УДК 517.97

Г. Ш. Тамасян

ГРАДИЕНТНЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧИ КОШИ*)

Введение. Для решения задачи Коши в настоящее время известно множество численных методов, например метод последовательных приближений Пикара, метод Эйлера, метод Рунге–Кутта [1]. В данной работе решение задачи Коши сводится к безусловной минимизации соответствующего функционала. С учетом специфики строения функционала для поиска минимизирующей последовательности применяются градиентные методы. Рассматриваемые ниже алгоритмы относятся к прямым методам вариационного исчисления [2, 3].

Постановка задачи Коши. Пусть $T > 0$ – фиксированное число. Рассмотрим линейную неоднородную систему дифференциальных уравнений

$$\dot{x} = P(t)x + g(t), \tag{1}$$

в которой x – n-мерный вектор искомых функций; $P(t)$ – $(n \times n)$-матрица; $g(t)$ – n-мерная вектор-функция. Предположим, что элементы матрицы $P(t)$ и вектор-функции $g(t)$ являются функциями, определенными и непрерывными при $t \in [0, T]$.

Требуется среди всех решений системы (1) найти такое, которое будет удовлетворять начальному условию $x(0) = x_0$, где $x_0 \in \mathbb{R}^n$ – заданный вектор. С учетом указанных требований к системе (1) решение задачи Коши существует и единственно [1].

Постановка вариационной задачи. Произведем следующую замену переменных:

$$x(t) = x_0 + \int\limits_0^t z(\tau)\, d\tau, \tag{2}$$

где $z(t)$ – непрерывная вектор-функция при $t \geqslant 0$. Отметим, что $z(t) = \dot{x}(t)$.

Итак, решение задачи Коши сводится к поиску такой вектор-функции $z(t)$, которая удовлетворяет системе

$$z(t) = P(t)\left(x_0 + \int\limits_0^t z(\tau)\, d\tau\right) + g(t). \tag{3}$$

Далее для краткости записи будем писать $x(t)$ вместо $x_0 + \int\limits_0^t z(\tau)\, d\tau$ (см. (2)).

Тамасян Григорий Шаликович – кандидат физико-математических наук, доцент кафедры математической теории моделирования систем управления факультета прикладной математики–процессов управления Санкт-Петербургского государственного университета. Количество опубликованных работ: 17. Научные направления: негладкий анализ, вариационное исчисление, недифференцируемая оптимизация. E-mail: grigoriytamasjan@mail.ru, grigoriytamasyan@yandex.ru.

*) Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 09-01-00360).

Обозначим через $C_n[0,T]$ пространство непрерывных вещественных n-мерных вектор-функций на отрезке $[0,T]$. В этом пространстве введем норму

$$\|z\| = \sqrt{\int\limits_0^T (z(t), z(t))\, dt}.$$

Поиск решения системы (3) сведем к минимизации следующего функционала на всем пространстве $C_n[0,T]$.

Рассмотрим функционал и некоторые его свойства:

$$I(z) = \|z(t) - P(t)x(t) - g(t)\|^2 = \int\limits_0^T \Big(z(t) - P(t)x(t) - g(t), z(t) - P(t)x(t) - g(t)\Big)\, dt,$$

где $x(t) = x_0 + \int\limits_0^t z(\tau)\, d\tau$.

Заметим, что $I(z) \geqslant 0$ при любых $z \in C_n[0,T]$. Несложно также показать, что функционал $I(z)$ достигает минимального значения, равного нулю $I(z^*) = 0$, тогда и только тогда, когда $z^*(t)$ – решение задачи Коши (1) или (3).

Покажем, что $I(z)$ является строго выпуклым функционалом [1, 3], доказав прежде следующие вспомогательные утверждения.

Обозначим через $f(z) = z(t) - P(t)x(t) - g(t)$.

Утверждение 1. *Пусть $z_1(t)$ и $z_2(t)$ – произвольные элементы из пространства $C_n[0,T]$. Для того чтобы $f(z_1) = f(z_2)$ для всех $t \in [0,T]$, необходимо и достаточно, чтобы $z_1(t) = z_2(t)$.*

Д о к а з а т е л ь с т в о. Достаточность очевидна. Докажем необходимость. Пусть $f(z_1) = f(z_2)$, но $z_1(t) \neq z_2(t)$. Тогда найдется такая функция $\psi(t) \not\equiv 0$ из пространства $C_n[0,T]$, что $z_1(t) = z_2(t) + \psi(t)$.

Итак,

$$f(z_1) = f(z_2 + \psi) = z_2(t) + \psi(t) - P(t)\left[x_0 + \int\limits_0^t (z_2(\tau) + \psi(\tau))d\tau\right] - g(t) =$$
$$= f(z_2) + \psi(t) - P(t)\int\limits_0^t \psi(\tau)d\tau.$$

Так как по условию $f(z_1) = f(z_2)$, то

$$\psi(t) - P(t)\int\limits_0^t \psi(\tau)d\tau = 0.$$

Полученное выражение – однородное интегральное уравнение Вольтерра второго рода. При $t = 0$ имеем $\psi(0) = 0$, значит, единственным решением интегрального уравнения является $\psi(t) = 0$. Получили противоречие.

Утверждение 2. *Функционал $I(z)$ строго выпуклый.*

Доказательство. Требуется показать, что

$$\alpha I(z_1) + (1-\alpha)I(z_2) - I(\alpha z_1 + (1-\alpha)z_2) > 0$$

для всех $\alpha \in (0,1)$, $z_1 \neq z_2$, z_1, $z_2 \in C_n[0,T]$.

Действительно,

$$I(\alpha z_1 + (1-\alpha)z_2) = \|f(\alpha z_1 + (1-\alpha)z_2)\|^2 =$$

$$= \left\| \alpha z_1(\tau) + (1-\alpha)z_2(\tau) - P(t)\left[x_0 + \int_0^t (\alpha z_1(\tau) + (1-\alpha)z_2(\tau))d\tau\right] - g(t) \right\|^2 =$$

$$= \|\alpha f(z_1) + (1-\alpha)f(z_2)\|^2 =$$

$$= \alpha^2 \int_0^T (f(z_1), f(z_1))\,dt + 2\alpha(1-\alpha)\int_0^T (f(z_1), f(z_2))\,dt + (1-\alpha)^2 \int_0^T (f(z_2), f(z_2))\,dt =$$

$$= \alpha \int_0^T (f(z_1), f(z_1))\,dt + (1-\alpha)\int_0^T (f(z_2), f(z_2))\,dt + 2\alpha(1-\alpha)\int_0^T (f(z_1), f(z_2))\,dt -$$

$$- \alpha(1-\alpha)\int_0^T (f(z_1), f(z_1))\,dt - \alpha(1-\alpha)\int_0^T (f(z_2), f(z_2))\,dt =$$

$$= \alpha\|f(z_1)\|^2 + (1-\alpha)\|f(z_2)\|^2 - \alpha(1-\alpha)\int_0^T (f(z_1) - f(z_2), f(z_1) - f(z_2))\,dt.$$

Используя утверждение 1, имеем

$$\alpha I(z_1) + (1-\alpha)I(z_2) - I(\alpha z_1 + (1-\alpha)z_2) =$$

$$= \alpha(1-\alpha)\int_0^T (f(z_1) - f(z_2), f(z_1) - f(z_2))\,dt > 0$$

для всех $\alpha \in (0,1)$, $z_1 \neq z_2$.

Итак, решение задачи Коши будем искать среди функций, доставляющих минимальное значение (равное нулю) функционалу

$$I(z) = \int_0^T \Big(z(t) - P(t)x(t) - g(t), z(t) - P(t)x(t) - g(t)\Big)\,dt \longrightarrow \min_{z \in C_n[0,T]} \tag{4}$$

на всем пространстве $C_n[0,T]$.

Используя теорию точных штрафных функций, в [2] были получены необходимые условия минимума для поставленной проблемы.

Описание алгоритмов. В классическом вариационном исчислении наиболее распространенный поиск экстремалей функционала осуществляется либо из решений уравнений Эйлера [4], либо при помощи прямых методов [5, 6] типа Ритца, Галеркина, Канторовича и т. п. Описанные ниже алгоритмы являются аналогами градиентных методов [2, 7], которые применяются в теории оптимизации конечномерных пространств, а именно методы наискорейшего спуска и сопряженных направлений. Указанные методы (в конечномерных задачах) обладают достаточно хорошей скоростью сходимости минимизирующей последовательности [7–9]. К примеру, в задачах минимизации выпуклых квадратичных функций метод сопряженных направлений сходится за конечное число шагов, которое не превосходит размерности пространства. Различные варианты градиентных и прямых методов решения задач вариационного исчисления можно найти в работах Б. Т. Поляка [3], Л. В. Канторовича [10], С. Г. Михлина [6], В. Ф. Демьянова [2] и др.

Рассмотрим классическую вариацию вектор-функции $z(t)$:

$$z_\varepsilon(t) = z(t) + \varepsilon v(t), \tag{5}$$

где $\varepsilon > 0$, $v(t)$ – произвольная вектор-функция пространства $C_n[0,T]$.

Применяя вариацию (5) к функционалу (4), получим классическую вариацию функционала $I(z)$

$$I(z_\varepsilon) = I(z) + \varepsilon \int\limits_0^T \left(G(z,t), v(t)\right)\, dt + o(\varepsilon),$$

здесь $\dfrac{o(\varepsilon)}{\varepsilon} \xrightarrow[\varepsilon\downarrow 0]{} 0$, $G(z,t)$ – градиент Гато [2, 3, 5] функционала $I(z)$ в точке z имеет вид

$$G(z,t) = z(t) - P(t)x(t) - g(t) - \int\limits_t^T P^*(\tau)\big[z(\tau) - P(\tau)x(\tau) - g(\tau)\big]\, d\tau.$$

Через $P^*(t)$ обозначается транспонированная матрица $P(t)$.

Далее рассмотрим несколько разновидностей градиентного метода. Изучаемые алгоритмы носят итерационный характер. Это значит, что строится некоторая последовательность $z_k(t)$, $k = 0, 1, \ldots$, относительно которой можно утверждать, что она в той или иной мере сходится к решению задачи минимизации. Так как функционал (4) строго выпуклый, то равенство нулю нормы градиента в точке z^* (стационарная точка) является необходимым и достаточным условием минимума функционала [2, 8, 9].

Градиентный метод. Выбираем начальное приближение $z_0(t)$. Строим последовательность $\{z_k(t)\}$ по правилу

$$z_{k+1}(t) = z_k(t) - \gamma_k G(z_k,t), \quad \gamma_k > 0, \quad k = 0, 1, \ldots,$$

если $\|G(z_k,t)\| = \sqrt{\int\limits_0^T \Big(G(z_k,t), G(z_k,t)\Big)\, dt} \neq 0$, то шаг $\gamma_k > 0$ [8, 9] можно выбрать так, чтобы $I(z_{k+1}) < I(z_k)$.

Если $\|G(z_k,t)\| = 0$, то z_k – стационарная точка, и процесс построения минимизирующей последовательности прекращается.

Метод наискорейшего спуска (МНС). Пользуясь тем, что градиент $G(z,t)$ линейный относительно z, а подынтегральная функция функционала (4) квадратичная относительно z, можно модифицировать предыдущий метод. А именно, шаг $\gamma_k > 0$ будем выбирать из условия минимума функционала (4). Несложно показать, что таким шагом является величина

$$\gamma_{k+1} = \frac{\int\limits_0^T \Big(z_k(t) - P(t)x_k(t) - g(t), G(z_k,t) - P(t)\int\limits_0^t G(z_k,\tau)\,d\tau \Big) dt}{\int\limits_0^T \Big(G(z_k,t) - P(t)\int\limits_0^t G(z_k,\tau)\,d\tau, G(z_k,t) - P(t)\int\limits_0^t G(z_k,\tau)\,d\tau \Big) dt}.$$

Метод сопряженных градиентов (направлений) (МСГ). Пусть $z_0(t)$ – некоторое начальное приближение. Будем строить последовательность $\{z_k(t)\}$ по формулам

$$z_{k+1}(t) = z_k(t) - \gamma_k W_k(t), \quad k = 0, 1, \ldots, \tag{6}$$

где

$$W_0(t) = G(z_0,t), \quad W_k(t) = G(z_k,t) + \beta_k W_{k-1}(t), \quad k = 1, 2, \ldots .$$

Величина γ_k может определяться так же, как и в вышеуказанных методах, а β_k – по одной из формул

$$\beta_k = \frac{\int\limits_0^T \Big(G(z_k,t), G(z_k,t) - G(z_{k-1},t) \Big) dt}{\int\limits_0^T \Big(G(z_{k-1},t), G(z_{k-1},t) \Big) dt}, \tag{7}$$

$$\beta_k = \frac{\int\limits_0^T \Big(G(z_k,t), G(z_k,t) \Big) dt}{\int\limits_0^T \Big(G(z_{k-1},t), G(z_{k-1},t) \Big) dt}, \tag{8}$$

$$\beta_k = \frac{\int\limits_0^T \Big(G(z_k,t), G(z_k,t) - G(z_{k-1},t) \Big) dt}{\int\limits_0^T \Big(W_{k-1}, G(z_{k-1},t) \Big) dt}, \tag{9}$$

$$\beta_k = \frac{\int\limits_0^T \Big(G(z_k,t), G(z_k,t) \Big) dt}{\int\limits_0^T \Big(W_{k-1}, G(z_{k-1},t) \Big) dt}. \tag{10}$$

Как показывает практика, в каждом шаге алгоритма с неизбежностью накапливаются погрешности. Это может привести к тому, что векторы W_k перестают указывать направление убывания функционала, и релаксация метода может нарушиться. Для борьбы с таким явлением метод сопряженных направлений время от времени обновляют, полагая в (6) $\beta_k = 0$, т. е. осуществляют градиентный спуск.

Заметим, что величины, вычисленные по формулам (7)–(10), будут одинаковыми, если градиенты попарно ортогональные.

Пример. Пусть $T = 15$. Решим задачу Коши

$$\dot{x} = -x$$

при начальном условии $x(0) = 1$.

Найдем решение данной задачи $x^*(t) = e^{-t}$ методом последовательного приближения Пикара. На k-м шаге получаем первые k членов разложенного в ряд Тейлора «искомого» решения, а именно

$$x_1(t) = 1,\ x_2(t) = 1 - t,\ x_3(t) = 1 - t + \frac{t^2}{2},\ \dots\ , x_k(t) = \sum_{s=1}^{k} (-1)^s \frac{t^s}{s!}.$$

Очевидно, что $x_k(t) \xrightarrow[k\to\infty]{} x^*(t) = e^{-t}$.

В табл. 1–3 приведены результаты решения задачи Коши алгоритмами Пикара, МНС и МСГ на отрезке $[0, 15]$. Вычисления проводились в математическом пакете Mathcad. В качестве начального приближения в методе Пикара взято $x_0(t) = 1$, а в алгоритмах МНС и МСГ выбрали $z_0(t) = 1$. В табл. 1 представлены выборочные итерации алгоритма Пикара; в табл. 2 и 3 – первые 5 шагов соответствующих алгоритмов.

Таблица 1. **Результаты алгоритма Пикара**

k	$\|z^* - z_k\|$	$\|x^* - x_k\|$	$I(z_k)$
7	$1.168 \cdot 10^4$	$2.224 \cdot 10^4$	$3.38 \cdot 10^4$
14	$1.204 \cdot 10^5$	$1.205 \cdot 10^5$	$2.4 \cdot 10^5$
21	$3.4 \cdot 10^4$	$2.3 \cdot 10^4$	$5.7 \cdot 10^4$
28	947	486.9	1440
35	4.58	1.897	6.48
40	0.043	0.016	0.058

Таблица 2. **Результаты алгоритма МНС**

k	$\|z^* - z_k\|$	$\|x^* - x_k\|$	$I(z_k)$	$\|G(z_k)\|$
0	4.183	36.894	1635	824.327
1	1.328	2.597	11.527	20.506
2	0.732	1.043	1.911	7.66
3	0.678	1.011	1.594	8.05
4	0.613	0.828	1.34	6.09
5	0.6	0.823	1.14	6.48

Таблица 3. **Результаты алгоритма МСГ**

k	$\|z^* - z_k\|$	$\|x^* - x_k\|$	$I(z_k)$	$\|G(z_k)\|$	$\|W_k\|$
0	4.183	36.894	1635	824.327	824.327
1	1.328	2.597	11.527	20.506	20.512
2	0.73	1.06	1.854	5.938	6.17
3	0.34	0.283	0.21	1.32	1.35
4	0.096	0.067	0.014	0.305	0.313
5	0.028	0.023	0.00132	0.105	–

Из примера видно, что рассматриваемые методы являются релаксационными.

Заключение. Результаты численных экспериментов показали эффективность использованных методов для решения задачи Коши. Метод сопряженных градиентов более трудоемкий по сравнению с методами наискорейшего спуска и градиентным, однако построенная минимизирующая последовательность имеет более высокую скорость сходимости. Указанные алгоритмы могут быть применены и для решения нелинейных систем дифференциальных уравнений. Остаются открытыми вопросы о сходимости и скорости сходимости построенных минимизирующих последовательностей.

Автор благодарит В. Ф. Демьянова и С. К. Мышкова за ценные указания и внимание к работе.

Литература

1. *Матвеев Н. М.* Методы интегрирования обыкновенных дифференциальных уравнений. Изд. 4-е, испр. и доп. Минск: Вышэйшая школа, 1974. 768 с.

2. *Демьянов В. Ф.* Условия экстремума и вариационные задачи. М.: Высшая школа, 2005. 335 с.

3. *Поляк Б. Т.* Градиентные методы минимизации функционалов // Журн. вычисл. математики и мат. физики. 1963. Т. 3, № 4. С. 643–653.

4. *Гюнтер Н. М.* Курс вариационного исчисления. М.: Гостехиздат, 1941. 308 с.

5. *Вайнберг М. М.* Вариационный метод и метод монотонных операторов в теории нелинейных уравнений. М.: Наука, 1972. 416 с.

6. *Михлин С. Г.* Численная реализация вариационных методов. М.: Наука, 1966. 432 с.

7. *Карманов В. Г.* Математическое программирование. М.: Наука, 1986. 228 с.

8. *Пшеничный Б. Н., Данилин Ю. М.* Численные методы в экстремальных задачах. М.: Наука, 1975. 320 с.

9. *Васильев Ф. П.* Численные методы решения экстремальных задач. М.: Наука, 1980. 550 с.

10. *Канторович Л. В., Акилов Г. П.* Функциональный анализ. 2-е изд., перераб. М.: Наука, 1977. 741 с.

Статья рекомендована к печати В. Ф. Демьяновым.

Статья принята к печати 28 мая 2008 г.