EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 31. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1982, NR. 1

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 31
ФИЗИКА * МАТЕМАТИКА. 1982, № 1

УДК 517.948

А. РООЗЕ

ОБ ИТЕРАЦИОННОМ МЕТОДЕ РЕШЕНИЯ НЕЛИНЕЙНЫХ УРАВНЕНИЙ С ПРИМЕНЕНИЕМ ПАРАЛЛЕЛЬНОЙ АППРОКСИМАЦИИ ОБРАТНОГО ОПЕРАТОРА

(Представил А. Хумал)

Предлагается общий подход к конструированию итерационных методов, в которых применяется параллельная аппроксимация обратного оператора. Эти методы предназначены для решения нелинейных уравнений на параллельно работающих процессорах. Выведены условия сходимости одного конкретного метода и проведено его сравнение с методом, использующим последовательную аппроксимацию обратного оператора.

1. Как известно [1], в рамках методов можно выделить два главных подхода к распараллеливанию вычислений. Первый — использование т. н. естественного параллелизма существующих последовательных методов, второй — построение новых методов и алгоритмов специально для параллельных вычислительных устройств. При этом естественный параллелизм методов может быть усилен искусственными построениями.

Обзор работ по параллельным вычислениям в линейной алгебре [1–3] говорит о том, что это, по-видимому, наиболее развитый раздел теории и практики параллельных вычислений. Публикаций же, посвященных параллельным методам решения нелинейных уравнений, значительно меньше. Отметим здесь работы [4–8].

В настоящей статье предлагается прием, который позволит строить методы, состоящие из двух параллельных ветвей. Одна из них предназначена для построения приближений к решению нелинейных уравнений, а другая — для аппроксимации обратных операторов, которые использует первая ветвь. При необходимости в каждой из ветвей можно также распараллеливать расчеты, применяя методы для линейной алгебры.

2. Итак, требуется найти решение нелинейного уравнения

$$P(x)=0, \tag{1}$$

где $P(x)$ действует из одного банахова пространства X в другое Y. Перейдем от уравнения (1) к т. н. расширенной системе

$$\begin{aligned} &P(x)=0,\\ &P'(x)A-I=0, \end{aligned} \tag{2}$$

где $P'(x)$ — производная Фреше оператора $P(x)$, I — единичный оператор. Ясно, что, решив систему уравнений (2) относительно x и A, получим решение $\bar{x}$ исходного уравнения (1).

Приложение идей методов Гаусса—Зейделя и Якоби [9] (§ 7.4) к расширенной системе уравнений (2) дает возможность генерировать итерационные методы, использующие аппроксимации обратного оператора $[P'(x)]^{-1}$, и решать ими уравнение (1).

А именно, применив к системе (2) идею метода Якоби, причем использовав на каждом шаге итерации один шаг метода Ньютона для каждого из уравнений (2), получим итерационный метод с параллельной аппроксимацией обратного оператора в виде:

$$x^{n+1}=x^n-A^nP(x^n), \tag{3а}$$

$$A^{n+1}=A^n(2I-P'(x^n)A^n), \quad n=0,\ 1,\ \ldots, \tag{3б}$$

где x^0, A^0 — начальные приближения.

Конечно, решая уравнения (2), можно применять на каждой итерации метода Якоби и несколько шагов ньютоновского процесса, а также и другие методы, использующие аппроксимации обратных операторов. Можно строить и релаксационные варианты этих методов. В результате каждый случай даст свои новые параллельные методы для решения уравнения (1).

Если же применим к системе (2) идею метода Гаусса—Зейделя, использовав на каждой его итерации один шаг метода Ньютона для каждого из уравнений (2), то получим итерационный метод с последовательной аппроксимацией обратного оператора, впервые предложенный и изученный С. Ульмом в [10]. В этой же работе им доказана квадратичная скорость сходимости метода. Аналогичный результат получен для метода с последовательной аппроксимацией псевдообратного оператора [11].

Ниже установим свойства сходимости метода (3), который по терминологии Ортеги—Рейнбольта [9] можно назвать и одношаговым блочным методом Якоби—Ньютона для решения расширенной системы уравнений (2). Полученные оценки позволят далее определить условия, при которых метод с параллельной аппроксимацей обратного оператора (3) обеспечит меньшее реальное время счета, чем метод с последовательной аппроксимацией.

Теорема. *Пусть*

1° *уравнение* (1) *имеет решение* $\bar{x}$ *и существует* $\bar{A}=[P'(\bar{x})]^{-1}$;

2° *в сфере* $\|x-\bar{x}\|\leqslant r_0$ *справедлива оценка* $\|P''(x)\|\leqslant L$;

3° $\|P'(\bar{x})\|\leqslant C, \quad \|[P'(\bar{x})]^{-1}\|\leqslant B$;

4° $h=\max\{Kr_0, G\}<1$, *где* $r_0=\max\{\|\bar{x}-x^0\|,\ \|\bar{A}-A^0\|\}$,

$$K=C+1{,}5BL+1{,}5Lr_0, \qquad G=Cr_0+B^2L+2BLr_0+Lr_0^2.$$

Тогда последовательности приближений $\{x^n\}_0^\infty$ *и* $\{A^n\}_0^\infty$ *сходятся соответственно к* $\bar{x}$ *и* $\bar{A}$, *причем найдутся константы* γ_1, γ_2 *такие, при которых справедливы оценки*

$$\|\bar{x}-x^n\|\leqslant h^{c_n}r_0, \tag{4}$$

$$\|\bar{A}-A^n\|\leqslant h^{g_n}r_0, \tag{5}$$

где $c_n=\gamma_1t_1^n+\gamma_2t_2^n-2$, $g_n=c_{n-1}+1$, $c_{-1}=-1$,

$$t_1=0{,}5(1+\sqrt{5})\approx 1{,}618, \qquad t_2=0{,}5(1-\sqrt{5})\approx -0{,}618, \qquad n=0,\ 1,\ \ldots .$$

Доказательство. На основании (3а) и формулы Тейлора имеем

$$\bar{x}-x^{n+1}=\bar{x}-x^n+A^n(P(x^n)-P(\bar{x}))=\bar{x}-x^n-A^n[P'(x^n)(\bar{x}-x^n)+$$

$$+\int_0^1 P''(x^n+t(\bar{x}-x^n))(\bar{x}-x^n)^2(1-t)\,dt]=[I-A^nP'(x^n)](\bar{x}-x^n)-$$

$$-A^n\int_0^1 P''(x^n+t(\bar{x}-x^n))(\bar{x}-x^n)^2(1-t)\,dt.$$

Поскольку

$$I-A^nP'(x^n)=\bar{A}P'(\bar{x})-A^nP'(x^n)=(\bar{A}-A^n)P'(\bar{x})+A^n(P'(\bar{x})-P'(x^n))$$

и

$$\|A^n\|\leqslant\|\bar{A}\|+\|\bar{A}-A^n\|\leqslant B+\|\bar{A}-A^n\|,$$

то, используя условия 2° и 3° теоремы, получим

$$\|\bar{x}-x^{n+1}\|\leqslant[C\|\bar{A}-A^n\|+(B+\|\bar{A}-A^n\|)L\|\bar{x}-x^n\|]\|\bar{x}-x^n\|+$$
$$+0{,}5(B+\|\bar{A}-A^n\|)L\|\bar{x}-x^n\|^2=$$
$$=C\|\bar{x}-x^n\|\|\bar{A}-A^n\|+1{,}5BL\|\bar{x}-x^n\|^2+1{,}5L\|\bar{x}-x^n\|^2\|\bar{A}-A^n\|. \qquad (6)$$

С другой стороны, на основании (3б)

$$\bar{A}-A^{n+1}=\bar{A}-A^n(2I-P'(x^n)A^n)=\bar{A}-A^n(I+P'(\bar{x})\bar{A}-P'(x^n)A^n)=$$
$$=\bar{A}-A^n[I+P'(\bar{x})(\bar{A}-A^n)+(P'(\bar{x})-P'(x^n))A^n]=$$
$$=[I-A^nP'(\bar{x})](\bar{A}-A^n)-A^n(P'(\bar{x})-P'(x^n))A^n=$$
$$=(\bar{A}-A^n)P'(\bar{x})(\bar{A}-A^n)-A^n(P'(\bar{x})-P'(x^n))A^n.$$

Следовательно:

$$\|\bar{A}-A^{n+1}\|\leqslant C\|\bar{A}-A^n\|^2+\|P'(\bar{x})-P'(x^n)\|\|A^n\|^2\leqslant$$
$$\leqslant C\|\bar{A}-A^n\|^2+B^2L\|\bar{x}-x^n\|+2BL\|\bar{x}-x^n\|\|\bar{A}-A^n\|+$$
$$+L\|\bar{x}-x^n\|\|\bar{A}-A^n\|^2. \qquad (7)$$

Докажем справедливость оценок (4), (5), предположив сперва, что $c_0\geqslant 0$, $c_1\geqslant 0$. Для этого покажем, что если до некоторого $n\geqslant 2$ можно писать

$$c_n=c_{n-1}+g_{n-1}+1=c_{n-1}+c_{n-2}+2, \qquad (8)$$
$$g_n=c_{n-1}+1, \qquad (9)$$

то эти формулы верны и при $n+1$. А именно, используя (6) и (7), получим

$$\|\bar{x}-x^{n+1}\|\leqslant C(h^{c_n}r_0)h^{g_n}r_0+1{,}5BLh^{2c_n}r_0^2+1{,}5L(h^{2c_n}r_0^2)h^{g_n}r_0=$$
$$=(C+1{,}5BLh^{c_n-g_n}+1{,}5Lh^{c_n}r_0)h^{c_n+g_n}r_0^2<$$
$$<(Kr_0)h^{c_n+g_n}r_0\leqslant h^{c_n+g_n+1}r_0=h^{c_{n+1}}r_0;$$

$$\|\bar{A}-A^{n+1}\|\leqslant Ch^{2g_n}r_0^2+B^2Lh^{c_n}r_0+2BL(h^{c_n}r_0)h^{g_n}r_0+L(h^{c_n}r_0)h^{2g_n}r_0^2=$$
$$=(Ch^{2g_n-c_n}r_0+B^2L+2BLh^{g_n}r_0+Lh^{2g_n}r_0^2)h^{c_n}r_0<$$
$$<Gh^{c_n}r_0\leqslant h^{c_n+1}r_0=h^{g_{n+1}}r_0.$$

Далее нетрудно проверить, что

$$c_n=\gamma_1t_1^n+\gamma_2t_2^n-2$$

есть общее решение разностного уравнения

$$c_n=c_{n-1}+c_{n-2}+2,$$

причем t_1, t_2 — наибольший и наименьший корни характеристического уравнения $t^2 - t - 1 = 0$ соответственно. На основе (6), (7) и условия 4° теоремы легко выписать оценки

$$\|\bar{x} - x^1\| \leqslant hr_0, \quad \|\bar{x} - x^2\| \leqslant h^3 r_0,$$

из коих видно, что формула (8) верна при $n = 2$. Теперь константы γ_1, γ_2 можно определить из условий

$$c_0 = \gamma_1 + \gamma_2 - 2 = 0, \tag{10}$$

$$c_1 = \gamma_1 t_1 + \gamma_2 t_2 - 2 = 1. \tag{11}$$

Ввиду того, что $t_2 - t_1 \neq 0$, система линейных уравнений (10)—(11) имеет единственное решение. Тем самым доказана справедливость оценок (4). Для доказательства правильности оценок (5) выпишем на основе (6), (7) и условия 4° теоремы оценки

$$\|\bar{A} - A^1\| \leqslant hr_0, \quad \|\bar{A} - A^2\| \leqslant h^2 r_0,$$

из коих явствует, что соотношение (9) верно при $n = 2$ и, согласно приведенным выше рассуждениям, оценки (5) остаются в силе и при $n > 2$. Теорема доказана.

3. Сравним теперь время сходимости методов с последовательной и параллельной аппроксимацией обратного оператора (3).

С. Ульмом доказано [10], что если

$$h_0 r_0 = \max\{K, G_0\} r_0 < 1,$$

где $G_0 = C + B^2LK + 2BLKr_0 + LKr_0^2$, а K, C, B, L, r_0 имеют тот же смысл, что и в теореме данной работы, то для приближений $\{x^k\}_0^\infty$, получаемых методом с последовательной аппроксимацией обратного оператора, справедливы оценки

$$\|\bar{x} - x^k\| \leqslant (h_0 r_0)^{2^k - 1} r_0, \quad k = 0,\ 1,\ \ldots, \tag{12}$$

где x^0 — начальное приближение. Предположим, что существует положительная константа $Q < 1$ такая, что $h \leqslant Q$ и $h_0 r_0 \leqslant Q$. Тогда вместо оценок (4) и (12) можем воспользоваться оценками

$$\|\bar{x} - x^n\| \leqslant Q^{c_n} r_0, \quad n = 0,\ 1,\ \ldots, \tag{13}$$

где $c_n = \gamma_1 t_1^n + \gamma_2 t_2^n - 2$, $t_1 \approx 1{,}618$, $t_2 \approx -0{,}618$, $\gamma_1 \approx 1{,}894$, $\gamma_2 \approx 0{,}106$, и

$$\|\bar{x} - x^k\| \leqslant Q^{2^k - 1} r_0, \quad k = 0,\ 1,\ \ldots. \tag{14}$$

Далее, при достаточно больших n и k (напр., n, $k > 6$) можем для оценки скорости сходимости последовательностей приближений использовать приближенные неравенства:

$$\|\bar{x} - x^n\| \lessapprox Q^{1{,}894(1{,}618)^n} r_0, \tag{15}$$

$$\|\bar{x} - x^k\| \lessapprox Q^{2^k} r_0. \tag{16}$$

Теперь по оценкам (15) и (16) видно, что методами с последовательной и параллельной аппроксимацией обратного оператора будут получены приблизительно одинаковые результаты (в смысле величины нормы $\|\cdot\|$) при выполнении равенства

$$2^k = 1{,}894\,(1{,}618)^n. \tag{17}$$

Далее, из неравенства (17) получаем *

* На практике переменная k принимает, конечно, только целочисленные значения. Однако ввиду того, что наши рассуждения в этой части работы носят приблизитель-

$$k = (\ln 1{,}894 + n \ln 1{,}618) / \ln 2 \approx 0{,}921 + 0{,}694n. \qquad (18)$$

Пусть один шаг итерационного метода с последовательной аппроксимацией обратного оператора реализуется за время $\tau_1 + \tau_2$, где τ_1 — время вычисления очередной аппроксимации обратного оператора и τ_2 — время вычисления очередного приближения к искомому решению $\bar{x}$. Тогда, при однотипности двух параллельных процессоров, которые реализуют метод (3), и процессора, который работает по методу с последовательной аппроксимацией обратного оператора [10], один шаг итерационного метода с параллельной аппроксимацией обратного оператора (3) займет время $\max\{\tau_1, \tau_2\}$. Для конкретности примем далее $\tau_1 = \max\{\tau_1, \tau_2\}$. Общая продолжительность времени, за которое сравниваемыми методами получим приблизительно одинаковые результаты, будет тогда $T_s = k(\tau_1 + \tau_2)$ и $T_p = n\tau_1$, где индексы s и p означают последовательный и параллельный метод ссответственно. Проведение счета в двух параллельных ветвях по методу (3) будет оправдано только в том случае, если

$$T_s - T_p > 0. \qquad (19)$$

Следовательно, должно выполняться неравенство:

$$k(\tau_1 + \tau_2) - n\tau_1 = (0{,}921 + 0{,}694n)(\tau_1 + \tau_2) - n\tau_1 = (0{,}921 - 0{,}306n)\tau_1 + \\ + (0{,}921 + 0{,}694n)\tau_2 > 0. \qquad (20)$$

Отсюда приходим к выводу, что неравенство (20) выполняется, если при достаточно больших n и k

$$\tau_2/\tau_1 > (0{,}306n - 0{,}921)/(0{,}694n + 0{,}921). \qquad (21)$$

Несложно проверить, что если $\tau_2/\tau_1 \geqslant 0{,}44$, то неравенство (19) выполняется при всех $n = 1, 2, \ldots$. Для конкретных значений n можно эту оценку ослабить. Например, для $n = 9$ неравенство (19) выполняется, если $\tau_2/\tau_1 \geqslant 0{,}25$. Наибольшее ускорение счета в параллельных ветвях достигается, конечно, при $\tau_1 = \tau_2$.

4. Пример. Для системы

$$112x_2 + 16x_2^2 + 7/(x_1+1)^2 - (244 + 16x_2^2)x_1 - 4 = 0,$$

$$48(x_3 + x_1) + 8(x_3 - x_1)^2 + 3/(x_2+1)^2 - [96 + 8(x_3 - x_1)^2]x_2 - 2 = 0,$$

$$80(1 + x_2) + 16(1 - x_2)^2 + 5/(x_3+1)^2 - [160 + 16(1 - x_2)^2]x_3 - 4 = 0$$

сравнили последовательный [10] и параллельный (3) методы путем имитации работы последнего на ЭВМ ЕС 1040. Результаты одного прогона задачи приведены в таблице:

Метод	Число шагов	Общее время счета, *мкс*	τ_2/τ_1	Начальное приближение	Найденное решение
Паралл.	7	6999,55	0,52	$x^0 = \begin{bmatrix} 0{,}1 \\ 0{,}3 \\ 0{,}4 \end{bmatrix}$	$\bar{x} \approx \begin{bmatrix} 0{,}2704 \\ 0{,}5106 \\ 0{,}7466 \end{bmatrix}$
Послед.	6	8879,42	—		

ный характер, мы без ущерба для конечного результата можем допустить в выкладках и нецелочисленные значения переменной k. Здесь же отметим, что приближенная функция (18) уже при $n > 4$ дает значения переменной k, которые достаточно хорошо совпадают со значениями, вычисленными на основе исходного уравнения $2^k - 1 = c_n$. Это подтверждают непосредственные расчеты.

При вычислении принималось $A^0 = [P'(x^0)]^{-1}$. Условие окончания счета $\max\limits_{1 \leqslant i \leqslant 3} |x_i^{n+1} - x_i^n| \leqslant 0{,}0001$.

Автор признателен А.-М. Мантсик и И. Талуру за проведение численных экспериментов.

ЛИТЕРАТУРА

1. Фаддеева В. Н., Фаддеев Д. К., Кибернетика, № 6, 28—40 (1977).
2. Heller, D., SIAM Rev., **20**, № 4, 740—777 (1978).
3. Conrad, V., Wallach, Y., IEEE Trans. Comput., **C-26**, № 9, 838—846 (1977).
4. Shedler, G. S., Communs ACM, **10**, № 5, 286—291 (1967).
5. Miranker, W. L., IBM J. Res. Develop., № 5, 297—301 (1969).
6. Miellou, J.-C., Compt. rend. Acad. sci., Ser. A, № 278, 957—960 (1974).
7. Baudet, G. M., J. Assoc. Comput. Mach., **25**, № 2, 226—244 (1978).
8. Casulli, V., Trigiante, D., Calcolo, **15**, № 2, 147—160 (1978).
9. Ortega, J. M., Rheinboldt, W. C., Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York—London, 1970.
10. Ульм С., Изв. АН ЭССР, Физ. Матем., **16**, № 4, 403—411 (1967).
11. Ваарманн О., Изв. АН ЭССР, Физ. Матем., **24**, № 2, 239—242 (1975).

Специальное научно-производственное объединение «Алгоритм»

Поступила в редакцию
23/IX 1981

A. ROOSE

PÖÖRDOPERAATORI PARALLEELSET APROKSIMEERIMIST SISALDAVAST ITERATSIOONIMEETODIST

Artiklis on esitatud üldine võte pöördmaatriksi aproksimeerimist sisaldavate iteratsioonimeetodite konstrueerimiseks. On uuritud ühe konkreetse meetodi koonduvusomadusi, mis võimaldab pöördoperaatorit aproksimeerida samaaegselt lähendi leidmisega mittelineaarse võrrandi otsitavale lahendile, ning võrreldud selle meetodi ja ühe pöördmaatriksi järkjärgulist aproksimeerimist sisaldava iteratsioonimeetodi koonduvusaegu.

A. ROOSE

ON ITERATIVE METHOD WITH PARALLEL APPROXIMATION OF THE INVERSE OPERATOR

In [10] S. Ulm first considered iterative methods with successive approximation of the inverse operator for solving nonlinear equation

$$P(x) = 0, \tag{1}$$

where P is an operator from one Banach space to another. Advent of the parallel computer systems requires development of new algorithms or revision of the old ones to allow an effective execution on these computer systems.

In this paper we present an iterative method with parallel approximation of the inverse operator

$$x^{n+1} = x^n - A^n P(x^n), \tag{2a}$$

$$A^{n+1} = A^n (2I - P'(x^n) A^n), \quad n = 0, 1, \dots, \tag{2b}$$

where x^0, A^0 are initial approximations to $\bar{x}$ and $\bar{A} = [P'(\bar{x})]^{-1}$ respectively ($\bar{x}$ is exact solution of (1)), $P'(x)$ is a Fréchet derivative of $P(x)$ and I is the identity operator. One can notice that branches (2a) and (2b) are executable in parallel.

The convergence properties of method (2) are given which show that the sequence $\{x^n\}_0^\infty$ converges to the exact solution of equation (1) with approximate convergence rate factor 1,618. The convergence times of the method (2) and that of [10] are being compared and the preference conditions for these methods are being suggested. The main points of the given comparison are illustrated by actual measurements obtained from simulation runs of both methods on a computer.