f(x) = (x-3)_1^2 + (x_2 + 4)^2 + e^{5x_3} \to \min$

$x_1 + x_2 + x_3 \le 1,$

$x_1, x_2, x_3 \ge 0.$

а) записать условия Куна-Таккера для этой задачи;

б) доказать, что выполнение условий Куна-Таккера является достаточным для существования оптимального решения данной задачи;

в) используя результаты пп. а) и б), доказать, что $x = (1;0;0)$ есть точка оптимума.

7. Дана следующая задача:

$f(x) = x_1^2 + 2x_2^2 \to \min$

$x_1^2 + x_2^2 \le 5$

$2x_1 - 2x_2 = 1.$

Библиотека ГГТУ им.П.О.Сухого

а) сформулировать эквивалентную задачу Куна-Таккера;

в) используя результат п. а), сделать вывод о следующих решениях задачи:

1) $x^{(1)} = (0;0)$;

2) $x^{(2)} = (1;1/2)$;

3) $x^{(3)} = (1/3;-1/6)$.

Проверить сделанные выводы, обосновать.

4. Элементы линейного программирования

Задачами линейного программирования называются оптимизационные задачи с *линейной* целевой функцией и ограничениями, заданными в виде *линейных равенств* или *неравенств.* Линейное программирование представляет собой наиболее часто используемый метод оптимизации.

Пример 4.1. Пекарня выпекает плюшки и ватрушки. На производство плюшки расходуется в два раза больше дрожжей и в два раза меньше сахара, чем на производство одной ватрушки. Расход муки одинаков. Предприниматель закупил муку в объеме, которого хватило бы на выпечку 100 кг плюшек, сахар – на 80 кг плюшек, а дрожжи – на 90 кг ватрушек. Плюшки продаются по цене 20 тыс. руб./кг, ватрушки – 30 тыс. руб./кг. Требуется спланировать производство плюшек-ватрушек так, чтобы получить максимальную прибыль.

Решение

Обозначим через x_1 и x_2 объем производства плюшек и ватрушек соответственно (в килограммах). Тогда функция прибыли примет вид:

$$f = 20x_1 + 30x_2,$$

а ограничения на сырье запишутся в виде:

$$x_1 + x_2 \le 100,$$

$$2x_1 + x_2 \le 180,$$

$$x_1 + 2x_2 \le 160.$$

Кроме того, очевидно, $x_1, x_2 \ge 0$.

Мы получили *математическую модель задачи*:

$$f = 20x_1 + 30x_2 \to \max,$$

Библиотека ГГТУ им. П.О. Сухого

$$\begin{cases} x_1 + x_2 \le 100, \\ 2x_1 + x_2 \le 180, \\ x_1 + 2x_2 \le 160, \\ x_1, x_2 \ge 0 \end{cases}$$

4.1. Задача линейного программирования в стандартной форме

Задача линейного программирования ***в стандартной форме*** с m ограничениями и n переменными имеет следующий вид:

Максимизировать (или минимизировать) целевую функцию

$$f = c_1x_1 + c_2x_2 + \dots + c_nx_n \to \max \qquad (4.1)$$

при наличии ограничений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \quad \dots \quad \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m, \\ x_i \ge 0, i = 1,2,\dots,n, \\ b_j \ge 0, j = 1,2,\dots,m. \end{cases} \qquad (4.2)$$

Задачу (4.1)-(4.2) можно записать в более компактном матричном виде следующим образом:

$$f = cx \to \max, \qquad (4.3)$$

$$Ax = b, \;\; x \ge 0, b \ge 0. \qquad (4.4)$$

Матрицу A называют ***матрицей коэффициентов***, x – ***вектором переменных***, b – ***вектором ресурсов***, c – ***вектором оценок*** задачи линейного программирования.

Основные понятия и определения

– ***допустимое решение*** – неотрицательный вектор x, для которого выполняются ограничения $Ax = b$;

– ***допустимая область*** состоит из всех допустимых решений (обозначается буквой S);

– ***оптимальное решение*** – такой допустимый вектор x^0, для которого значение целевой функции $f_0 = cx^0$ дольше, чем для любого другого допустимого решения; в этом случае значение cx^0 называется ***оптимальным значением*** задачи линейного программирования.

При решении задачи ЛП симплекс-методом требуется, чтобы задача была представлена в *стандартной форме*, однако так бывает

Библиотека ГГТУ им.П.О.Сухого

не всегда. Таким образом, первый этап решения задачи линейного программирования состоит в приведении ее к стандартному виду путем введения дополнительных ***избыточных*** или ***остаточных*** переменных для преобразования неравенств в равенства. Кроме того, в некоторых задачах необходимо рассмотреть переменные, принимающие как положительные, так и отрицательные значения. Но в стандартной задаче ЛП все переменные полагаются неотрицательными, следовательно, неограниченные переменные необходимо заменить разностью двух неотрицательных. Например, если x_1 – неограниченная переменная, то следует ввести замену $x_1 = s_1 - s_2$.

Пример 3.2. Преобразовать к стандартному виду задачу

$$f = x_1 - 2x_2 + 3x_3 \to \max$$

$$\begin{cases} x_1 + x_2 + x_3 \le 7, \\ x_1 - x_2 + x_3 \le -2, \\ x_1, x_2 \ge 0 \end{cases}$$

Решение

1. Т.к. переменная x_3 не ограничена по знаку, то заменим x_3 на $x_4 - x_5$, где $x_4, x_5 \ge 0$. Добавим в первое уравнение остаточную переменную x_6:

$$x_1 + x_2 + x_4 - x_5 + x_6 = 7.$$

2. Умножим обе части второго неравенства на (-1), т.к. все b_i должны быть неотрицательны:

$$-x_1 + x_2 - x_3 \ge 2.$$

3. Преобразуем последнее неравенство в равенство с помощью введения избыточной переменной x_7:

$$-x_1 + x_2 - (x_4 - x_5) - x_7 = 2.$$

Запишем задачу в стандартном виде:

$$f = x_1 - 2x_2 + 3x_4 - 3x_5 \to \max$$

$$\begin{cases} x_1 + x_2 + x_4 - x_5 + x_6 = 7, \\ -x_1 + x_2 - x_4 + x_5 - x_7 = 2, \\ x_1, x_2, x_4, x_5, x_6, x_7 \ge 0 \end{cases}$$

Библиотека БГУИР

4.2. Основы симплекс-метода

Рассмотрим задачу линейного программирования (4.1)-(4.2) в стандартном виде.

Как правило, число ограничений m меньше числа неизвестных n, поэтому множество допустимых решений бесконечно, а значит, выбор наилучшего допустимого решения, максимизирующего функцию f, нетривиален. Назовем m переменных, входящих с единичными коэффициентами только в одно уравнение системы и с нулевыми – во все остальные, ***базисными переменными***, а остальные $n-m$ переменных – ***небазисными*** (зависимыми).

Базисным решением системы в каноническом виде называется решение, полученное при нулевых значениях небазисных переменных. Базисное решение называется ***допустимым***, если значения входящих в него базисных переменных неотрицательны.

Алгоритм симплекс-метода для задачи (4.1)-(4.2), записанной в стандартном виде, включает следующие шаги:

1. Выбор начального допустимого базисного решения.
2. Переход к другому допустимому базисному решению с наилучшим значением целевой функции. На этом шаге исключаются все допустимые базисные решения, которые хуже текущего решения, что как раз и делает симплекс-метод гораздо более эффективным, чем прямой перебор всех допустимых базисных решений.
3. Продолжение поиска допустимых базисных решений, улучшающих значение целевой функции. Если некоторое решение нельзя улучшить, то оно и является оптимальным.

Рассмотрим подробнее ***шаг 2*** этого алгоритма.

Пусть допустимое базисное решение задано в виде

$$x_1=\overline{b}_1, x_2=\overline{b}_2, \ldots, x_m=\overline{b}_m, \ldots, x_{m+1}=x_{m+2}=\cdots=x_n.$$

Назовем вектор базисных переменных $x_B=(x_1,x_2,\ldots,x_m)$ ***базисом***. Обозначим вектор коэффициентов при базисных переменных в целевой функции через $c_B=(c_1,c_2,\ldots,c_m)$. Т.к. небазисные переменные равны нулю, то значение целевой функции, соответствующее начальному допустимому базисному решению, равно

$$f=c_Bx_B=c_1\overline{b}_1+c_2\overline{b}_2+\cdots+c_m\overline{b}_m.$$

Симплекс метод позволяет проверить существование допустимого базисного решения с большим значением целевой функции f. Для этого проверяют полученное решение на

Библиотека ГГТУ им. П.О. Сухого

оптимальность и в случае, если оно не оптимально, переходят к смежному (отличающемуся от рассматриваемого только одной базисной переменной) допустимому базисному решению с большим значением функции f.

Для получения смежного допустимого базисного решения симплекс-метод превращает одну из базисных переменных в небазисную и вводит одну из небазисных переменных в базис. Для выбора вводимой в базис переменной следует присвоить небазисной переменной значение, равное единице, и вычислить изменение целевой функции.

Рассмотрим для определенности небазисную переменную x_S. Присвоим ей значение, равное единице, и вычислим изменение целевой функции. Т.к. мы рассматриваем смежное допустимое базисное решение, то все остальные небазисные переменные по-прежнему равны нулю. Подставив в систему (4.2) $x_S = 1$, получим:

$$x_1 = \overline{b}_1 - \overline{a}_{1S},\ x_2 = \overline{b}_2 - \overline{a}_{2S}, \ldots, x_m = \overline{b}_m - \overline{a}_{mS},\ x_{m+1} = 0, \ldots,$$
$$x_S = 1, x_{S+1} = \ldots = x_n = 0 \qquad (4.5)$$

Тогда новое значение целевой функции равно $f_H = \sum_{i=1}^{m} c_i(\overline{b}_i - \overline{a}_{iS}) + c_S$. Обозначим через $\overline{c}_S$ приращение целевой функции. Тогда

$$\overline{c}_S = f_{новое} - f_{старое} = f_H - f = \sum_{i=1}^{m} c_i(\overline{b}_i - \overline{a}_{iS}) + c_S - \sum_{i=1}^{m} c_i \overline{b}_i =$$
$$= c_S - \sum_{i=1}^{m} c_i \overline{a}_{iS}. \qquad (4.6)$$

Величина $\overline{c}_S$ называется ***относительной оценкой*** небазисной переменной x_S. Если $\overline{c}_S > 0$, то, вводя переменную x_S в базис, можно добиться увеличения целевой функции. Уравнение (4.6), определяющее относительную оценку, называется ***правилом скалярного произведения***.

До каких пор можно улучшать значение целевой функции? Ответ на этот вопрос дает

Условие оптимальности. Если относительные оценки небазисных переменных допустимого базисного решения задачи (4.1)-(4.2) отрицательны или равны нулю, то решение оптимально.

Пример 4.3. Решим симплекс-методом задачу 4.1.

Библиотека ГГТУ им. П.О.Сухого

Решение

В примере 4.1 мы составили математическую модель задачи:

$$f = 20x_1 + 30x_2 \to \max,$$

$$\begin{cases} x_1 + x_2 \le 100, \\ 2x_1 + x_2 \le 180, \\ x_1 + 2x_2 \le 160, \\ x_1, x_2 \ge 0 \end{cases}$$

Преобразуем задачу к стандартному виду:

$$f = 20x_1 + 30x_2 \to \max,$$

$$\begin{cases} x_1 + x_2 + x_3 = 100, \\ 2x_1 + x_2 + x_4 = 180, \\ x_1 + 2x_2 + x_5 = 160, \\ x_1, x_2, x_3, x_4, x_5 \ge 0 \end{cases}$$

В качестве базисных переменных выберем x_3, x_4, x_5 (каждая из них входит в одно из уравнений с коэффициентом 1 и не входит в остальные). Последовательные итерации симплекс-метода удобно представлять в виде ***таблицы***.

Элементы таблицы представляют собой коэффициенты задачи. Столбец c_B состоит из коэффициентов целевой функции при базисных переменных. Из таблицы видно, что начальное допустимое базисное решение имеет вид:

$$x_1 = 0, x_2 = 0, x_3 = 100, x_4 = 180, x_5 = 160.$$

Значение целевой функции в этом случае равно нулю:

$$f = c_B x_B = (0,0,0)(100,180,160) = 0.$$

Для проверки оптимальности найденного допустимого решения вычислим относительные оценки переменных, пользуясь правилом скалярного произведения (4.6).

c_B	Базис	c_j					Постоянные
		20	30	0	0	0	
		x_1	x_2	x_3	x_4	x_5	
0	x_3	1	1	1	0	0	100
0	x_4	2	1	0	1	0	180
0	x_5	1	2	0	0	1	160
c-строка		20	30	0	0	0	$f = 0$

Библиотека ГГТУ им. П.О. Сухого

Небазисные переменные x_1 и x_2 имеют положительные относительные оценки, что говорит о том, что построенное решение не оптимально. Небазисная переменная x_2 имеет самую большую относительную оценку, поэтому ее следует ввести в базис. Для выделения переменной, выводимой из базиса, применяется ***принцип минимального отношения***: вычисляем отношения $\dfrac{\bar{b}_i}{\bar{a}_{i2}}$ для всех ограничений с положительным коэффициентом при x_2, и выбираем наибольшее из них. У нас это будут числа $\dfrac{100}{1}, \dfrac{180}{1}, \dfrac{160}{2}$, наименьшее из них – 80 (соответствует третьей строке), следовательно, назовем третью строку нашей таблицы ***ведущей строкой***, а коэффициент при x_2 в ведущей строке, – ***ведущим элементом*** (в данной задаче он равен двум). Выводим из базиса переменную x_5 и приходим к новому базису x_3, x_4, x_2.

Строим новую таблицу, снова вычисляем вектор относительных оценок и, в случае хотя бы одной положительной, с помощью принципа минимального отношения определяем ведущую строку.

c_B	Базис	c_j					Постоянные
		20	30	0	0	0	
		x_1	x_2	x_3	x_4	x_5	
0	x_3	½	0	1	0	-1/2	20
0	x_4	3/2	0	0	1	-1/2	100
30	x_2	½	1	0	0	1/2	80
c-строка		5	0	0	0	-15	$f = 2400$

Значение функции $f = c_B x_B = (0,0,30)(20,100,80) = 2400$.

В первом столбце относительная оценка положительна, следовательно, значение целевой функции еще можно улучшить. Найдем минимальное отношение:

$$\min\left\{\frac{80}{1/2}, \frac{100}{3/2}, \frac{20}{1/2}\right\} = \left\{160, 66\frac{2}{3}, 40\right\} = 40.$$

Таким образом, первая строка является ведущей, переменная x_3 выводится из базиса, а переменная x_1 в базис вводится.

Строим новую симплекс-таблицу.

		c_j	

Библиотека ГГТУ им. П.О.Сухого

c_B	Базис	20 x_1	30 x_2	0 x_3	0 x_4	0 x_5	Постоянные
20	x_1	1	0	2	0	-1	40
0	x_4	0	0	-3	1	1	40
30	x_2	0	1	-1	0	1	60
c-строка		0	0	-10	0	-10	$f = 2600$

В последней таблице все относительные оценки неположительны, следовательно, построено оптимальное решение. Оно имеет вид $x_1 = 40$, $x_2 = 60$, оптимальное значение функции $f = 2600$. Следовательно, если выпекать по 40 кг плюшек и 60 кг ватрушек, то мы получим максимальную прибыль, равную 2600 тыс. руб.

Замечание 4.1. Задачу минимизации можно превратить в задачу максимизации путем умножения целевой функции на (-1). К полученной задаче «на максимум» можно применить описанный симплекс-метод.

Замечание 4.2. При использовании симплекс-метода может возникнуть ситуация, когда все коэффициенты в ограничениях при вводимо в базис переменной неположительны (применение принципа минимального отношения невозможно). Это говорит о том, что вводимая в базис небазисная переменная может неограниченно возрастать, не нарушая множество ограничений. Отсюда следует, что у задачи нет конечного оптимального решения и говорят, что «она имеет неограниченный оптимум».

Контрольные вопросы и задания

1. Записать задачу линейного программирования в общем виде.
2. В чем различие между симплекс-методом и методом полного перебора?
3. Какие переменные называются базисными?
4. Что такое остаточные (избыточные) переменные и для чего они служат?
5. В чем заключается правило скалярного произведения и какова его роль в симплекс методе?
6. В чем заключается правило минимального отношения и какова его роль в симплекс методе?
7. Опишите алгоритм симплекс-метода.

Библиотека БГУИР

8. Преобразовать следующую задачу ЛП к стандартной форме:
минимизировать $f(x) = -3x_1 + 4x_2 - 2x_3 + 5x_4$
при ограничениях $4x_1 - x_2 + 2x_3 - x_4 = -2$,

$$-2x_1 + 3x_2 - x_3 + 2x_4 \geq 2,$$

$$x_1 + x_2 + 3x_3 - x_4 \leq 14,$$

$$x_1, x_2 \geq 0, \ x_3 \leq 0, \ x_4 \text{ не ограничена по знаку.}$$

9. Имеется три магазина, товар в которые доставляется с двух складов. Ежедневно с первого склада вывозится 12 условных единиц товара, со второй – 15. Первый магазин принимает 8 единиц товара, второй – 9, третий – 10. Стоимость перевозки единицы товара со складов в магазины приведена в таблице:

Склад	Магазин		
	№1	№2	№3
Первый	0,8	1,1	0,9
Второй	1,0	0,7	1,2

10. Стройотряд располагает следующими ресурсами: 100м^2 стекла, 300 кг металла и 160 человеко-часов рабочего времени. Отряд изготавливает изделия типа А по цене 500 тыс. руб. за штуку и изделия типа Б по 600 тыс. руб. за штуку. Расход ресурсов приведен в таблице. Требуется спланировать объем выпуска продукции типов А и Б так, чтобы получить наибольшую прибыль.

Изделия	Расход ресурсов		
	Металл, кг	Стекло, м^2	Человеко-часы
А	4	2	2
Б	5	1	3

11. Для производства трех различных картофельных продуктов – картофельных кубиков, картофельных долек и чипсов – предприятие может закупать картофель у двух поставщиков – А и В. Объемы продуктов, получающиеся из одной тонны картофеля каждого поставщика, приведены в таблице. Там же приведены ограничения на объем выпуска продукции каждого вида.

Продукт	Поставщик		Ограничения на объем выпускаемой продукции
	А	В	
Кубики	0,2	0,1	1,2
Дольки	0,3	0,3	1,8
Чипсы	0,2	0,3	2,4

Библиотека ГГТУ им.П.О.Сухого

Прибыли предприятия при закупке 1 т сырья у производителя, соответственно, А и В составляют 5 и 6 у.е. Какое количество картофеля нужно закупать у каждого из поставщиков, чтобы обеспечить предприятию максимально возможную прибыль?

5. Методы поиска для решения задач многомерной условной оптимизации

В технических приложениях часто приходится сталкиваться с задачами, в которых целевая функция либо функции-ограничения разрывны или недифференцируемы. В таких случаях применимы только методы, использующие лишь значения самой функции и не использующие производных. Рассмотрим некоторые методы, которые основаны на применении идей, процедур и алгоритмов методов прямого поиска, рассмотренных в разделе о безусловной оптимизации.

Общая задача условной оптимизации содержит ограничения в виде равенств или неравенств, а также верхние и нижние границы значений переменных.

$$f(x_1, x_2, \ldots, x_n) \to \min, \tag{5.1}$$

$$g_l(x_1, x_2, \ldots, x_n) \ge 0,\ l = 1,2,\ldots m,$$

$$h_j(x_1, x_2, \ldots, x_n) = 0,\ j = 1,2,\ldots k. \tag{5.2}$$

$$x_i^{(l)} \le x_i \le x_i^{(u)}.$$

Для произвольной точки x^0 можно так видоизменить схему метода безусловной оптимизации, что точки, получаемые на последующих итерациях, оказываются допустимыми. Например, если пробная точка v не удовлетворяет какому-либо ограничению-неравенству, то можно заменить ее другой, изменяя величину шага так, чтобы это неравенство оказалось выполненным. Ограничения же в виде равенств доставляют больше трудностей, т.к. даже если x^0 и v удовлетворяют ограничениям-равенствам, то точки на прямой, через них проходящей, не обязательно им удовлетворяют (за исключением случая линейных ограничений). По существу все методы прямого поиска требуют, чтобы ограничения задачи были заданы только в виде неравенств, следовательно, ограничения в виде равенств должны быть явно или неявно исключены перед решением задачи.

Простейший способ исключения ограничений в виде равенств заключается в решении его относительно одной из переменных с последующим исключением этой переменной из задачи путем

Библиотека ГГТУ им. П.О. Сухого

подстановки. При этом верхние и нижние границы значений переменных сохраняются в задаче в виде ограничений-неравенств.

Пример 5.1. Минимизировать функцию $f(x) = x_1^2 + 4x_2^2 + x_3^2$ при ограничениях

$$x_1 + x_2^2 - 2 = 0,$$
$$-1 \le x_1 \le 1,$$
$$0 \le x_2 \le 2,$$
$$0 \le x_3 \le 2.$$

Решение

Выразим $x_1 = 2 - x_2^2$, подставим в целевую функцию и остальные ограничения. Задача примет вид:

$$f(x_2, x_3) = (2 - x_2)^2 + 4x_2^2 + x_3^2 \to \min,$$
$$1 \le x_2 \le \sqrt{3},$$
$$0 \le x_3 \le 2.$$

Следующий этап решения задачи состоит в выборе начальной допустимой точки. Простейший способ – метод случайного поиска с последующей проверкой на допустимость путем подстановки в ограничения. По заданным верхним и нижним границам каждой переменной $x_i^{(u)}, x_i^{(l)}$ подсчитываются

$$x_i = x_i^{(l)} + r_i (x_i^{(u)} - x_i^{(l)}), \qquad (5.3)$$

$i = 1,2,\ldots,n$, где r_i – случайные числа из интервала (0;1). Процесс повторяется n раз (для каждой координаты) и продолжается, пока не будет найдена точка, удовлетворяющая всем ограничениям (он может оказаться очень долгим, особенно если там присутствуют ограничения-равенства – именно поэтому их надо исключить).

5.1 Методы прямого поиска

В принципе, любой из методов прямого поиска может быть использован для решения задач с ограничениями в виде неравенств.

Рассмотрим применение метода сопряженных направлений. Пусть задана начальная точка x^0 и множество направлений d_i. Метод включает процедуру одновременного поиска вдоль этих направлений, в результате чего определяется сопряженное направление и улучшается значение целевой функции. Но такой поиск может привести к тому, что решение соответствующей задачи одномерной оптимизации окажется недопустимым. Этого можно избежать, если ограничить поиск вдоль выбранного направления точкой пересечения

Библиотека ГГТУ им. П.О.Сухого

луча с границей области. Однако и в этом случае решение задачи не гарантируется: вычисления могут свестись к поиску вдоль фиксированных направлений и либо выводить из допустимой области, либо давать точку с худшим значением целевой функции. Тогда требуется изменить направления поиска, используя информацию о градиентах функций-ограничений.

Аналогичный прием используется в методе Хука-Дживса, однако и тут следует внести некоторые изменения в алгоритм. Если в результате поиска по образцу получена недопустимая базовая точка, то можно уменьшать шаг, например, вдвое, до тех пор, пока не будет найдена допустимая точка. Аналогично проводится «исследующий поиск». Однако и здесь, как и в методе сопряженных направлений, может возникнуть ситуация, когда уменьшение шага не дает улучшения. Тогда следует поменять направление и производить поиск вдоль прямой, параллельной ограничивающей поверхности.

Вывод: несложный прием уменьшения шага при поиске очередной точки позволяет получить пробную допустимую точку. Начинать надо в любом случае с допустимой точки. Однако, если не корректировать множество направлений поиска, то результат оказывается неудовлетворительным.

5.2. Метод комплексов

Этот метод представляет собой метод поиска по симплексу, адаптированный на случай оптимизации с ограничениями.

Бокс предложил строить множество пробных точек P случайным образом, генерируя их по формуле (5.3). Заметим, что для этого необходимо знать границы значений всех переменных. Они определяются либо естественным образом из постановки задачи, либо с помощью оценки интервалов возможных значений переменных. Их следует включить в ограничения и сделать как можно короче для большей эффективности.

Каждая построенная в соответствии с этой формулой точка проверяется на допустимость, и в случае, если какое-то из ограничений нарушается, сдвигается к центру тяжести построенных точек до тех пор, пока не получится допустимая точка. Формула (5.3) при этом примет вид:

$$x^m = \bar{x} + \alpha(\bar{x} - x^R), \tag{5.4}$$

где x^R – исключаемая точка, $\bar{x}$ – центр тяжести остальных точек.

Библиотека БГУИР

Когда найдена новая точка и вычислены значения целевой функции, возможны следующие случаи:

1. Новая точка допустима, значение целевой функции в ней не совпадает с максимальным значением на всей совокупности точек. В этом случае снова выполняем процедуру отражения.

2. Новая точка допустимая и значение в ней совпадает с максимальным значением целевой функции. Тогда вместо отражения передвигаем эту точку на половину расстояния до ранее найденного центра тяжести.

3. Новая точка недопустимая. Тогда уменьшаем в два раза расстояние до вычисленного ранее центра тяжести. Процедура поиска продолжается до тех пор, пока многогранник не окажется стянутым в центр тяжести в пределах заданной точности, либо пока разница значений целевой функции не станет достаточно малой.

Пример 5.2. Требуется спроектировать прямоугольную конструкцию в виде параллелепипеда с открытой передней стенкой. Конструкция должна иметь объем 16 000 фут^3, периметр основания не должен превышать 220 футов. Глубина не должна быть больше 60 футов, ширина – 70 футов. Кроме того, ширина не должна превышать утроенной глубины, а высота – 2/3 ширины. Стоимость гофрированного материала, из которого изготовляются крыша и три стенки конструкции составляет 30$/фут^2. Определить размеры конструкции так, чтобы минимизировать стоимость материалов.

Решение

Введем следующие обозначения: x_1 – глубина, x_2 – ширина, x_3 – высота.

Тогда стоимость крыши составит $30x_1x_2$, стоимость задней стенки – $30x_2x_3$, стоимость боковых стенок – $2 \cdot 30x_1x_3$.

Таким образом, целевая функция имеет вид:

$$f(x) = 30x_1x_2 + 30x_2x_3 + 60x_1x_3.$$

Ограничение на объем:

$$x_1x_2x_3 = 16000,$$

Ограничение на периметр:

$$2(x_1 + x_2) \le 220,$$

Условия на линейный размер:

$$x_2 \le 3x_1,\ x_3 \le \frac{2}{3}x_2.$$

Границы изменения переменных:

$$0 \le x_1 \le 60,$$
$$0 \le x_2 \le 70.$$

Задача в стандартной форме примет вид:

$$f(x) = 30x_1x_2 + 30x_2x_3 + 60x_1x_3 \to \min,$$
$$h_1(x) = x_1x_2x_3 - 16000 = 0,$$
$$g_1(x) = 100 - x_1 - x_2 \ge 0,$$
$$g_2(x) = 3x_1 - x_2 \ge 0,$$
$$g_3(x) = \frac{2}{3}x_2 - x_3 \ge 0,$$
$$0 \le x_1 \le 60,$$
$$0 \le x_2 \le 70,$$
$$0 \le x_3$$

Исключим ограничение-равенство.

$$x_3 = \frac{16000}{x_1 x_2},$$

Тогда $f(x) = 30x_1x_2 + \dfrac{48 \cdot 10^4}{x_1} + \dfrac{96 \cdot 10^4}{x_2} \to \min$

$$g_2(x) = 3x_1 - x_2 \ge 0,$$
$$g_3(x) = \frac{2}{3}x_2 - \frac{16 \cdot 10^3}{x_1x_2} \ge 0,$$
$$0 \le x_1 \le 60,$$
$$0 \le x_2 \le 70.$$

Возьмем начальную точку $x^{(1)} = (50;50)$.

Построим еще три точки:

$x_1^{(k)} = 0 + 60r_1$, $x_2^{(k)} = 0 + 70r_1$, где r_1, r_2 – случайные числа, равномерно распределенные на (0;1).

Пусть $r_1 = 0{,}5$, $r_2 = 0{,}9$. Тогда $x^{(2)} = (30;63)$ – вторая допустимая точка.

Пусть $r_1 = 0{,}1$, $r_2 = 0{,}15$. Тогда $x^{(3)} = (6;10{,}5)$ – недопустимая точка. Т.к. $\bar{x} = \frac{1}{2}\left(x^{(1)} + x^{(2)}\right) = (40;56{,}6)$, то $\bar{x} = \frac{1}{2}\left(x^{(3)} + \bar{x}\right) = (23;33{,}5)$ – третья допустимая точка.

Библиотека ГГТУ им. П.О. Сухого

Пусть $r_1 = 0{,}9$, $r_2 = 0{,}33$. Тогда $x^{(4)} = (54;23{,}1)$ – четвертая допустимая точка.

Вычислим значения целевой функции:

$f\left(x^{(1)}\right) = 103{,}8 \cdot 10^3$, $f\left(x^{(2)}\right) = 87{,}94 \cdot 10^3$, $f\left(x^{(3)}\right) = 73{,}64 \cdot 10^3$, $f\left(x^{(4)}\right) = 8787 \cdot 10^3$.

Новая возможная точка (при величине шага α=1,3):

$$x^{(5)} = \bar{x} + 1{,}3(\bar{x} - x^{(1)}) = (35{,}67;\ 39{,}87) + 1{,}3\big((35{,}67;\ 39{,}87) - (50;50)\big) = \\ = (17{,}03;26;69).$$

Но в этой точке не выполнено ограничение g_3, следовательно, сдвинем точку в направлении центра тяжести.

$$x^{(5)} = \frac{1}{2}(x^{(5)} + \bar{x}) = (26{,}35;\ 33{,}28),\ f\left(x^{(5)}\right) = 73{,}37 \cdot 10^3.$$

Имеем точки $x^{(2)}, x^{(3)}, x^{(4)}, x^{(5)}$, «худшая» из них – $x^{(2)}$.

Центр тяжести остальных вершин: $\bar{x} = (34{,}45;29{,}96)$.

$x^{(6)} = (34{,}81;\ 27{,}28)$ (после двух сдвигов к центру тяжести), $f\left(x^{(6)}\right) = 77{,}47 \cdot 10^3$.

Последующие итерации приводят к оптимальному решению $x^* = (20;\ 40)$

Контрольные вопросы и задания

1. Записать в общем виде задача условной оптимизации функции n переменных.
2. Почему необходимо исключить ограничения в виде равенств при использовании методов прямого поиска?
3. Какова причина использования более чем $N+1$ точки в алгоритме комплексов?
4. Возможен ли пропуск минимального решения при использовании случайного поиска с уменьшением интервала? Если да, то по какой причине и как этого можно избежать?
5. Дана задача

$$\text{минимизировать } f(x) = x_1^2 + x_2^2 + x_3^2$$

при ограничениях

$$g_1(x) = 1 - x_2^{-1}x_3 \ge 0,$$
$$g_2(x) = x_1 - x_3 \ge 0,$$
$$h_1(x) = x_1 - x_2^2 + x_2x_3 - 4 = 0,$$
$$0 \le x_1 \le 5, \quad 0 \le x_2 \le 3, \quad 0 \le x_3 \le 3.$$

Библиотека ГГТУ им.П.О.Сухого

Какие преобразования необходимо выполнить, чтобы можно было использовать метод комплексов? Приведите окончательную форму задачи.

6. При решении задачи

минимизировать $f(x) = 3(x_2 - 4)^2 - 2x_1$

при ограничениях $g_1(x) = 10 - x_1^2 - x_2^2 \ge 0$,

$$g_2(x) = 9 - x_1^2 - (x_2 - 4)^2 \ge 0,$$

$$0 \le x_1 \le 4, \quad 0 \le x_2 \le 4$$

методом комплексов получены следующие точки:

	x_1	x_2	$f(x)$
1	2,0	2,0	8
2	0,5	3,0	2
3	0,6	1,1	24,03
4	2,3	2,1	6,23

Полагая $\alpha = 1{,}3$, выполните две итерации метода комплексов.

7. Дана задача

минимизировать $f(x) = -x_1 x_2^2 x_3 / 81$

при ограничениях $h_1(x) = x_1^2 + x_2^2 + x_3 - 13 = 0$,

$$h_2(x) = x_2^2 x_3^{-1} - 1 = 0,$$

$$x_i \ge 0.$$

Переписать задачу в форме, пригодной для использования метода комплексов. Оценить верхние и нижние границы значений независимой переменной при использовании метода комплексов. Найти начальную допустимую точку.

6. Методы штрафных функций

Рассмотрим общую задачу условной оптимизации, содержащую ограничения в виде равенств и неравенств, а также верхние и нижние границы значений переменных.

$$f(x_1, x_2, \ldots, x_n) \to \min, \tag{6.1}$$

$$g_l(x_1, x_2, \ldots, x_n) \ge 0, \; l = 1,2,\ldots m,$$

$$h_j(x_1, x_2, \ldots, x_n) = 0, \; j = 1,2,\ldots k. \tag{6.2}$$

$$x_i^{(l)} \le x_i \le x_i^{(u)}$$

Как мы знаем, решение задач условной оптимизации гораздо более трудоемко по сравнению с задачами безусловной оптимизации,

Библиотека ГГТУ им. П.О.Сухого

т.к. ограничения в виде равенств или неравенств требуют проверки их выполнения на каждом шаге. Поэтому одним из способов решения задач нелинейного программирования является сведение их к последовательности задач безусловной оптимизации. К этому направлению относится и метод штрафных функций.

Суть метода: предположим, что для вектора x^*, являющегося решением задачи (6.1)-(6.2), известно некоторое начальное приближение $x^{(0)}$, возможно недопустимое. С помощью специальных алгоритмов в пространстве R^n строится конечная последовательность точек $x^{(1)}, x^{(2)}, \ldots, x^{(T)}$, начинающейся с точки $x^{(0)}$ и заканчивающейся точкой $x^{(T)}$, дающей наилучшее приближение к x^* среди всех точек рассматриваемой последовательности. В качестве точек последовательности $x^{(1)}, x^{(2)}, \ldots, x^{(T)}$ берутся стационарные точки так называемой *штрафной функции* – целевой функции вспомогательной задачи безусловной минимизации. Конкретные методы, основанные на указанной общей схеме, определяются видом штрафной функции, а также правилами, по которым производится пересчет штрафных параметров по окончании очередного цикла безусловной оптимизации. Методы штрафных функций классифицируются в соответствии со способами учета ограничений-неравенств, поскольку ограничения-равенства учитываются во всех методах более или менее одинаково.

Штрафная функция определяется выражением

$$P(x,R) = f(x) + \Omega(R, g(x), h(x)), \qquad (6.3)$$

где R – набор штрафных параметров, а так называемый штраф Ω является функцией R и функций, задающей ограничения.

Штраф Ω определяется так, чтобы допустимые точки имели преимущество перед недопустимыми в отношении безусловной минимизации штрафной функции.

Основные типы штрафов

1. Квадратичный штраф (используется для учета ограничений-равенств)

$$\Omega = R\{h(x)\}^2$$

При минимизации этот штраф препятствует отклонению величины $h(x)$ от нуля (как в сторону положительных, так и в сторону отрицательных значений). Легко видеть, что при увеличении

R стационарная точка соответствующей штрафной функции $P(x,R)$ приближается к x^*, поскольку в пределе $h\left(x^{(T)}\right)=0$.

Штрафы, используемые для ограничений-неравенств

2. Бесконечный барьер

Соответствующее выражение принимает бесконечно большие значения в недопустимых точках и нулевое значение в допустимых. Строго говоря, машинная реализация бесконечных штрафов невозможна, однако, можно в качестве бесконечности использовать большое положительное число, например,

$$\Omega = 10^{20}\sum\left|g_j(x)\right|$$

3. Логарифмический штраф

$$\Omega = -R\ln\left|g(x)\right|.$$

Этот штраф положителен при всех x, таких что $0 < g(x) < 1$, и отрицателен при $g(x) > 1$. В данном случае как бы вводится искусственная дискриминация точек допустимой области: внутренним точкам отдается предпочтение перед всеми остальными. Итерационный процесс начинается из допустимой начальной точки при положительном начальном значении R (например, $R=10$ или $R=100$). После решения каждой подзадачи безусловной минимизации параметр R уменьшается и в пределе стремится к нулю.

4. Штраф, заданный обратной функцией

$$\Omega = R/g(x)$$

Не имеет отрицательных значений в допустимой области. Этот штраф, как и предыдущий, является барьером. В этом случае также возможны трудности, связанные с появлением недопустимых точек.

5. Штраф типа квадрата срезки

$$\Omega = R < g(x) >^2,$$

где $< g(x) >= \begin{cases} g(x), \; если \; g(x) \le 0, \\ 0, \; если \; g(x) > 0. \end{cases}$

Этот штраф внешний и стационарные точки функции $P(x,R)$ действительно могут оказаться недопустимыми. Однако, в этом случае они не создают дополнительных сложностей по сравнению с допустимыми, т.к. различие состоит только в том, что в допустимых и граничных точках штраф равен нулю.

Библиотека БГТУ им. П.О. Сухого

Этот штраф удобен тем, что функция $P(x,R)$ определена и непрерывна всюду. Вычисления производят с положительными значениями R и после очередной итерации увеличивается.

Пример 6.1 Минимизировать функцию

$$f(x) = (x_1 - 4)^2 + (x_2 - 4)^4$$

при ограничении $h(x) = x_1 + x_2 - 5 = 0$.

Решение

Используя квадратичный, штраф введем штрафную функцию:

$$P(x,R) = (x_1 - 4)^2 + (x_2 - 4)^4 + R(x_1 + x_2 - 5)^2$$

Найдем ее стационарную точку:

$$\frac{\partial P}{\partial x_1} = 2(x_1 - 4) + 2R(x_1 + x_2 - 5) = 0$$

$$\frac{\partial P}{\partial x_2} = 2(x_2 - 4) + 2R(x_1 + x_2 - 5) = 0$$

Отсюда $x_1 = x_2 = \dfrac{4 + 5R}{1 + 2R}$.

Переходя к пределу при $R \to \infty$, получаем $\lim\limits_{R \to \infty} \dfrac{4 + 5R}{1 + 2R} = 2{,}5$.

Таким образом, метод сходится к точке $x^* = (2{,}5;2{,}5)$, $f_{\min} = 4{,}5$.

***Замечание* 6.1.** На практике точное значение получить удается крайне редко, обычно для поиска минимума приходится использовать знакомые нам приближенные методы. При этом применяют такую стратегию: начав со сравнительно маленького R (например, с $R = 0{,}01$), находят точку минимума функции $P(x;0{,}01)$. Затем увеличивают R, например, до 0,1 и находят точку минимума $P(x;0{,}1)$ в качестве начальной. Увеличивают R до тех пор, пока не окажется, что элементы итерационных последовательностей $x^{(t)}, f\left(x^{(t)}\right)$ и $P\left(x^{(t)}, R\right)$ изменяются от шага к шагу достаточно мало.

Пример 6.2. Минимизировать функцию

$$f(x) = (x_1 - 4)^2 + (x_2 - 4)^4$$

при ограничении

$$g(x) = 5 - x_1 - x_2 \ge 0.$$

Решение

Используем штраф типа квадрата срезки. Введем штрафную функцию:

Библиотека ГГТУ им. П.О.Сухого

$$P(x,R) = f(x) + R < g(x) >^2$$

$$P(x,R) = (x_1 - 4)^2 + (x_2 - 4)^4 + R < 5 - x_1 - x_2 >^2$$

Найдем ее стационарную точку:

$$\frac{\partial P}{\partial x_1} = 2(x_1 - 4) + 2R < 5 - x_1 - x_2 > (-1) = 0$$

$$\frac{\partial P}{\partial x_2} = 2(x_2 - 4) + 2R < 5 - x_1 - x_2 > (-1) = 0$$

Отсюда, как и в предыдущем случае, следует, что $x_1 = x_2$, а значит,

$$x_1 - 4 - R < 5 - 2x_1 >= 0 .$$

Рассмотрим отдельно случаи, когда аргумент оператора срезки больше нуля, равен нулю и меньше нуля.

Если $2x_1 \ge 5$, то

$$x_1 - 4 - R(5 - 2x_1) = 0 ,$$

откуда $x_1 = \dfrac{4+5R}{1+2R}$, или, в пределе

$$\lim_{R \to \infty} \frac{4+5R}{1+2R} = 2{,}5 .$$

Таким образом, метод сходится к точке $x^* = (2{,}5;2{,}5)$, $f_{\min} = 4{,}5$.

Можно построить таблицу значений и убедиться, что при увеличении R от 0 до ∞ стационарная точка функции $P(x,R)$ перемещается вдоль отрезка прямой, соединяющей точку (4;4) безусловного минимума функции $f(x)$ и точку (2,5;2,5) ее условного минимума.

Рассмотрим применение логарифмического штрафа на примере той же задачи. Положим

$$P(x,R) = f(x) - R\ln(g(x)) .$$

В данном случае,

$$P(x,R) = (x_1 - 4)^2 + (x_2 - 4)^4 - R\ln(5 - x_1 - x_2)$$

Условия стационарности примут вид:

$$\frac{\partial P}{\partial x_1} = 2(x_1 - 4) + \frac{R}{5 - x_1 - x_2} = 0$$

$$\frac{\partial P}{\partial x_2} = 2(x_2 - 4) + \frac{R}{5 - x_1 - x_2} = 0$$

Библиотека ГГТУ им.П.О.Сухого

откуда $x_1 = x_2$.

Тогда из уравнения $2(x_1 - 4) + \frac{R}{5 - 2x_1} = 0$ приходим к уравнению $2x_1^2 - 13x_1 + 20 - \frac{R}{2} = 0$.

Один из его корней $x_1 = \frac{13}{4} - \frac{1}{4}\sqrt{9 + 4R}$ (второй корень соответствует недопустимой точке).

В пределе при $R \to 0$ получим $x_1 = x_2 = 2{,}5$.

Рассмотрим еще один штраф – штраф, задаваемый обратной функцией.

$$P(x, R) = f(x) + \frac{R}{g(x)}.$$

В данном случае,

$$P(x, R) = (x_1 - 4)^2 + (x_2 - 4)^4 + \frac{R}{5 - x_1 - x_2}.$$

Условия стационарности примут вид:

$$\frac{\partial P}{\partial x_1} = 2(x_1 - 4) + \frac{R}{(5 - x_1 - x_2)^2} = 0$$

$$\frac{\partial P}{\partial x_2} = 2(x_2 - 4) + \frac{R}{(5 - x_1 - x_2)^2} = 0$$

откуда $x_1 = x_2$ и $4x_1^3 - 36x_1^2 + 105x_1 - 100 + \frac{R}{2} = 0$.

Положив $R = 100$, 10, 1, 0,1, 0,01, 0 можно численными методами найти корни.

Общие недостатки всех методов: сходимость метода связана со степенью вытянутости линий уровня штрафной функции, что увеличивает вероятность тог, что они окажутся неэффективными при отсутствии ограничений. Описания всех методов основывались на предположении, что все подзадачи безусловной оптимизации могут быть решены с достаточной точностью. Если уверенности в этом нет, то практически ставится под сомнение эффективность всего метода в целом.

Библиотека ГГТУ им. П.О.Сухого

***Замечание* 6.2:** При решении задачи (6.1)-(6.2), имеющей несколько ограничений равенств и неравенств, строят штрафную функцию вида

$$P(x,R) = f(x) + R\sum \Omega(x, g_j(x)) + R\sum (h_k(x))^2 .$$

Контрольные вопросы и задания

1. Для решения каких задач применяется метод штрафных функций? В чем его суть?
2. Что называется штрафной функцией?
3. Какие виды штрафов используются для ограничений-равенств?
4. Какие виды штрафов используются для ограничений-неравенств?
5. Дана задача:

$$\text{минимизировать } f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

при ограничениях $g_1(x) = x_1 + 1 \geq 0$,

$$g_2(x) = 1 - x_2 \geq 0,$$

$$g_3(x) = 4x_2 - x_1 - 1 \geq 0,$$

$$g_4(x) = 1 - 0{,}5x_1 - x_2 \geq 0.$$

Записать для этой задачи штрафные функции, используя:

1) логарифмический штраф; 2) штраф, задаваемый обратной функцией.

6. Дана задача:

$$\text{минимизировать } f(x) = (x_1 - 1)^2 + (x_2 - 1)^2$$

при ограничениях $g_1(x) = x_1 + x_2 - 1 \geq 0$,

$$g_2(x) = 1 - x_1 \geq 0$$

$$g_3(x) = 1 - x_2 \geq 0$$

и штрафная функция вида $P(x,R) = f(x) + R\sum_{i=1}^{J}\left[\dfrac{1}{g_i(x)}\right]$.

Найти зависимость координат стационарных точек функции $P(x,R)$ от R.

7. Выбрав по своему усмотрению один из штрафов, найти с точностью до трех десятичных знаков решение задачи:

$$\text{минимизировать } f(x) = x_1^2 + x_2^2 + x_3^2$$

при ограничениях $g_1(x) = 1 - x_2^{-1}x_3 \geq 0$,

$$g_2(x) = x_1 - x_3 \geq 0,$$

Библиотека ГГТУ им. П.О.Сухого

$$h_1(x) = x_1 - x_2^2 + x_2 x_3 - 4 = 0,$$
$$0 \le x_1 \le 5, \quad 0 \le x_2 \le 3, \quad 0 \le x_3 \le 3.$$

Библиотека ГГТУ им. П.О. Сухого

ЛИТЕРАТУРА

1. Аттетков, А. В. Методы оптимизации : учебник для втузов / под ред. В. С. Зарубина, А. П. Крищенко. – 2-е изд., стер. – Москва : МГТУ, 2003. – 439 с.
2. Аттетков, А. В. Введение в методы оптимизации : учеб.пособие / А. В. Аттетков, В. С. Зарубин, А. Н. Канатников. – Москва : Финансы и статистика; ИНФРА-М, 2008. – 272 с.
3. Банди, Брайн. Методы оптимизации: ввод. курс / Б. Банди ; пер. с англ. О. В. Шихеевой ; под ред. В. А. Волынского. – Москва : Радио и связь, 1988. – 128 с.
4. Васильев, О.В. Методы оптимизации и их приложения / О.В.Васильев, В. А. Срочко, В.А.Терлецкий; Отв. ред. А. П. Маренков; Сиб.энегет.ин-т СО АН СССР. – Новосибирск : Наука, 1990. – 150 с.
5. Реклейтис, Г. Оптимизация в технике: В 2-х кн. Кн.1. / Г. Реклейтис, А. Рейвиндран, К. Рэгсдел. – Москва : Мир, 1986. – 349 с.
6. Струченков, В. И. Методы оптимизации : основы теории, задачи, обучающие компьютерные программы: учеб. пособие / В. И. Струченков. – Изд.2-е, перераб. – Москва : Экзамен, 2007. – 255 с.

Библиотека ГГТУ им. П.О.Сухого

МАТЕМАТИЧЕСКИЕ МЕТОДЫ ПОИСКА ОПТИМАЛЬНЫХ РЕШЕНИЙ

Практикум
по одноименной дисциплине для студентов специальности 1-43 01 02 «Электроэнергетические системы и сети» дневной формы обучения

Составители: **Задорожнюк** Мария Викторовна
Бородин Николай Николаевич

Подписано к размещению в электронную библиотеку ГГТУ им. П. О. Сухого в качестве электронного учебно-методического документа 02.02.17.

Рег. № 56E.
http://www.gstu.by

Библиотека ГГТУ им. П.О.Сухого