Московский Авиационный Институт
(государственный технический университет)

Кафедра 308

Дисциплина
«Теория и методы оптимизации»

Лабораторная работа №4

«Методы прямого поиска в задачах условной минимизации»

(методические указания)

Объем – 4 часов

Составил доцент, к.т.н. В.Г.Герасименко

Москва, 2006г.

Работа 4. МЕТОДЫ ПРЯМОГО ПОИСКА В ЗАДАЧАХ УСЛОВНОЙ МИНИМИЗАЦИИ

1. ОБЩИЕ ЗАМЕЧАНИЯ

Идея использования прямых методов в задачах с ограничениями вполне прозрачна: после вычисления текущей итерации любым из уже известных методов прямого поиска (Хука-Дживса, симплекс-метод) нужно проверить эту точку на допустимость и, если хотя бы одно из ограничений нарушается, вычислить другую точку, вновь ее проверить и т.д.

В данной работе изучаются 2 метода прямого поиска минимума в задачах с ограничениями-неравенствами:

- модифицированный метод Хука-Дживса;
- метод комплексов Бокса.

Идеи, заложенные в методы прямого поиска, достаточно подробно изложены в лекциях по данному курсу (гл. 4, разд.4.3). Также основные процедуры алгоритмов прямого поиска для безусловных задач обсуждались в работе №2 данного курса «Алгоритмы прямого поиска в многомерной минимизации». Кроме того, можно использовать рекомендованную литературу.

Поэтому, в данном методическом пособии

- мы не повторяем уже известный теоретический материал;
- обращается внимание только на модификацию алгоритмов, которая обеспечивает применение этих алгоритмов к условным задачам;
- приводится необходимая последовательность выполнения практической части.

Напомним, что структура любых алгоритмов прямого поиска в задачах с ограничениями-неравенствами включает три основные части.

1) **Первая часть** реализует автоматический поиск допустимой начальной точки, поскольку траектория минимизации должна начинаться («стартовать») из допустимой точки.

2) **Вторая часть** – проверка правила «останова», т.е. окончания поиска.

3) **Третья часть** алгоритма вычисляет точки поиска, последовательность которых должна постепенно приближаться к окрестности искомого решения. Эта часть алгоритма – основная, она и определяет метод поиска. Важно четко понимать, что именно здесь каждая вновь вычисляемая точка поиска проверяется на допустимость, т.е. на принадлежность допустимой области. Конечно, если она не принадлежит

допустимой области пространства решений, то точка должна быть скорректирована.

Теперь разберем конкретное содержание указанных частей алгоритмов Хука-Дживса и Бокса, разработанных для минимизации условных задач.

2. МОДИФИЦИРОВАННЫЙ МЕТОД ХУКА-ДЖИВСА

2.1. О случайном поиске начальной допустимой точки для алгоритмов условной минимизации

Вначале заметим, что вычисление начальной допустимой точки, основанное на случайном выборе, пригоден для любого метода условной минимизации.

Графическая иллюстрация условной задачи и вычисления начальной допустимой точки дается на рис.1.1.

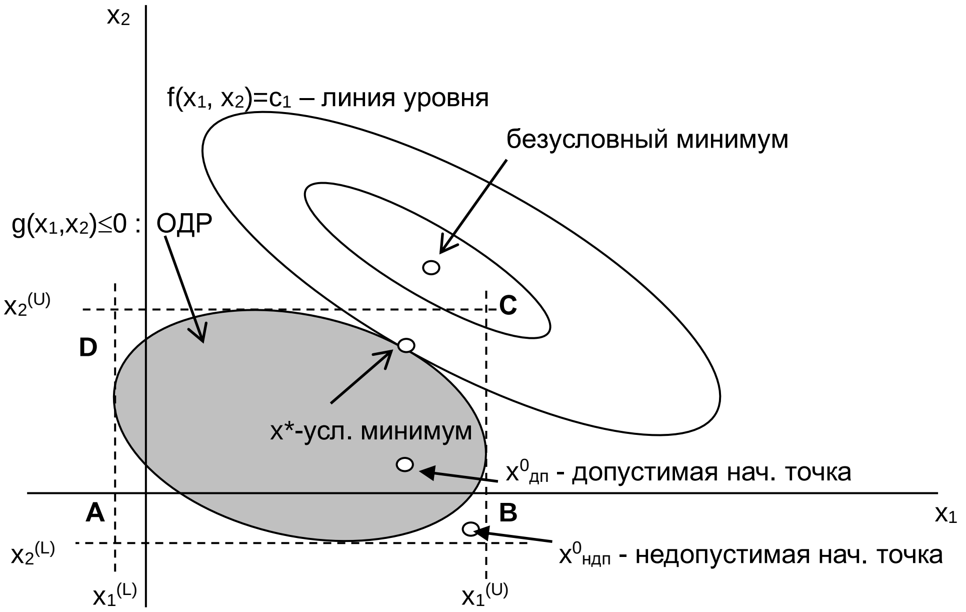

Рис. 1.1. Иллюстрация «проблемы» начальной точки в задачах условной минимизации

Здесь изображены 2 линии уровня $f(x_1,x_2)=c_i$ целевой функции и область допустимых решений. ОДР задается одним нелинейным уравнением-неравенством $g(x_1,x_2)\le 0$ и она обозначена серой заливкой. Ясно, что искомая точка условного минимума x* лежит на границе ОДР и не совпадает с точкой безусловного минимума.

На этом рисунке также показаны две начальные точки поиска – одна допустимая, а вторая точка – недопустимая.

Понятно, что **задание с клавиатуры** координат начальной допустимой точки $\mathbf{x^0_{дп}}$ даже в двумерном случае весьма затруднительно, а практически невозможно (**поясните – почему?**).

Поэтому возникает вопрос: как быстро и надежно вычислить начальную точку поиска, которая принадлежит ОДР?

Для этого, на примере задачи рис. 1.1 рассмотрим алгоритм **случайного поиска допустимой начальной точки.**

Алгоритм случайного поиска начальной допустимой точки

Вначале нужно провести **вспомогательную операцию:**

«вписываем» ОДР, заданную в начальных условиях задачи в **прямоугольник ABCD с границами** $x_i^{(U)}$ и $x_i^{(L)}$ (рис. 1.1):

- по оси $0x_1$ верхняя граница $x_1^{(U)}$ и нижняя $x_1^{(L)}$;
- по оси $0x_2$ верхняя граница $x_2^{(U)}$ и нижняя $x_2^{(L)}$.

Смысл буквенных индексов должен быть понятен:

(U) – «Upper», (L) – «Low».

Формально, на этом шаге пользователь аналитически задает простые ограничения задачи. Значит, для этого надо вычислить верхние $x_i^{(U)}$ и нижние $x_i^{(L)}$ границы каждой переменной x_i, i=1,2 (т.к. рассматривается 2-мерная задача).

После ввода в алгоритм вышеуказанных границ прямоугольника, начинет «работать» алгоритм случайного поиска.

Шаг 1. Найдем некоторую **пробную точку** $\mathbf{x^0}$ и **«равномерно бросим» ее на площадь прямоугольника ABCD**, в котоый висана ОДР.

Это означает, что здесь нужно вычислить:

$x^0_i = x_i^{(L)} + R_i(x_i^{(U)} - x_i^{(L)})$,

где $\mathbf{i = 1,2}$;

$\mathbf{R_i}$**-случайные числа, равномерно распределенные на интервале (0,1).**

Шаг 2. Очередную пробную точку x^0 нужно **проверить на принадлежность ОДР**: если она **принадлежит ОДР** – допустимая точка $x^0_д$ получена: $x^0 = x^0_д$; в противном случае нужно вычислить очередную случайную точку.

Формально, на этом шаге выполняется проверка условия:

если $x^0 \in$ ОДР, **то** $x^0 = x^0_д$; **перейти** к поиску минимума;

иначе - перейти к **шагу 1**.

Замечание. *При программировании этого алгоритма разработчик должен предусмотреть разумное ограничение по числу обращений к шагу 2, т.к. в случае некорректно сформулированной задачи, допустимая точка может отсутствовать.*

2.2. Модифицированный метод Хука-Дживса для условной минимизации

Этот раздел будет достаточно кратким, поскольку метод Хука-Дживса уже изучался в л/р №2.

Модификация метода Хука-Дживса для условной минимизации вполне очевидна. При вычислении целевой функции для учета ограничений-неравенств будет вполне достаточно присвоить функции **«очень большое»** значение там, где ограничения нарушаются.

Эту идею очень просто реализовать при программировании. Заметим, что «очень большое» значение целевой функции вне допустимой области может быть таким: 10^{10}. Например[1], иногда используется еще большее число 1e+30.

Бейсик-программа модифицированного метода Хука-Дживса для условных задач минимизации

Конкретная модификация алгоритма для задач условной минимизации должна быть понятна из основной части нижеприведенного текста готовой Бейсик-программы[1].

```
10   PRINT "МЕТОД ХУКА-ДЖИВСА ПРИ НАЛИЧИИ ОГРАНИЧЕНИЙ"
20   REM Целевая функция вычисляется в строке 2000
30   PRINT "Введите число переменных": INPUT N
40   DIM X(N),B(N),Y(N),P(N)
50   PRINT "Введите начальную точку X1,X2,… ,XN"
60   FOR I=1 TO N: INPUT X(I): NEXT I
70   PRINT "Введите длину исследующего шага": INPUT H
80   K=H: FE=0
90   FOR I=1 TO N
100  Y(I)=X(I): P(I)=X(I): B(I)=X(I): NEXT I
110  GOSUB 2000: FI=Z
120  PRINT "НАЧАЛЬНОЕ ЗНАЧЕНИЕ ФУНКЦИИ" Z
130  FOR I=1 TO N: PRINT X(I);"   ";:NEXT I:PRINT ""
140  PS=0: BS=0
150  REM Исследование вокруг базовой точки
180  J=1: FB=FI
200  X(J)=V(J)+K
210  GOSUB 2000
220  IF Z<FI THEN GOTO 280
230  X(J)=V(J)-K
240  GOSUB 2000
250  IF Z<FI THEN GOTO 280
260  X(J)=Y(J)
270  GOTO 290
280  Y(J)=X(J)
```

[1] Б.Банди. Методы Оптимизации Вводный курс. М.,«Радио и связь». 1988.

```
290  GOSUB 2000
300  FI=Z
310  PRINT “ИССЛЕДУЮЩИЙ ПОИСК” Z
320  FOR I=1 TO N: PRINT X(I);”   “;:NEXT I:PRINT “”
330  IF J=N THEN GOTO 360
340  J=J+1
350  GOTO 200
```

Для составления собственной программы метода Хука-Дживса внимательно проанализируйте структуру и вычисления вышеприведенной программы.

Вычисление целевой функции осуществляется в подпрограмме с меткой 2000. Эта подпрограмма здесь не отображена, поскольку вам нужно разработать самостоятельно. В связи с этим укажем, что в задачах безусловной минимизации, целевая функция в этой подпрограмме 2000 вычислялась без ограничений как обычно, например:

```
2000 Z=(X(1)-2)^2+(X(2)-5)^2+(X(3)+2)^4
2010 FE=FE+1
2020 REM Счетчик количество вычислений функции
2030 RETURN
```

Переходим к изучению метода комплексов Бокса, не забывая о том, что алгоритм случайного поиска начальной допустимой точки, являющийся необходимой частью любого метода условного поиска, уже был рассмотрен выше.

3. МЕТОД КОМПЛЕКСОВ БОКСА

Рассмотрим достаточно эффективный метод прямого поиска, разработанный Боксом[2], который обычно называют **методом комплексов.**

Бокс предложил строить на каждой итерации **не регулярный симплекс** с n+1 вершинами, где n-размерность задачи,

а произвольный симплекс, состоящий из $P \geq n+1$ точек, причем **каждая точка начального симплекса строится случайным образом.**

Алгоритм метода комплексов

Шаг 1. Задать начальные параметры поиска:

α - коэффициент отражения;

ε, δ - параметры останова;

[2] Г.Реклейтис и др. Оптимизация в технике. Кн.1.М.,«Мир», 1986.

$x_i^{(L)}, x_i^{(U)}$ - нижние и верхние границы значений всех переменных ($i=1,\dots,n$);

2. Построить начальный комплекс, из P допустимых точек x^p ($p=0$, P) с помощью генератора стандартных случайных чисел:

2.1. определить координаты точки x_i^p ;

2.2. если x^p недопустима, то найти центр тяжести x^c уже найденных точек (если такие существуют) и положить

$x^p = x^p + 0.5(x^c - x^p)$; повторять процедуру до тех пор, пока x^p не станет допустимой;

иначе, перейти к 2.1. и повторять (2.1) до тех пор, пока $p=P$.

2.3. вычислить $f(x^p)$ для $p=0,1,\dots,P-1$.

3. Отражение комплекса:

3.1. выбрать отражаемую точку x^R, для которой

$$f(x^R) = \max f(x^p) \equiv F_m.$$

3.2. вычислить центр тяжести x^c оставшихся точек (без отражаемой), а затем новую, отраженную точку:

$$x^m = x^c + \alpha(x^c - x^R);$$

3.3. если x^m - недопустимая точка и $f(x^m) \geq F_m$, то уменьшить в два раза расстояние между x^m и центром тяжести x^c, продолжать поиск, пока $f(x_m) < F_m$; перейти к 5.

иначе,

если x^m -допустимая точка и $f(x^m) < F_m$, то перейти к 5;

4. Скорректировать отраженную точку для обеспечения допустимости:

4.1. если $x_i^m < x_i^{(L)}$, то положить $x_i^m = x_i^{(L)}$;

если $x_i^m > x_i^{(U)}$, то положить $x_i^m = x_i^{(U)}$;

4.2. если x^m – недопустимая точка, то уменьшить в два раза расстояние до центра тяжести; продолжать до тех пор, пока x^m не станет допустимой точкой;

5. Проверить условия окончания поиска:

5.1. вычислить $f_s = \frac{1}{P}\sum f(x^p)$ и $x = \frac{1}{P}\sum x^p$;

5.2. если $\sum(f(x^p) - f_s)^2 \leq \varepsilon$ и $\sum(x^p - x)^2 \leq \delta$, то печатать результаты поиска и «Останов»;

иначе перейти к 2.1.

Замечание. Практика показывает, что рекомендуется **α=1.3** и значение P выбирать от **(n+2)** до **2n**, где n – размерность задачи.

5. КОНТРОЛЬНЫЕ ВОПРОСЫ

5.1. Почему необходимо исключать ограничения-равенства при использовании методов прямого поиска?

5.2. Какова причина использования более чем n+1 точки в алгоритме комплексов?

5.3. Покажите графически возможные трудности работы метода комплексов при невыпуклой области допустимых решений.

5.4. Почему в алгоритме комплексов используется процедура возврата?

5.5. Является ли выпуклой область:

$$x_1+x_2+x_3=1,$$

$$x_1^2+x_2^2\le 1,$$

$$x_3+x_2^2\ge 1?$$

Обоснуйте ответ.

5.6. Дана задача

минимизировать $f(x)= x_1^2+x_2^2+x_3^2$

при ограничениях $1-x_2^{-1}x_3\ge 0,$

$$x_1-x_3\ge 0,$$

$$x_1-x_2^2+x_2x_3-4=0,$$

$$0\le x_1\le 5,\ 0\le x_2\le 3,\ 0\le x_3\le 3.$$

Какие преобразования необходимо выполнить, чтобы можно было использовать метод комплексов?

Приведите окончательную форму задачи.

6. ПРАКТИЧЕСКАЯ ЧАСТЬ

5.2. Практическое исследование алгоритма комплексов

5.2.1. Запрограммируйте и отладьте программу минимизации методом комплексов.

С помощью программы решите следующую задачу.

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

при ограничениях $10-x_1^2-x_2^2\ge 0$,

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

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

Положите вначале $\alpha=1.3$.

Дайте геометрическую интерпретацию задачи. Изобразите на этом рисунке около 10 начальных точек поиска.

5.2.2. Методом комплексов решите задачу[3]:

минимизировать $f(x)= (x_1-1)^2+(x_1-x_2)^2+(x_2-x_3)^2$

при ограничениях $x_1(1+x_2^2)+x_2^4-4-3\sqrt{2}=0$,

$0\le x_i\le 3$.

5.2.4. Для разработки конструкции водонепроницаемых переборок минимального веса для танкера требуется решить следующую задачу[4]:

минимизировать $f(x)=\dfrac{5.885x_4(x_1+x_3)}{x_1+(x_3^2-x_2^2)^{1/2}}$

при ограничениях $x_2x_4(0.4x_1+\frac{1}{6}x_3)-8.94[x_1+(x_3^2-x_2^2)^{1/2}]\ge 0$,

$x_2^4x_4(0.2x_1+{}^1/_{12}x_3)-2.2\{8.94[x_1+(x_3^2-x_2^2)^{1/2}]\}^{4/3}\ge 0$,

$x_4-0.0156x_1-0.15\ge 0$,

$x_4-0.0156x_3-0.15\ge 0$,

$x_4-1.05\ge 0$,

$x_3-x_2\ge 0$.

Переменные x_i в этой задаче соответствуют следующим линейным размерам переборок: x_1-ширина, x_2-глубина и x_4-толщина.

5.2.5. Необходимо спроектировать прямоугольную конструкцию в форме параллелепипеда с открытой передней стенкой.

---

[3] См.Miele et al.,J.Optim. Theory Appl.,10,1 (1972).(f*=0.032567)

[4] См.Kvalie et al.,Tech. Rep., Norwegian Technical Institute, Trondheim, Norway, 1966. (f*=6.84241)

Конструкция должна иметь объем 16 000 фут^3, а периметр ее основания не должен превышать 220 фут.

Глубина не должна быть больше 60 фут, а ширина – 80 фут.

Кроме того, ширина не должна превышать утроенной глубины, а высота – 2/3 ширины.

Стоимость гофрированного материала, из которого изготовляются крыша и три стенки конструкции, составляют 1долл за квадратный фут.

Требуется определить размеры конструкции таким образом, чтобы минимизировать стоимость материалов.

Для решения этой задачи пользуйтесь инструкциями:

- нарисуйте чертеж конструкции, обозначив, например, x_1-глубина, x_2-ширина, x_3-высота;
- выпишите целевую функцию как стоимость крыши и всех стенок;
- напишите ограничение-равенство на требуемый объем, ограничения-неравенства на линейные размеры, а также простые ограничения изменения переменных и получите стандартную форму данной задачи, которая будет содержать 1 ограничение-равенство, 2 ограничения неравенства и 3 простых ограничений, среди которых будет одностороннее ограничение;
- преобразуйте задачу, исключив ограничение-равенство;
- изобразите линии уровня полученной целевой функции и область допустимых значений;
- покажите на фоне линий уровня 10-12 начальных итераций поиска минимума методом комплекса;
- решите задачу с точностью до 10^{-6};
- найдите минимум этой же (преобразованной) целевой функции без ограничений.