

МЕТОДЫ ТИПА РУНГЕ-КУТТА ДЛЯ РЕШЕНИЯ НЕЛИНЕЙНЫХ УРАВНЕНИЙ В ГИЛЬБЕРТОВОМ ПРОСТРАНСТВЕ

И. ПЕТЕРСЕН,

кандидат физико-математических наук

Рассмотрим вопрос о приближенном решении уравнения

$$P(x) = 0, \quad (1)$$

где $P(x)$ нелинейный оператор из вещественного гильбертова пространства H в то же пространство. Из большого числа методов для решения уравнения (1) наиболее изучены методы типа Ньютона. Однако для практического использования на электронных вычислительных машинах эти методы, в случаях когда размерность пространства H больше 2—3, мало пригодны, потому что они требуют оперирования с $P'(x)$ и для сходимости мы должны иметь очень точные первые приближения. С точки зрения практики больше интереса представляют различные варианты метода «наискорейшего спуска», детально рассмотренные в работах М. Альтмана и Л. Кивистика (см. библиографию в [1]). Но и для этих методов надо вычислять $P'(x)$.

Ниже предлагается способ вывода алгоритмов для решения уравнения (1), при которых на каждом шагу придется вычислять лишь значения оператора $P(x)$. Кроме того, эти алгоритмы содержат параметр «шаг» метода, при достаточно малом значении которого можно всегда добиться сходимости метода (при ограниченности производных до требуемого порядка). Существенным ограничением остается лишь положительная (или отрицательная) определенность производной $P'(x)$. Идея этого способа заключается в рассмотрении, как предложено в работе М. К. Гавурина [2], вместо уравнения (1) задачи о решении дифференциального уравнения

$$x'(t) = -P(x(t)), \quad x(0) = x_0, \quad (2)$$

и в применении к этому уравнению методов приближенного решения дифференциальных уравнений.

Первая теорема связывает проблему решения уравнения (1) с задачей (2). Ее аналог для нахождения минимумов функционалов получен П. Росенблумом [3]. Аналогичная локальная теорема с более жесткими условиями имеется и в [2]. Теорема 2 устанавливает связь между приближенным решением задачи (2) и решением уравнения (1). Теоремы 3 и 4 являются примерами применения теоремы 2. Результаты, близкие в том или ином смысле к теореме 3, имеются в [1], [4] и [5]. Метод,

предложенный в теореме 4, и подобные методы высших порядков, являются, по-видимому, новыми и довольно многообещающими, особенно когда к ним присоединить некоторый алгоритм выбора h (вместе со знаком) на каждом шагу из условия приближенного минимума функции одной переменной $\|P(x_{n+1}(h))\|$, что легко осуществимо на электронных вычислительных машинах.

Теорема 1. Если оператор $P(x)$ дифференцируем в замкнутой сфере $S[x_0, r]$, где $r \geq M \|P(x_0)\|$, удовлетворяет в этой сфере условию Липшица и условию¹

$$(P'(x)y, y) \geq \frac{1}{M} \|y\|^2, \quad (M > 0), \quad (3)$$

при любом $y \in H$, то уравнение (1) имеет в сфере $S[x_0, r]$ единственное решение x^* . При этом дифференциальное уравнение (2) имеет для $t > 0$ решение $x(t) \in S(x_0, r)$, $\lim_{t \rightarrow +\infty} x(t) = x^*$

и имеют место оценки

$$\|P(x(t))\| \leq \|P(x_0)\| e^{-\frac{t}{M}}, \quad (4)$$

$$\|x(t) - x(t_1)\| \leq M \|P(x_0)\| (e^{-\frac{t_1}{M}} - e^{-\frac{t}{M}}) \quad (0 \leq t_1 < t), \quad (5)$$

$$\|x(t) - x^*\| \leq M \|P(x_0)\| e^{-\frac{t}{M}}. \quad (6)$$

Доказательство. Решение уравнения (2) существует для достаточно малых $t > 0$. Для таких t

$$x''(t) = -P'(x(t))x'(t) = P'(x(t))P(x(t))$$

и

$$\frac{d}{dt} (x'(t), x'(t)) = 2(x''(t), x'(t)) = -2(P'(x(t))P(x(t)), P(x(t))),$$

откуда на основании (3) получаем

$$\frac{d}{dt} \|x'(t)\|^2 \leq -\frac{2}{M} \|x'(t)\|^2$$

или

$$\frac{d}{dt} \ln \|x'(t)\| \leq -\frac{1}{M},$$

так что

$$\|x'(t)\| \leq \|x'(0)\| e^{-\frac{t}{M}},$$

т. е. (4) имеет место.

Дальше, при $t > t_1$

$$\|x(t) - x(t_1)\| = \left\| \int_{t_1}^t x'(t) dt \right\| \leq \int_{t_1}^t \|x'(t)\| dt \leq \|x'(0)\| \int_{t_1}^t e^{-\frac{t}{M}} dt =$$

¹ Аналогичная теорема имеет место при условии $(P'(x)y, y) \leq -\frac{1}{M} \|y\|^2$.

$$= M \| P(x_0) \| (e^{-\frac{t_1}{M}} - e^{-\frac{t}{M}}),$$

так что и (5) имеет место. Из последнего неравенства легко вывести существование $x(t)$, включение $x(t) \in S[x_0, r]$ для всех $t > 0$ и существование предела $\lim_{t \rightarrow \infty} x(t) = x^* \in S[x_0, r]$. При $t \rightarrow +\infty$ из (4) получим $P(x^*) = 0$ и из (5) неравенство (6).

Наконец имеем для любого $y \in S[x_0, r]$

$$(P(y) - P(x^*), y - x^*) = (P'(x^* + \Theta(y - x^*))(y - x^*), y - x^*) \geqslant \geqslant \frac{1}{M} \| y - x^* \|^2 \quad \text{с } 0 < \Theta < 1,$$

откуда следует

$$\| P(y) \| \geqslant \frac{1}{M} \| y - x^* \|,$$

что доказывает единственность решения x^* в сфере $S[x_0, r]$.

Теорема доказана.

Для сформулирования следующей теоремы вводим некоторые обозначения. Пусть $\rho(\delta)$ — полином степени $s - 2$ с положительными коэффициентами,

$$q(h, \delta) = e^{-\frac{h}{M}} + A h^s \rho(\delta), \tag{7}$$

$$r(h, \delta) = \frac{M(1 - e^{-\frac{h}{M}}) + h^s \rho(\delta)}{1 - q(h, \delta)} \delta, \tag{8}$$

где A и M положительные постоянные, и пусть для любого $\delta_0 > 0$ и $h > 0$ последовательность $\{\delta_n\}$ определена формулой

$$\delta_{n+1} = q(h, \delta_n) \delta_n \quad (n = 0, 1, \dots). \tag{9}$$

Теорема 2. Пусть оператор $P(x)$ дифференцируем в сфере $S[x_0, r(h, \delta_0)]$, где $\| P(x_0) \| \leqslant \delta_0$, и удовлетворяет в этой сфере условию (3) и условию Липшица с постоянной A . Пусть дальше h^* положительный корень уравнения

$$q(h, \delta_0) = 1, \tag{10}$$

h некоторое число, $0 < h < h^*$, и последовательность $\{x_n\}$ элементов пространства H построена рекуррентной формулой

$$x_{n+1} = F(x_n, h) \quad (n = 0, 1, \dots), \tag{11}$$

причем для всех n выполнено условие

$$\| F(x_n, h) - z_n(h) \| \leqslant h^s \rho(\| P(x_n) \|) \| P(x_n) \| \quad s \geqslant 2, \tag{12}$$

где $z_n(t)$ решение задачи

$$z'(t) = -P(z(t)), \quad z(0) = x_n. \tag{13}$$

Тогда $x_n \in S[x_0, r(h, \delta_0)]$ и последовательность $\{x_n\}$ сходится к единственному решению x^* уравнения (1) в сфере $S[x_0, r(h, \delta_0)]$ при любом таком h , что $0 < h < h^*$, со скоростью

$$\|x_{n+1} - x^*\| \leq [h^s p(\|P(x_n)\|) + Me^{-\frac{h}{M}}] \|P(x_n)\| \leq [h^s p(\delta_n) + Me^{-\frac{h}{M}}] \delta_n. \quad (14)$$

Доказательство. При любом фиксированном δ имеем $q(0, \delta) = 1$, $q'(0, \delta) < 0$ и $q''(h, \delta) > 0$ для $h > 0$, так что $q(h, \delta) < 1$ для всех h из $0 < h < \bar{h}$, если $q(\bar{h}, \delta) \leq 1$. По условию теоремы, следовательно, $q(h, \delta_0) < 1$ для $0 < h < h^*$. Согласно (9) имеем тогда $\delta_1 < \delta_0$. Так как $q(h, \delta)$ является при фиксированном h возрастающей функцией от δ , то $q(h, \delta_1) < 1$. Таким образом, по индукции получаем, что при $0 < h < h^*$

$$\delta_{n+1} < \delta_n, \quad q(h, \delta_n) < 1 \quad (n = 0, 1, \dots). \quad (15)$$

Так как $\delta_n > 0$ для всех n , то существует $\lim_{n \rightarrow \infty} \delta_n = \bar{\delta} \geq 0$. Если $\bar{\delta} > 0$, то из (9) мы получили бы $q(h, \bar{\delta}) = 1$ и вследствие $\delta_0 > \bar{\delta}$ имели бы $q(h, \delta_0) > 1$, что противоречит условию $h < h^*$.

Таким образом, при $0 < h < h^*$

$$\lim_{n \rightarrow \infty} \delta_n = 0. \quad (16)$$

Предполагаем теперь, что $\|P(x_n)\| \leq \delta_n$ и

$S[x_n, r(h, \delta_n)] \subset S[x_0, r(h, \delta_0)]$, и покажем, что тогда и

$$\|P(x_{n+1})\| \leq \delta_{n+1}, \quad (17)$$

$$S[x_{n+1}, r(h, \delta_{n+1})] \subset S[x_n, r(h, \delta_n)]. \quad (18)$$

Действительно, по (12) и (5) $(S[x_n, r(h, \delta_n)] \supset S[x_n, M \|P(x_n)\|])$

имеем

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq \|F(x_n, h) - z_n(h)\| + \|z_n(h) - x_n\| \leq \\ &\leq h^s p(\|P(x_n)\|) \|P(x_n)\| + M(1 - e^{-\frac{h}{M}}) \|P(x_n)\|, \end{aligned}$$

так что

$$\|x_{n+1} - x_n\| \leq r(h, \delta_n) [1 - q(h, \delta_n)]. \quad (19)$$

Следовательно, $x_{n+1} \in S[x_0, r(h, \delta_0)]$. По теореме 1 и по нашему предположению $z_n(h) \in S[x_0, r(h, \delta_0)]$.

Применение условия Липшица дает теперь

$$\|P(x_{n+1})\| \leq \|P(z_n(h))\| + A \|F(x_n, h) - z_n(h)\|$$

или, воспользуясь (4) и (12),

$$\|P(x_{n+1})\| \leq q(h, \delta_n) \delta_n = \delta_{n+1},$$

т. е. (17) имеет место. Так как $\delta_{n+1} < \delta_n$ и $p(\delta)$

возрастающая функция, то

$$r(h, \delta_{n+1}) = \frac{M(1 - e^{-\frac{h}{M}}) + h^s p(\delta_{n+1})}{1 - e^{-\frac{h}{M}} - A h^s p(\delta_{n+1})} \delta_{n+1} < \\ < \frac{M(1 - e^{-\frac{h}{M}}) + h^s p(\delta_n)}{1 - e^{-\frac{h}{M}} - A h^s p(\delta_n)} \delta_n q(h, \delta_n) = r(h, \delta_n) q(h, \delta_n),$$

что вместе с (19) даст

$$\|x_{n+1} - x_n\| \leq r(h, \delta_n) - r(h, \delta_{n+1}),$$

т. е. и (18) имеет место. Таким образом, (17) и (18) выполнены для всех $n = 0, 1, \dots$

Теперь можно для каждой сферы $S[x_n, r(h, \delta_n)]$ применять теорему 1. По этой теореме единственное решение x^* уравнения (1) в сфере $S[x_0, r(h, \delta_0)]$ находится в сфере $S[x_n, r(h, \delta_n)]$ и

$$\|x_{n+1} - x^*\| \leq \|F(x_n, h) - z_n(h)\| + \|z_n(h) - x^*\| \leq \\ \leq h^s p(\|P(x_n)\|) \|P(x_n)\| + M e^{-\frac{h}{M}} \|P(x_n)\|,$$

что вместе с (17) дает оценку (14) и в силу (16) доказывает теорему.

Теорема 2 позволяет, исходя из некоторого метода приближенного решения дифференциальных уравнений, который имеет вид (11) и удовлетворяет условию (12), построить и установить условия сходимости метода для решения уравнений вида (1). С точки зрения вычислительных машин, вероятно, представляют наибольший интерес такие методы, при которых в основу взяты методы Рунге-Кутты решения дифференциальных уравнений. Применяя к задаче (13) метод Рунге-Кутты первого порядка, т. е. метод Эйлера, получим для уравнения (1) метод «наискорейшего спуска». Условия сходимости при этом принимают следующий вид:

Теорема 3. Если оператор $P(x)$ непрерывно дифференцируем в замкнутой сфере $S[x_0, r]$, где

$$r \geq \frac{2M(1 - e^{-\frac{h}{M}}) + Ah^2}{2(1 - e^{-\frac{h}{M}}) - A^2h^2} \|P(x_0)\|, \quad (20)$$

удовлетворяет в этой сфере условиям (3) и, кроме того,

$$\|P'(x)\| \leq A, \quad (21)$$

$$0 < h \leq \frac{2M}{1 + A^2M^2}, \quad (22)$$

то последовательность

$$x_{n+1} = x_n - hP(x_n) \quad (23)$$

сходится к единственному в этой сфере решению x^* уравнения (1) со скоростью

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq \left(\frac{A}{2} h^2 + Me^{-\frac{h}{M}}\right) \|P(x_n)\| \leq \\ &\leq \left(\frac{A}{2} h^2 + Me^{-\frac{h}{M}}\right) \left(\frac{A^2}{2} h^2 + e^{-\frac{h}{M}}\right)^n \|P(x_0)\|. \end{aligned} \quad (24)$$

Доказательство. По формуле Тэйлора можно решение $z_n(t)$ задачи (13) представить в виде

$$z_n(t) = x_n - tP(x_n) + \int_0^1 (1-\tau) P'(z_n(\tau t)) P(z_n(\tau t)) t^2 d\tau,$$

а по (23) в данном случае

$$F(x_n, t) = x_n - tP(x_n),$$

так что на основании (4)

$$\|F(x_n, h) - z_n(h)\| \leq h^2 \frac{A}{2} \|P(x_n)\|.$$

Следовательно (12) выполнено с $s = 2$ и $p(\delta) = \frac{A}{2}$.

Дальше имеем

$$\begin{aligned} q(h, \delta_0) - 1 &= e^{-\frac{h}{M}} + \frac{A^2}{2} h^2 - 1 < 1 - \frac{h}{M} + \frac{h^2}{2M^2} + \frac{A^2}{2} h^2 - 1 = \\ &= \frac{h}{M} \left(\frac{1 + A^2 M^2}{2M} h - 1 \right), \end{aligned}$$

так что

$$h^* > \frac{2M}{1 + A^2 M^2}.$$

В данном случае $q(h, \delta)$ не зависит от δ и поэтому

$$\delta_n = q^n(h, \delta_0) \delta_0.$$

Остается применить теорему 2.

Дальше рассмотрим один вариант метода Рунге-Кутты второго порядка.

Теорема 4. Если оператор $P(x)$ два раза непрерывно дифференцируем в замкнутой сфере $S[x_0, r]$, где

$$r \geq \frac{12M \left(1 - e^{-\frac{h}{M}}\right) + h^3(5AB\|P(x_0)\| + 2A^2)}{12\left(1 - e^{-\frac{h}{M}}\right) - Ah^3(5AB\|P(x_0)\| + 2A^2)} \|P(x_0)\| \quad (25)$$

удовлетворяет в этой сфере условиям (3) и (21), а в сфере $S[x_0, r + h\|P(x_0)\|]$ условию

$$\|P''(x)\| \leq B \tag{26}$$

и h удовлетворяет неравенству

$$0 < h \leq \frac{2\sqrt{3}M}{2,118\,117 + \sqrt{5}ABM^3\|P(x_0)\| + 2A^3M^3 - 2}, \tag{27}$$

то последовательность

$$x_{n+1} = x_n - \frac{h}{2} [P(x_n) + P(x_n - hP(x_n))] \tag{28}$$

сходится к единственному в этой сфере решению x^* уравнения (1) со скоростью

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq \left[h^3 \left(\frac{5B}{12} \|P(x_n)\| + \frac{A^2}{6} \right) + M e^{-\frac{h}{M}} \right] \|P(x_n)\| \leq \\ &\leq \left[h^3 \left(\frac{5B}{12} \delta_n + \frac{A^2}{6} \right) + M e^{-\frac{h}{M}} \right] \delta_n, \end{aligned} \tag{29}$$

где

$$\begin{aligned} \delta_0 &\geq \|P(x_0)\|, \\ \delta_k &= \frac{5h^3AB}{12} \delta_{k-1}^2 + \left(e^{-\frac{h}{M}} + \frac{h^3A^3}{6} \right) \delta_{k-1}. \end{aligned}$$

Доказательство. В данном случае

$$\begin{aligned} z_n(t) &= x_n - t(P(x_n) + \frac{t^2}{2} P'(x_n)P(x_n) - \\ &- \frac{t^3}{2} \int_0^1 (1-\tau)^2 [P''(z_n(\tau t))P^2(z_n(\tau t)) + P'^2(z_n(\tau t))P(z_n(\tau t))] d\tau \end{aligned}$$

и

$$\begin{aligned} F(x_n, t) &= x_n - \frac{t}{2} [P(x_n) + P(x_n - tP(x_n)) + \\ &+ t^2 \int_0^1 (1-\tau) P''(x_n - \tau tP(x_n)) P^2(x_n) d\tau]. \end{aligned}$$

Следовательно,

$$\|F(x_n, h) - z_n(h)\| \leq h^3 \left[\frac{5B}{12} \|P(x_n)\| + \frac{A^2}{6} \right] \|P(x_n)\|,$$

т. е. условие (12) выполнено с $s = 3$ и

$$\rho(\delta) = \frac{5B}{12} \delta + \frac{A^2}{6}.$$

Дальше имеем ([6], теор. 143)

$$\begin{aligned}
 q(h, \delta_0) - 1 &= e^{-\frac{h}{M}} + A h^3 \left(\frac{5B}{12} \delta_0 + \frac{A^2}{6} \right) - 1 < \\
 < 1 - \frac{h}{M} + \frac{h^2}{2M^2} - \frac{h^3}{6M^3} + \frac{h^4}{24M^4} + \left(\frac{5AB}{12} \delta_0 + \frac{A^3}{6} \right) h^3 - 1 = \\
 &= \frac{h}{3M} \left[\left(\frac{h}{2M} \right)^3 + (5ABM^3 \delta_0 + 2A^3M^3 - 2) \left(\frac{h}{2M} \right)^2 + 3 \frac{h}{2M} - 3 \right] = \frac{h}{3M} \varphi \left(\frac{h}{2M} \right),
 \end{aligned}$$

где

$$\varphi(t) = t^3 + kt^2 + 3t - 3$$

и

$$k = 5ABM^3 \delta_0 + 2A^3M^3 - 2 \geq 0$$

в силу $AM \geq 1$. Очевидно $q(h, \delta_0) - 1 < 0$ для $0 < h \leq \frac{2M}{N}$, если N является верхней границей положительных корней полинома

$$\psi(t) = -t^3 \varphi \left(\frac{1}{t} \right) = 3t^3 - 3t^2 - kt - 1.$$

Убедимся, что можно принять

$$N = c + \sqrt{\frac{k}{3}},$$

где $c = 1,222895$ положительный корень уравнения $3t^3 - 3t^2 - 1 = 0$.

Действительно, так как

$$\psi(N) = (2c - 1)k + 3c(3c - 2) \sqrt{\frac{k}{3}},$$

$$\psi'(N) = 3c(3c - 2) + 2k + 6(3c - 1) \sqrt{\frac{k}{3}},$$

$$\psi''(N) = 6(3c - 1) + 18 \sqrt{\frac{k}{3}},$$

$$\psi'''(N) = 18,$$

и $c > 1$, то $\psi(t)$ и ее производные в точке $t = N$ неотрицательны для всех $k \geq 0$ и поэтому $\psi(t)$ не имеет корней, превосходящих N . Уравнение (10), таким образом, не имеет положительных корней при h , удовлетворяющем неравенству (27). Для завершения доказательства теперь достаточно применить теорему 2.

Аналогично можно указать условия и скорость сходимости методов для решения уравнения (1), полученных из методов Рунге-Кутты более высоких порядков. Однако формулировки этих условий будут слишком громоздкими, чтобы приводить их здесь.

ЛИТЕРАТУРА

1. Л. А. Кивистик, Об итерационных методах в пространстве Гильберта, автореферат, Тарту, 1961.
2. М. К. Гавурин, Нелинейные функциональные уравнения и непрерывные аналоги итерационных методов, Изв. высших уч. зав., сер. матем. № 5, 1958.

3. P. C. Rosenbloom, The method of steepest descent, Numerical Analysis, vol. VI, Proceedings of the Sixth Symposium in Applied Mathematics of the American Mathematical Society, 1956.
4. Е. И. Линьков, О методе наискорейшего спуска для нелинейных уравнений, Уч. записки Московск. обл. пед. инст. ХСVI, вып. VI, 1960.
5. М. М. Вайнберг, О сходимости процесса наискорейшего спуска для нелинейных уравнений, Сибирский мат. журн. II, № 2, 1961.
6. Г. Г. Харди, Д. Е. Литтльвуд, Г. Поля, Неравенства, М., 1948.

*Институт кибернетики
Академии наук Эстонской ССР*

Поступила в редакцию
19. VI 1962

RUNGE-KUTTA TÕUPI MEETODID MITTELINEAARSETE VÕRRANDITE LAHENDAMISEKS HILBERTI RUUMIS

I. Petersen,

füüsika-matemaatikateaduste kandidaat

Resümee

Artiklis tuletatakse seos võrrandi (1) ja diferentsiaalvõrrandi (2) lahendite vahel ja käsitletakse võrrandi (1) ligikaudse lahendamise algoritme, mis vastavad diferentsiaalvõrrandi (2) lahendamisele Runge-Kutta meetodil.

*Eesti NSV Teaduste Akadeemia
Küberneetika Instituut*

Saabus toimetusse
19. VI 1962

RUNGE-KUTTA-METHODEN ZUR LÖSUNG NICHTLINEARER GLEICHUNGEN IM HILBERTSCHEN RAUM

I. Petersen

Zusammenfassung

In der Abhandlung wird die Beziehung zwischen den Lösungen der Gleichung (1) und der Differentialgleichung (2) abgeleitet und die Algorithmen zur angenäherten Lösung der Gleichung (1) behandelt, die der Lösung der Differentialgleichung (2) mittels der Runge-Kutta-Methode entspricht.

*Institut für Kybernetik
der Akademie der Wissenschaften der Estnischen SSR*

Eingegangen
am 19. Juni 1962