

МЕТОД РЕШЕНИЯ СИСТЕМ НЕЛИНЕЙНЫХ УРАВНЕНИЙ ПОТЕНЦИАЛЬНОГО ТИПА

С. В. Симеонов (София, Болгария)

В работе рассматриваются системы нелинейных уравнений в случае, когда они являются условием минимума функционала, заданного в n -мерном евклидовом пространстве E_n . Для решения системы применяется специальный метод спуска, который приводит к последовательному включению неизвестных. Таким образом, задача решена после включения последней неизвестной. Показано, что в случае линейной системы этот спуск приводит к вычислительной схеме метода окаймления.

Исследование большого числа физических проблем, особенно в области механики и автоматики, приводит к решению систем уравнений, представляющих условие экстремума некоторого функционала $f(x)$. Такие системы будем называть системами потенциального типа. Здесь будет изложен метод решения таких систем в случае, когда (x) задан в некоторой области G n -мерного евклидова пространства E_n .

1. Пусть задана система нелинейных уравнений

$$f_i(x) = 0 \quad (i = 1, \dots, n) \quad x = (x_1, x_2, \dots, x_n) \in G \quad (1.1)$$

Будем предполагать, что система (1.1) является условием минимума функционала $f(x)$, о котором известно, что он ограничен снизу и имеет единственную стационарную точку x^* . Очевидно,

$$f_i(x) = \partial f(x) / \partial x_i \quad (1.2)$$

При определении точки x^* условие потенциальности оператора $f_x = (f_1, \dots, f_n)$ позволяет вместо уравнений (1.1) рассматривать функционал $f(x)$ независимо от применяемых способов исследования. Это важное свойство широко используется ([1-4] и др.) как для качественного исследования этого типа операторов, заданных в довольно общих банаховых пространствах, так и для эффективных методов их решения.

Пусть $x^{(0)} \in G$ — некоторая выбранная начальная точка. При наших условиях $f(x^{(0)}) \geq f(x^*)$, где равенство имеет место только, если $x^{(0)} = x^*$. Предполагая, что необходимо разработать такой план (стратегия) спуска с $f(x^{(0)})$, чтобы обязательно и быстрее достигнуть $f(x^*)$. Подобные методы ([4, 5] и др.) обычно осуществляют спуск по прямой самого крутого наклона, исходящего из $f(x^{(0)})$. Длина пути спуска по этому направлению определяется по различным критериям, которые часто приводят к процессам довольно медленно сходящимся, особенно при подходе к $f(x^*)$.

Здесь принят другой план спуска, который в случае двух переменных можно условно назвать спуском по «долинам (уврагам)». Эти «увраги» отличаются от тех, которые были введены И. М. Гельфандом. Такой спуск при принятых предположениях относительно функционала $f(x)$ всегда приводит к требуемой точке $f(x^*)$.

Выполним, во-первых, спуск с $f(x^{(0)})$ по направлению x_1 до самой нижней точки $f(x^{(1)})$, где $x_i^{(1)} = x_i^{(0)}$ при $i \geq 2$. Очевидно, что $x_1^{(1)}$ будет решением уравнения

$$f_1(x_1; x_2^{(1)}, \dots, x_n^{(1)}) = 0 \quad (1.3)$$

рассматриваемое как уравнение одной переменной x_1 .

Точку $x^{(1)}$ будем рассматривать как дно какого-то «уврага». Чтобы дальше спуститься по нему, необходимо иметь еще одну точку $x^{(1)'}$. Ее получим также из (1.3), если дадим некоторой другой координате, например x_2 , приращение Δx_2 . Для этого необходимо решить уравнение

$$f_1(x_1; x_2^{(0)} + \Delta x_2, x_3^{(0)}, \dots, x_n^{(0)}) = 0 \quad (1.4)$$

Корень его будет $x_1^{(1)'} = x_1^{(1)} + \Delta x_1$. Направление t , по которому будем осуществлять дальнейший спуск, выражается при помощи параметра t по формулам

$$x_i(t) = x_i^{(1)} + \frac{\Delta x_i}{\Delta s} t \quad (i = 1, 2), \quad x_i(t) = x_i^{(1)} = x_i^{(0)} \quad (i = 1, 2) \quad (1.5)$$

$$(\Delta s = \sqrt{(\Delta x_1)^2 + (\Delta x_2)^2})$$

Можно положить также $\Delta s = 1$

Спуск по t продолжается до достижения самой низкой точки $f(x(t^*))$. Нетрудно видеть, что t^* будет решением уравнения

$$f_1(x(t)) \Delta x_1 + f_2(x(t)) \Delta x_2 = 0 \quad (1.6)$$

Левая часть этого уравнения выражает приращение функционала по направлению t . Спуск с $x^{(0)}$ до $x(t^*) = x^{(1,1)}$ будет называть первым подциклом.

Принимая дальше $x^{(1,1)}$ в качестве $x^{(0)}$, выполняем совершенно аналогично второй подцикл и т. д. до тех пор, пока $x^{(1,m)}$ и $x^{(1,m+1)}$ не окажутся достаточно близкими. Предельную точку обозначим через $x^{(2)}$. Ее существование следует из принятых предположений относительно функционала.

Покажем, что достигнутая точка $x^{(2)}$ будет решением системы двух уравнений

$$f_1(x_1, x_2; x_3^{(0)}, \dots, x_n^{(0)}) = 0 \quad f_2(x_1, x_2; x_3^{(0)}, \dots, x_n^{(0)}) = 0 \quad (1.7)$$

В самом деле, чтобы $x^{(2)}$ была решением системы (1.7), она должна удовлетворять условию $\Delta f = 0$, т. е. условию (1.6) при любых достаточно малых Δx_1 и Δx_2 . Но именно это условие осуществляется описанной выше вычислительной процедурой.

Поскольку Δx_2 — подобранная достаточно малая величина, то, если $f_1(x)$ допускает частные производные по x_1 и x_2 , (1.4) можно записать в виде

$$a_{11} \Delta x_1 + a_{12} \Delta x_2 = 0, \quad \text{или} \quad \Delta x_1 = -\frac{a_{12}}{a_{11}} \Delta x_2 \quad (1.8)$$

Здесь

$$a_{11} = a_{11}(x^{(1)}) = \left. \frac{\partial f_1(x)}{\partial x_1} \right|_{x=x^{(1)}}, \quad a_{12} = \left. \frac{\partial f_1(x)}{\partial x_2} \right|_{x=x^{(1)}} \quad (1.9)$$

Эти коэффициенты можно определить и при помощи численного дифференцирования, что особенно удобно при применении вычислительных машин.

Решение уравнений типа (1.6) рассмотрим ниже.

Продолжим теперь начатый процесс следующим образом. Дадим достигнутой точке $x^{(2)}$ малое приращение Δx_3 и снова решим систему (1.7)

$$\begin{aligned} f_1(x_1, x_2; x_3^{(0)} + \Delta x_3, x_4^{(0)}, \dots, x_n^{(0)}) &= 0 \\ f_2(x_1, x_2; x_3^{(0)} + \Delta x_3, x_4^{(0)}, \dots, x_n^{(0)}) &= 0 \end{aligned} \quad (1.10)$$

Отсюда получим $x_1^{(2)'}$ и $x_2^{(2)'}$, при помощи которых вычисляем $\Delta x_1 = x_1^{(2)'} - x_1^{(2)}$, $\Delta x_2 = x_2^{(2)'} - x_2^{(2)}$. Эти значения вместе с принятым Δx_3 определяют новое направление спуска t . Это направление имеет следующие текущие координаты:

$$\begin{aligned} x_i(t) = x_i^{(2)} + \frac{\Delta x_i}{\Delta s} t \quad (i = 1, 2, 3), \quad x_i(t) = x_i^{(2)} = x_i^{(0)} \quad (i > 3) \\ \Delta s = \sqrt{(\Delta x_1)^2 + (\Delta x_2)^2 + (\Delta x_3)^2} \end{aligned} \quad (1.11)$$

Значение параметра t , при котором функционал $f(x)$ достигает наименьшего значения на t , определяется как корень уравнения

$$f_1(x(t)) \Delta x_1 + f_2(x(t)) \Delta x_2 + f_3(x(t)) \Delta x_3 = 0 \quad (1.12)$$

где $x(t)$ выражается через (1.11). Как и выше, это есть первый подцикл, который повторяется до получения с необходимой точностью точки $x^{(3)}$. И здесь при принятом малом Δx_3 можно вычислить Δx_1 и Δx_2 из линейной системы

$$\begin{aligned} a_{11} \Delta x_1 + a_{12} \Delta x_2 &= -a_{13} \Delta x_3 \\ a_{21} \Delta x_1 + a_{22} \Delta x_2 &= -a_{23} \Delta x_3 \end{aligned} \quad (1.13)$$

где

$$a_{ik} = a_{ki} = \left. \frac{\partial f_i(x)}{\partial x_k} = \frac{\partial f_k(x)}{\partial x_i} \right|_{x=x^{(2)}} \quad (1.14)$$

Предполагается, что функции f_i дифференцируемы.

Отметим, что системы (3.2) и (3.3) имеют одну и ту же обратную матрицу A_{k-1}^{-1} . Составив уравнение (1.17) и имея в виду (3.3), нетрудно получить

$$x_k^{(k)} = \sum_{j=1}^k \beta_{kj}^{(k)} b_j, \quad \beta_{kj}^{(k)} = -\frac{\Delta x_j^{(k)}}{\alpha^k}, \quad \alpha_k = \sum_{j=1}^k a_{jk} \Delta x_j \quad (3.4)$$

Очевидно $\beta_{kj}^{(k)}$ являются элементами матрицы A_k^{-1} . Для определения остальных $\beta_{ij}^{(k)}$, исключив из (1.17) $t/\Delta s$, получим

$$x_i^{(k)} = x_i^{(k-1)} + \Delta x_i^{(k)} x_k^{(k)} = \sum_{j=1}^k \beta_{ij}^{(k)} b_j, \quad \beta_{ij}^{(k)} = \beta_{ij}^{(k-1)} + \Delta x_i^{(k)} \beta_{kj}^{(k)} \\ (i, j = 1, 2, \dots, k-1) \quad (3.5)$$

Вычисление $\Delta x_i^{(k+1)}$ осуществляется при помощи той же обратной матрицы A_k^{-1}

$$\Delta x_i^{(k+1)} = \sum_{j=1}^k \beta_{ij}^{(k)} a_{j,k+1} \quad (i = 1, 2, \dots, k), \quad \Delta x_{k+1}^{(k+1)} = 1 \quad (3.6)$$

При $k = n$ получаем требуемое решение x^* системы (3.1).

Полученный вычислительный процесс, как и следовало ожидать, соответствует известному методу окаймления [7]. Здесь, однако, пришли к нему по совершенно иному пути, вследствие чего вычисляемые промежуточные величины имеют другой смысл.

4. На базе изложенного метода есть возможность строить различные варианты, позволяющие вносить одни или другие упрощения в вычислительную схему. Это будут разные процессы последовательных приближений, число циклов в которых теперь возрастет и будет зависеть от требуемой точности. Здесь рассмотрим два таких варианта.

Примем, во-первых, следующий план спуска. Сначала выполним спуск по x_1 и x_2 , как это было сделано в п. 1. После этого, приняв достигнутую точку $x^{(2)}$ как исходной, осуществим такой же спуск по x_2 и x_3 . Далее, выбирая последовательно разные пары координат, в которых принимают участие все координаты, выполняем тот же спуск до тех пор, пока по всем парам не получим пренебрежимо малые поправки. Пары координат лучше выбирать так, чтобы в одной паре участвовали такие координаты, которые наисущественно зависят одна от другой. Сходимость изложенного процесса следует непосредственно из теоремы 2 работы [2, 3], но ее можно доказать и так, как было сделано в п. 1. Вычислительная схема этого процесса выражается в следующем.

Предположим, что, спускаясь по x_i и x_j , достигли точки $x^{(v)}$. Выбираем новую пару, содержащую и одну из предыдущих координат, например, x_i, x_k . Тогда, выбрав Δx_k , как в п. 1, аналогично (1.8), находим

$$\Delta x_i = -\frac{f_{ik}(x^{(v)})}{f_{ii}(x^{(v)})} \Delta x_k$$

и решаем уравнение

$$f_i(x(t)) \Delta x_i + f_k(x(t)) \Delta x_k = 0$$

Здесь

$$x_i(t) = x_i^{(v)} + \Delta x_i t, \quad x_k(t) = x_k^{(v)} + \Delta x_k t, \quad x_j(t) = x_j^{(v)} \quad (j \neq i, k)$$

Так получаем точку $x^{(v+1)}$. Потом выбираем новую пару координат и т. д. Изложенная процедура избавляет от необходимости многократно решать нарастающие системы линейных уравнений типа (1.19), но может привести к уменьшению скорости сходимости.

Предположим теперь, что неизвестные можно разбить на несколько групп так, чтобы в каждой группе участвовали те неизвестные, которые больше всего зависели бы

одно от другого. Потом группируем и уравнения системы (1.1) в подсистемы так, как группировали и неизвестные. По методу п. 1 последовательно решаем каждую подсистему, принимая достигнутую точку при решении данной подсистемы в качестве исходной при решении следующей.

Процесс заканчивается, когда уточнение неизвестных станет пренебрежимо малым. Сходимость этого процесса доказывается как и в предыдущем случае.

Поступила 5 VIII 1966

ЛИТЕРАТУРА

1. В а й н б е р г М. М. Вариационные методы исследования нелинейных операторов. М., Гостехиздат, 1956.
2. С и м е о н о в С. В. Некоторые методы решения нелинейных задач механики деформируемого тела. ПММ, 1964, т. 28, вып. 3.
3. С и м е о н о в С. В. Замечания к статье «Некоторые методы решения нелинейных задач механики деформируемого тела». ПММ, 1964, т. 28, вып. 6.
4. К а н т о р о в и ч Л. В., А к и л о в Г. П. Функциональный анализ в нормированных пространствах. М., Физматгиз, 1959.
5. Д е м и д о в и ч Б. П., М а р о н И. А. Основы вычислительной математики. Изд. 2, М., Физматгиз, 1963.
6. О с т р о в с к и й А. М. Решение уравнений и систем уравнений. М., Изд-во иностр. лит., 1963.
7. Ф а д д е е в Д. К., Ф а д д е е в а В. Н. Вычислительные методы линейной алгебры, М., Физматгиз, 1960.

О СРЕДНЕКВАДРАТИЧНОЙ УСТОЙЧИВОСТИ СИСТЕМЫ СТОХАСТИЧЕСКИХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Г. Н. Мильштейн, Ю. М. Репин

(Свердловск)

В заметке рассматриваются объекты, поведение которых описывается решениями системы линейных дифференциальных уравнений со случайными коэффициентами. Каждое такое решение в рассматриваемом случае является диффузионным процессом. Для вторых моментов этого диффузионного процесса выводится система линейных дифференциальных уравнений с постоянными коэффициентами, которая, в частности, может быть применена для исследования на устойчивость в среднеквадратичном тривиального решения соответствующей системы стохастических дифференциальных уравнений.¹

1. Рассмотрим систему стохастических дифференциальных уравнений вида

$$dx_i = \sum_{j=1}^n a_{ij} x_j dt + \sum_{k=1}^m \left(\sum_{j=1}^n b_{ij}^k x_j \right) dw_k(t) \quad (1)$$

Здесь $w_k(t)$ ($k = 1, \dots, m$) — независимые процессы броуновского движения.

При заданных начальных условиях уравнение (1), как известно [1-4], определяет диффузионный процесс. Стационарная марковская переходная функция $P(t, x, \Gamma)$ этого процесса определяет на банаховом пространстве B всех борелевских ограниченных функций $f(x)$ ($x = (x_1, \dots, x_n)$) полугруппу линейных операторов T_t :

$$T_t f(x) = \int_X P(t, x, dy) f(y) \quad (2)$$

Все ограниченные функции $f(x)$, имеющие непрерывные ограниченные производные до второго порядка включительно, принадлежат области определения D_A инфини-

¹ При чтении корректуры авторам стала известна работа И. И. Гихмана [9], в которой другим способом получены уравнения для вторых моментов.