12167

сообщения
объединенного
института
ядерных
исследований
дубна



P5 - 12167

В.Ц.Банчев

# О МИНИМИЗАЦИИ ФУНКЦИЙ С ПОМОЩЬЮ ОБОБЩЕННЫХ МЕТОДОВ ИНТЕГРИРОВАНИЯ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

1979

P5 - 12167

В.Ц.Банчев

# О МИНИМИЗАЦИИ ФУНКЦИЙ С ПОМОЩЬЮ ОБОБЩЕННЫХ МЕТОДОВ ИНТЕГРИРОВАНИЯ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Банчев В.Ц. P5 - 12167

О минимизации функций с помощью обобщенных методов интегрирования обыкновенных дифференциальных уравнений

Показано, что использование обобщенных методов интегрирования обыкновенных дифференциальных уравнений позволяет получить новые методы минимизации без ограничений со сверхлинейной скоростью сходимости, которые не используют процедуры одномерного поиска, минимизируют квадратичные функции за один шаг и удачно объединяют преимущества метода наискорейшего спуска и метода Ньютона. Величина шага контролируется путем оценки локальной ошибки усечения.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1979

Banchev V.Ts. P5 - 12167

On Function Minimization by Means of Generalized Methods for Integration of Ordinary Differential Equations

It is shown that the use of generalized methods for the integration of ordinary differential equations allows one to formulate new methods for unconstrained minimization with superlinear rate of convergence, which do not perform a unidimensional search, minimize quadratic functions in one step and successfully combine the merits of Newton's method and that of steepest descent. A step size is controlled by estimating the local truncation error.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Communicotion of the Joint Institute for Nuclear Research. Dubna 1979



Задача минимизации без ограничений функции $N$ переменных $f(x)$, $f: D \subset R^N \longrightarrow R^1$, в работе формулируется как задача нахождения асимптотически устойчивого по Ляпунову решения системы $N$ обыкновенных дифференциальных уравнений

$$x'(t) = -Q^{-1}(x)\,g(x), \qquad x(0) = x^o, \tag{1}$$

где $g^T(x) = (\partial f/\partial x_1, \dots, \partial f/\partial x_N)$, а $Q(x)$ – положительно определенная симметрическая матрица. Система (1) является непрерывным аналогом метода наискорейшего спуска /1/. Как показано в /1-4/, такая формулировка задачи минимизации без ограничений позволяет: 1) определить кривую наискорейшего спуска; 2) найти естественное обобщение метода Ньютона на более широкий класс минимизируемых функций, для которых уже не обязаны выполняться ограничения на знакоопределенность матрицы вторых производных $H(x)$ (матрицы Гессе); 3) непосредственно связать собственные числа и собственные векторы матрицы Гессе $H(x)$ с процессом минимизации. Кроме того, в /4/ отмечается, что общим недостатком большинства ныне используемых алгоритмов является резкое ухудшение их свойств при минимизации "овражных" функций, т.е. таких, у которых матрица Гессе $H(x)$ имеет большое спектральное число обусловленности $k = |M|/|m|$, где $M$, $m$ – наибольшее и наименьшее по модулю ненулевые собственные числа $H(x)$. Система (1) для таких функций является жесткой (в случае $Q(x) = H(x)$ задача решения (1) некорректна /4/), и, следовательно, для эффективного нахождения ее решения нужно применять методы интегрирования жестких систем. В связи с этим в /4/ предлагается использовать принцип квазистационарности производных совместно с методом, разработанным независимо в /2,3,5,6/, а в /3/ подчеркивается необходимость модификации существующих численных методов для эффективного интегрирования (1).

Цель нестоящей работы – показать целесообразность применения обобщенных методов интегрирования обыкновенных дифференциальных уравнений /9/ для нахождения асимптотически устойчивых точек покоя системы (1), что, по нашему мнению, является дальнейшим развитием методов из /2–4/ , а также указать на возможность более полного использования преимуществ вышеприведенной формулировки задачи минимизации без ограничений. В частности, в работе показано, что это позволяет получить методы минимизации со сверхлинейной скоростью сходимости, которые не используют процедуры одномерного поиска, минимизируют квадратичные функции за один шаг и удачно объединяют преимущества метода наискорейшего спуска и метода Ньютона. При этом величина шага контролируется путем оценки локальной ошибки усечения.

Пусть $f(x) \in C^2(x \in D)$ и пусть $x^*$ –локальный минимизатор /7/ для $f(x)$, т.е. $g(x^*) = 0$ и

$$
\begin{aligned}
&g(x) = H(x^*)(x - x^*) + \psi(x), \\
&\|\psi(x)\| = o(\|x - x^*\|) \quad (x \to x^*, \; x \in D).
\end{aligned} \tag{2}
$$

В дальнейшем будем предполагать, что $H(x^*)$ – положительно определенная матрица и $Q(x) = I$ , т.е. ограничимся рассмотрением уравнения

$$x'(t) = -g(x), \qquad x(0) = x^{\circ}. \tag{3}$$

Пусть, кроме того $x^{\circ} \in \bar{S}(x^*, r) = \{x \in R^N \mid \|x - x^*\| \le r\}$, $r > 0$, $\bar{S}(x^*, r) \subset D_0 \subset D$ , где $D_0$ – область притяжения положения равновесия $x^*$ для (3). Тогда /8/ $x^*$ является асимптотически устойчивой точкой покоя системы (3) и

$$\|x(t) - x^*\| \le L \|x^{\circ} - x^*\| \exp(-\gamma t), \quad t \ge 0, \tag{4}$$

где $L$ и $\gamma$ – некоторые положительные постоянные, не зависящие от выбора $x^{\circ}$ .

Для $0 = t_0 < t_1 < t_2 < \ldots < +\infty$ обозначим через $x^k$ $(x^{\circ} \in \bar{S}(x^*, r))$ приближения к значениям $x(t_k)$, $k = 0, 1, \ldots$, $x(t_0) = x^{\circ}$, решения задачи Коши (3), полученные с помощью какого-нибудь численного метода. Пусть $e_k$ – локальная ошибка усечения этого метода на $k$-м шагу интегрирования (3) и $h_k = t_{k+1} - t_k$ $(k=0,1,\ldots)$. Тогда, если

$$\|e_k\| \le \beta_k \|x^{k+1} - x^k\|, \quad 0 \le \beta_k < 1/2, \tag{5}$$

$$\delta = \max_k \delta_k = \max_k [L \exp(-\gamma h_k) + \beta_k]/(1 - \beta_k) < 1, \tag{6}$$



используя (4), получаем

$$\|x^{k+1} - x^*\| \le \delta_k \|x^k - x^*\|, \quad k = 0, 1, \ldots, \tag{7}$$

и $\lim_{k \to \infty} x^k = x^*$. . Очевидно, при заданных $L$ , $\gamma$ и $\beta_k$ всегда можно найти такое $h > 0$ , что при $h_k > h$ (6) выполняется. Скорость сходимости $x^k$ к $x^*$ теоретически можно сделать любой, если одновременно $h_k \to \infty$ и $\beta_k \to 0$ , например, если $\beta_k = O(\exp(-\gamma h_k))$.

Дадим теперь более эффективную оценку ошибки $\|x^{k+1} - x^*\|$. Пусть $h < h_k \le h_{max}$ $(k=0,1,\ldots)$, $h_k = h_{max}$ при $k \ge m$ . Определим отображение $\mathcal{G}: D_0 \subset R^N \to R^{N}$ следующим образом

$$\mathcal{G} u^{\circ} = u(h_{max}), \quad u^{\circ} \in D_0, \tag{8}$$

$$u'(t) = -H(x^*)(u - x^*), \quad u(0) = u^{\circ}. \tag{9}$$

Очевидно, $\mathcal{G}$ существует, определено единственным образом и

$$
\begin{aligned}
&\|\mathcal{G} p - \mathcal{G} q\| \le \alpha \|p - q\| \;\; \forall p, q \in D_0, \\
&\alpha = R \exp(-\mu h_{max}),
\end{aligned} \tag{10}
$$

где $R$ и $\mu$ – некоторые положительные постоянные. При этом $\min_i \lambda_i(H(x^*)) > \mu$ и $x^*$ – асимптотически устойчивая точка покоя (9). В дальнейшем будем предполагать значение $h_{max}$ таким, что $\alpha < 1$ . Тогда $\mathcal{G}$ является сжатием на $D_0$ и $\mathcal{G}\bar{S}(x^*, r) \subset \bar{S}(x^*, r)$. Положим $u^{k+1} = \mathcal{G} u^k$, $u^{\circ} \in \bar{S}(x^*, r)$, $\epsilon_k = \|x^{k+1} - \mathcal{G} x^k\|$, $k = 0, 1, \ldots$ . Согласно теореме об аппроксимативных сжатиях /7, с. 379/ имеем (в силу (7) $x^k \in \bar{S}(x^*, r)$)

$$\|x^{k+1} - x^*\| < [1/(1 - \alpha)][\alpha \|x^{k+1} - x^k\| + \epsilon_k], \quad k = 0, 1, \ldots, \tag{11}$$

$$\|x^{k+1} - x^*\| < \|u^{k+1} - x^*\| + \sum_{j=0}^{k} \alpha^{k-j} \epsilon_j + \alpha^{k+1} \|u^{\circ} - x^{\circ}\|, \quad k = 0, 1, \ldots, \tag{12}$$

где

$$\epsilon_k \le \|e_k\| + \left\| \int_0^{h_k} \exp[-(h_k - s) H(x^*)] \psi(\bar{x}(s)) \, ds \right\|, \quad k \ge m, \tag{13}$$

$\bar{x}(t)$ – решение (3) с $\bar{x}(0) = x^k$ и $\lim_{k \to \infty} \epsilon_k = 0$ в силу (2)–(7).

Остановимся теперь на применении обобщенных методов интегрирования обыкновенных дифференциальных уравнений /9/ для решения задачи Коши (3). Используя преобразование

$$z(t) = T(t) - \int_0^t T(s) \, a(s) \, ds$$

с матрицей $T(t)$, удовлетворяющей уравнению $T' = -TA(t)$, $T(0) = I$, сведем решение задачи (3) к интегрированию системы

$$z' = f(t, x(z)) = T(t)[-g(x(z)) - A(t)x(z) - a(t)], \quad z(0) = x^{\circ}, \qquad (14)$$

обычными численными методами (например, методами Рунге-Кутта или Адамса). При этом произвольные пока $N \times N$ – матрица $A(t)$ и $N$-мерный вектор $a(t)$ интегрируемы в смысле Римана в любом конечном интервале $0 \le t \le t_f < +\infty$. Заметим, что в (14)

$$x = T^{-1}(t)[z + \int_0^t T(s)\,a(s)\,ds],$$

где $(T^{-1})' = A(t)T^{-1}$, $T^{-1}(0) = I$, является решением (3). Как и выше, через $z^k$ $(k = 0, 1, \ldots)$ обозначим приближения к значениям $z(t_k)$, полученные при численном интегрировании (14), $0 = t_0 < t_1 < t_2 < \ldots < +\infty$. Пусть $z^k$ соответствуют приближения $x^k$ к значениям $x(t_k)$. Выберем $A(t) = -H(x^k)$ и $a(t) = -g(x^k) + H(x^k)x^k$ при $t \in [t_k, t_{k+1}]$ $(k = 0, 1, \ldots)$. Кроме того, будем предполагать, что матрица Гессе $H(x)$ удовлетворяет условию Гельдера

$$\| H(p) - H(q) \| \le c \| p - q \|^{\nu} \;\; \forall \, p, q \in D_0,$$

$0 < \nu \le 1$, $c > 0$. Тогда

$$f(t, x) = T(t)\rho(x), \;\; \rho(x) = -g(x) + g(x^k) + H(x^k)(x - x^k), \qquad (15)$$

$$\| \rho(x) \| = o(\| x - x^k \|), \;\; \| \partial\rho/\partial x \| \le c \| x - x^k \|^{\nu} \qquad (16)$$

для $t \in [t_k, t_{k+1}]$, $x(t)$ – решение уравнения (3) с $x(t_k) = x^k$. Система (14) нежесткая в интервале $[t_k, t_{k+1}]$, если $\| x - x^k \| \le \delta$ для $t \in [t_k, t_{k+1}]$, где $\delta \ge 0$ – достаточно малое число /9/. Пусть $x^{\circ} \in \bar{S}(x^*, r)$ и пусть $x^k$ удовлетворяют условиям (5) – (7). Так как в силу (4)

$$\| x - x^k \| < (L + 1) \| x^k - x^* \|,$$

то величина $h_k = t_{k+1} - t_k$ в интервале $[t_k, t_{k+1}]$, в котором $z(t) \approx \mathrm{const}$ и собственные значения матрицы Якоби системы (14) $\partial f/\partial z = T(t)[\partial\rho/\partial x]T^{-1}(t)$ малы, зависит от близости $x^k$ к $x^*$. Отсюда, согласно (7), следует, что обобщенные методы позволяют существенно увеличить шаг интегрирования в достаточно малой окрестности $x^*$, т.е. при достаточно большом $t$. Более того, покажем, что одновременно с увеличением шага $h_k$ они позволяют уменьшить значение $\beta_k$ в (5), если выполнено естественное требование, чтобы на $k$-м шагу локальная ошибка усечения $r_k$ выбранного нами метода для интегрирования (14) удовлетворяла неравенству

$$\| r_k \| < \| \int_{t_k}^{t_{k+1}} f(s, x(s))\,ds \|,$$

где $x(t)$ – решение уравнения (3) с $x(t_k) = x^k$. В самом деле, тогда в силу (16) для локальной ошибки усечения $e_k$ соответствующего обобщенного метода интегрирования (3) на том же шагу имеем оценку /9/

$$\| e_k \| < \epsilon K \| x^{k+1} - x^k \| \, \| T^{-1}(t_{k+1}) \| \int_{t_k}^{t_{k+1}} \| T(s) \| \, ds, \qquad (17)$$

где константа $k > 0$ такая, что

$$\sup_{t \in [t_k, t_{k+1}]} \| x(t) - x^k \| / \| x^{k+1} - x^k \| \le K,$$

$x(t)$ – решение уравнения (3) с $x(t_k) = x^k$ и $\epsilon \to 0$ при $\| x^{k+1} - x^k \| \to 0$.

Наконец, заметим, что с помощью обобщенных методов интегрирования обыкновенных дифференциальных уравнений, в силу (14), точку минимума строго выпуклых квадратичных функций можно найти за один шаг (см. также /2-4/).

Проведенные численные эксперименты с помощью программы 'GERUN' /9/, в которой реализован один из обобщенных методов Рунге-Кутта, подтвердили полученные выше результаты. При этом выбор $\beta_k = O(\exp(-\gamma h_k))$ в (6) оказался не только возможным, но и необходимым для того, чтобы избежать ненужного увеличения шага интегрирования, при котором становится существенным влияние ошибок округления.

## ЛИТЕРАТУРА


1. Гаврурин М.К. Изв. Вузов. Математика, 1958, №5, с.18-31.
2. *Botsaris C.A., Jacobson D.H. J.Math.Anal.Appls, 1976, v.54, p. 217-229.*
3. *Botsaris C.A. J. Math. Anal. Appls., 1978, v. 63, p. 177-198.*
4. Ракитский Ю.В., Устинов С.М., Черноруцкий И.Г. Численные методы решения жестких систем обыкновенных дифференциальных уравнений. "Наука", Л., 1977.
5. Ракитский Ю.В. ДАН СССР, 1972, т.207, №4, с.793-795.
6. Ракитский Ю.В. Труды ЛПИ, 1973, №332, с.88-97.
7. Ортега Дж., Рейнболдт В. Итерационные методы решения нелинейных систем уравнений со многими неизвестными. "Мир", М., 1975.

8. Демидович Б.П. Лекции по математической теории устойчивости. "Наука", М., 1967.
9. Банчев В.Ц. ОИЯИ, Р5-12174,Дубна, 1979.

Рукопись поступила в издательский отдел
10 января 1979 года.