

ОСНОВНЫЕ ВОПРОСЫ ПОСТРОЕНИЯ СТАНДАРТНОЙ ПРОГРАММЫ «ОПТИМИЗАТОР»

С. А. РОДИОНОВ, Е. И. ГУТМАН

Излагаются основные вопросы построения оптимизирующего алгоритма, оформленного авторами в виде стандартной программы, позволяющей успешно решать задачи, связанные с оптимизацией нелинейных систем.

Оптимизация нелинейных систем в настоящее время представляет собой одну из наиболее важных и распространенных задач прикладной математики и встречается практически во всех отраслях техники.

В статье излагаются основные вопросы построения оптимизирующего алгоритма широкого спектра применения, оформленного авторами в виде стандартной программы (СП) для ЦВМ «Минск-22» и испытанного большим числом задач в вычислительном центре ЛИТМО [1].

Постановка задачи. В общем виде задачу оптимизации можно сформулировать следующим образом. Имеется система m нелинейных функций от n параметров (переменных):

$$\begin{aligned} \tilde{f}_1 &= \tilde{f}_1(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n) \\ \tilde{f}_m &= \tilde{f}_m(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n) \end{aligned}$$

или в матричной форме:

$$\tilde{F} = \tilde{F}(\tilde{X}).$$

Необходимо в пространстве параметров \tilde{X} найти такую точку $\tilde{X}_{\text{опт}}$, которая была бы оптимальной по отношению к некоторой оценке функций $\varphi(\tilde{F})$.

Для наибольшей общности задачи не требуется явного задания функций \tilde{F} , поэтому ограничимся требованиями однозначности, непрерывности и дифференцируемости. Предположим, что имеется некоторый алгоритм или устройство для получения значений \tilde{F} в любой точке \tilde{X} . Процесс получения этих значений, т. е. обращение к такому алгоритму или устройству, называется «пробой» функций.

Содержание и объем «пробы» определяется конкретной технической задачей оптимизации. Можно предположить также, что имеется алгоритм «проба производных», который в любой точке \tilde{X} дает значения частных производных функций по параметрам, т. е. значения матрицы производных: $A = \{d\tilde{f}_i / d\tilde{x}_j\}$, $\varphi i = \overline{1, m}$; $\varphi j = \overline{1, n}$. В простейшем случае, при вычислении производных методом конечных разностей, алгоритм «проба производных» состоит из n раз повторенного алгоритма «проба».

Выбор оценочной функции. Нормирование пространств. Вид функционала, определяющего оценочную функцию $\varphi(\tilde{F})$, произволен, но должен удовлетворять двум очевидным требованиям: 1) адекватность

величины оценочной функции и критерия качества оптимизируемой системы, т. е. меньшая (или большая) величина φ должна соответствовать лучшей системе; 2) простота и удобство этого функционала для построения алгоритма. Наиболее распространенным и, по-видимому, наиболее удобным является выбор φ в виде квадрата длины вектора \tilde{F} . Для определения длины вектора в m -мерном пространстве необходимо ввести некоторую метрику. Если M_f — матрица-метрика пространства функций (M_f — симметричная, положительно определенная матрица), то φ представляет собой квадратичную форму \tilde{F} :

$$\varphi = \tilde{F}^T M_f \tilde{F}. \quad (1)$$

Тогда (1) можно представить в виде:

$$\varphi = F^T F, \quad (2)$$

где $F = M_f^{1/2}$, \tilde{F} — вектор-функция в нормированном пространстве с единичной метрикой.

В большинстве случаев, а также в СП, разработанной авторами, метрика M_f есть диагональная матрица; тогда нормирование можно толковать как преобразование масштабов по осям пространства.

В пространствах параметров \tilde{X} также вводится метрика M_x , после чего это пространство преобразуется в нормированное пространство X с единичной метрикой. В дальнейшем будем считать, что оптимизатор имеет дело с нормированными пространствами. Правильный выбор метрик M_x , M_f , а также и самих функций f_i заключается как раз в обеспечении адекватности оценочной функции и критерия качества системы. Итак, в нормированных пространствах X , F задача состоит в минимизации суммы квадратов функций F .

Рассмотрим только основные аспекты построения детерминированных локальных методов поиска. Локальные методы поиска представляют собой итеративные процессы, каждая итерация которых состоит из следующих этапов: анализ поведения системы в окрестности исходной точки $X_0^{(k)}$ и построение простой математической модели системы; построение одномерной траектории $\Delta X(s)$ движения в сторону предполагаемого минимума; выбор точки на траектории (выбор параметра s), которая является конечной точкой данной итерации $X_i^{(k)}$. Далее точка $X_i^{(k)}$ принимается за начальную точку $X_0^{(k+1)}$ следующей итерации, и процесс повторяется. Предполагается, что процесс удовлетворяет условию $\varphi(X_i^{(k)}) < \varphi(X_0^{(k)})$, т. е. в процессе оптимизации происходит монотонное уменьшение оценочной функции φ и, следовательно, процесс неизбежно сходится к локальному минимуму.

Анализ поведения системы в исходной точке и построение математической модели системы. Полную модель поведения системы в окрестности исходной точки дает разложение ее в ряд Тейлора. Однако «проба»

и «проба производных» дают нам информацию, достаточную только для построения первых двух членов ряда. Таким образом, приходим к линейной модели поведения системы в окрестности X_0 :

$$F_L(X) = F_0 + A\Delta X, \quad F_0 = F(X_0), \quad \Delta X = X - X_0. \quad (3)$$

Рассмотрим градиент ϕ :

$$\vec{\text{grad}}\phi = \left\{ \frac{\partial \phi}{\partial x_j} \right\}; \quad j = \overline{1, n}. \quad (4)$$

Обозначим $G = \frac{1}{2} \vec{\text{grad}} \phi$ (впоследствии G будем называть градиентом), тогда линейная аппроксимация для градиента G

$$G_L(X) = N + M\Delta X = G_0 + M\Delta X, \quad (5)$$

где $N = A^T F_0$ – матрица-столбец, размерности $n \times 1$; $M = A^T A$ – квадратная, симметричная неотрицательно определенная матрица порядка $n \times n$.

Построение траектории движения. Выбор той или иной траектории движения является основной чертой, отличающей различные методы и обычно определяющей и название метода. Выбрать траекторию – значит указать конкретный вид зависимости вектора $\Delta X(s)$ точки на этой траектории от параметра s . Рассмотрим траектории нескольких наиболее употребительных методов, заложенных в СП, разработанных авторами [1].

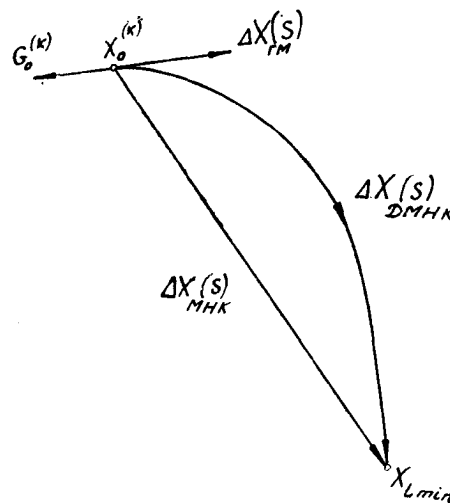


Рис.1. Траектория оптимизации: $X_0^{(k)}$ – начальная точка k -той оптимизации; $G_0^{(k)}$ – направление градиента в начальной точке $X_0^{(k)}$; $\Delta X_{ГМ}(s)$ – траектория градиентного метода (ГМ); $\Delta X_{МНК}(s)$ – траектория стандартного метода наименьших квадратов; $\Delta X_{ДМНК}(s)$ – траектория демпфированного метода наименьших квадратов (ДМНК); X_{Lmin} – минимум линейной модели системы

В градиентном методе (ГМ) (методе быстрого спуска) траектория движения представляется прямой линией, обратной по направлению градиенту в исходной точке (рис. 1).

$$\Delta X(s) = -s\alpha G_0 = -s\alpha N, \quad (6)$$

где $\alpha = \frac{NN}{N^T M N^T}$.

Метод Ньютона (МН) применяется только в случае $m = n$. Он основан на решении линейной системы, полученной из приравнивания $F_L(X)$ нулю:

$$F_L(X) = F_0 + A\Delta X = 0 \quad (7)$$

Отсюда траектория представляет собой прямую линию, проходящую через минимум (обязательно нулевой) линейной модели (рис. 1):

$$\Delta X(s) = -sA^{-1}F_0. \quad (8)$$

Развитием метода Ньютона на случай $m \neq n$ можно считать известный метод наименьших квадратов (МНК) [2]. Траектория определяется из решения нормальной системы, полученной из линейной модели (5) градиента, на основании условия

$$G_L = N + M\Delta X = 0. \quad (9)$$

Следовательно, траектория представляет собой прямую линию, проходящую через минимум линейной модели, т. е. точку, в которой (см. рис. 1) $G = 0$:

$$\Delta X(s) = -sM^{-1}N. \quad (10)$$

Наиболее эффективным и универсальным является демпфированный метод наименьших квадратов (ДМНК) [2]. В этом методе вводится «демпфирование» траектории, а именно, в оценочную функцию φ добавляется квадрат длины вектора траектории $\Delta X^T \Delta X$:

$$\varphi_p = \varphi + p^2 \Delta X^T \Delta X, \quad (11)$$

где p — демпфер. Траектория движения определяется из решения так называемой демпфированной нормальной системы (см. рис. 1):

$$N + M_p \Delta X = 0, \quad (12)$$

где $M_p = M + p^2 I$.

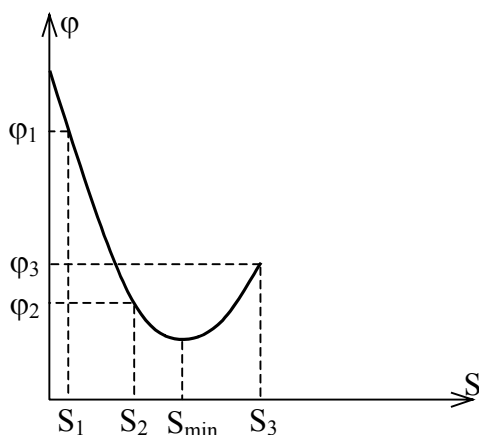


Рис. 2. Зависимость оценочной функции φ от длины шага s .

Выбор длины шага. Выбор длины шага должен обеспечивать наиболее быструю сходимость метода. В основном применяются три различных метода для выбора s : метод релаксации, метод слежения за областью линейности и метод оптимизации оценочной функции по длине шага. В программе, разработанной авторами, используется последний. Так как оптимизация здесь производится по одному параметру s , наиболее эффективны глобальные методы оптимизации, например, метод параболической аппроксимации.

Сначала путем нескольких проб на траектории находят три точки: s_1, s_2 и s_3 так, чтобы

$$\phi_1 = \phi(s_1) > \phi_2 = \phi(s_2) < \phi_3 = \phi(s_3). \quad (13)$$

Затем через эти точки проводится парабола, и ее вершина принимается за минимум однопараметрической оптимизации (рис. 2). Методы, использующие такой выбор длины шага, обычно называют оптимальными.

К задачам, успешно решаемым разработанной программой, следует отнести: оптимизацию нелинейных оптических систем, решение задач анализа и синтеза в теории зубчатых зацеплений, оптимизацию химических процессов и многие другие инженерные задачи.

ЛИТЕРАТУРА

1. Правила и инструкция по программированию на ЭВМ «Минск-22» на языке «АКИ» с использованием транслятора АКИ-400. Под ред. Г.Л. Головановского, Г.Я. Малахова и Г.И. Новикова. Л., 1969.
2. Levenberg K.A. Method for Solution of certain non-linear problems in lest squares. „Quarterly of Applied Mathematics", vol. 2, p. 164, No 2, 1944.