EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 32. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1983, NR. 4

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 32
ФИЗИКА * МАТЕМАТИКА. 1983, № 4

https://doi.org/10.3176/phys.math.1983.4.04

УДК 517.988:519.615

Т. МЕРЕССОО

ОБ ОДНОМ МЕТОДЕ ПОРЯДКА $1+\sqrt{2}$ РЕШЕНИЯ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

(Представил Н. Алумяэ)

1. Рассмотрим систему нелинейных уравнений

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

где $H=(H_1, \ldots, H_m)$, $\beta=(\beta_1, \ldots, \beta_m)$, а $x=(x_1, \ldots, x_n)$ определяется как решение нелинейных зависящих от вектора параметров β задач

$$F_i(x_i, \beta) \to \min_x, \quad x_i \in \Gamma_i(\beta), \tag{2}$$

где F_i — критерий эффективности i-й вспомогательной задачи, $\Gamma_i(\beta) \subset R^{n_i}$ — ее допустимая область $(i=1, \ldots, n)$★. Считая, что задачи (2) имеют решения $x_i=x_i(\beta)$ $(i=1, \ldots, n)$, исследуем в дальнейшем существование и нахождение решения системы (1), т. е. такого вектора $\beta^* \in R^m$, что

$$H(x(\beta^*), \beta^*)=0. \tag{3}$$

На практике часто невозможно найти аналитических выражений для функций $x_i(\beta)$ $(i=1, \ldots, n)$, можно вычислить только их значения при фиксированных векторах параметров β.

Поэтому здесь удобно применять методы решения систем нелинейных уравнений, использующие соответственно только значения функций $H_j(x(\beta), \beta)$ $(j=1, \ldots, m)$. Это методы простой итерации, секущих, Стеффенсена и другие [1, 2]. Кроме того, вычисление значений $x_i(\beta)$ из вспомогательных задач (2) обычно связано с большими затратами времени. Поэтому целесообразно для решения системы (1) применять такие методы, которые обладают высокими порядками сходимости и тем самым требуют меньшего числа итераций, а следовательно, меньшего числа расчетов значений функций, чем методы более низкого порядка. Известно, что порядки сходимости названных выше методов не превосходят 2 [2, 3].

2. Для получения сходимости более высокого порядка рассмотрим следующий метод решения уравнения $J(\beta)=0$:

$$\begin{gathered}\beta_{k+1}=\mu_k - J(2\mu_k-\beta_k;\ \beta_k)^{-1} J(\mu_k), \\ \mu_k=\beta_k - J(2\mu_{k-1}-\beta_{k-1};\ \beta_{k-1})^{-1} J(\beta_k).\end{gathered} \tag{4}$$

Здесь $J(\beta; \mu)$ — разделенная разность первого порядка [4] нелинейного оператора $J(\beta): E_1 \to E_2$, где E_1, E_2 — в общем случае банаховы пространства, $\{\mu_k\}$ — последовательность вспомогательных точек [5]. Этот метод эффективен в том смысле, что для его реализации требуются на каждом шаге итерации (за исключением первого) вычисления лишь двух значений оператора и одной первой разделенной разности, а асимптоти-

★ Решение систем уравнений в аналогичной постановке исследовано в [1, 2].

ческий порядок сходимости [6] в условиях доказываемой теоремы не ниже $1+\sqrt{2}$.

Теорема. *Пусть* $\beta_0 \in E_1$, $S = \{\beta \in E_1 \mid \|\beta-\beta_0\| \leqslant \varrho\}$ *и на множестве* S *выполнены следующие условия:*

1° *оператор* $J(\beta)$ *дифференцируем (по Фреше);*

2° *разделенная разность* $J(\beta;\mu)$ *удовлетворяет условию* $\|J(\beta;\mu)-J(\mu;\omega)\| \leqslant L/2(\|\beta-\mu\|+\|\mu-\omega\|)$;

3° $\|J(\beta;\mu)^{-1}\| \leqslant \lambda$;

4° $\delta = \lambda^2 L\alpha(2+1/2\alpha)\|J(\beta_0)\| < 1$, *где* $\alpha = \lambda^2 L(5/2\|J(\beta_0)\| + \|J(\mu_0)\|)$ (μ_0 — *начальная вспомогательная точка, удовлетворяющая условиям* $\|\mu_0-\beta_0\| \leqslant \lambda\|J(\beta_0)\|$, $\|J(\mu_0)\| \leqslant \alpha\|J(\beta_0)\|$). *Тогда, если* $\varrho \geqslant \max\{r=\lambda d^{-1}\times (1+\alpha)T_0(\delta);\ 2\lambda d^{-1}T_0(\delta)\}$, *где* $T_k(\delta)=\sum_{i=k}^{\infty}\delta^{2^i}$, $d=2\lambda^2 L\alpha$, *то уравнение* $J(\beta)=0$ *имеет в* S *решение* β^*, $\|\beta_0-\beta^*\| \leqslant r$, *к которому сходится последовательность* $\{\beta_k\}$, *определяемая методом (4), и справедлива оценка*

$$\|\beta_k-\beta^*\| \leqslant \lambda d^{-1}(1+\alpha)T_k(\delta).$$

При этом асимптотический порядок сходимости метода (4) равен $\tau = 1+\sqrt{2}$.

Доказательство. Обозначим $A_k = J(2\mu_k-\beta_k;\beta_k)^{-1}$ и предположим временно (а ниже это будет доказано), что существуют $\gamma_k, \tilde{\gamma}_k$ ($k=0, 1, 2\ldots$) такие. что $\|I-J'(\beta_k)A_k\| \leqslant \gamma_k$, $\|I-J'(\beta_k)A_{k-1}\| \leqslant \tilde{\gamma}_k$, где $\beta_{-1}=\beta_0$, $\mu_{-1}=\mu_0$. Воспользуемся методикой Б. Т. Поляка [7]. Поскольку согласно определению разделенной разности первого порядка [4] из предположения 2° следует, что $J'(\beta)$ удовлетворяет условию Липшица с константой L, то, используя формулу

$$J(\beta+\mu) = J(\beta) + \int_0^1 J'(\beta+t\mu)\,\mu\,dt,$$

получаем

$$\|J(\beta_{k+1})\| = \|J(\mu_k - A_kJ(\mu_k))\| = \|(I-J'(\beta_k)A_k)J(\mu_k) + J'(\beta_k)A_kJ(\mu_k) -$$

$$- J'(\mu_k)A_kJ(\mu_k) + J'(\mu_k)A_kJ(\mu_k) - \int_0^1 J'(\mu_k - tA_kJ(\mu_k))A_kJ(\mu_k)\,dt\| =$$

$$= \|(I-J'(\beta_k)A_k)J(\mu_k) + (J'(\beta_k)-J'(\mu_k))A_kJ(\mu_k) + \int_0^1 [J'(\mu_k)-J'(\mu_k -$$

$$- tA_kJ(\mu_k))]A_kJ(\mu_k)\,dt\| \leqslant \gamma_k\|J(\mu_k)\| + \lambda L\|J(\mu_k)\|\,\|\beta_k-\mu_k\| +$$

$$+ 1/2\lambda^2 L\|J(\mu_k)\|^2 \leqslant \gamma_k\|J(\mu_k)\| + \lambda^2 L\|J(\mu_k)\|\,\|J(\beta_k)\| +$$

$$+ 1/2\lambda^2 L\|J(\mu_k)\|^2 = (\gamma_k + \lambda^2 L\|J(\beta_k)\| + 1/2\lambda^2 L\|J(\mu_k)\|)\|J(\mu_k)\|. \tag{5}$$

Аналогично

$$\|J(\mu_k)\| = \|J(\beta_k - A_{k-1}J(\beta_k))\| = \|(I-J'(\beta_k)A_{k-1})J(\beta_k) +$$

$$+ J'(\beta_k)A_{k-1}J(\beta_k) - \int_0^1 J'(\beta_k - tA_{k-1}J(\beta_k))A_{k-1}J(\beta_k)\,dt\| =$$

$$= \|(I-J'(\beta_k)A_{k-1})J(\beta_k) + \int_0^1 [J'(\beta_k) - J'(\beta_k - tA_{k-1}J(\beta_k))]\times \tag{6}$$

$$\times A_{k-1}J(\beta_k)\,dt\| \leqslant \tilde{\gamma}_k\|J(\beta_k)\| + 1/2\lambda^2 L\|J(\beta_k)\|^2 = (\tilde{\gamma}_k + 1/2\lambda^2 L\|J(\beta_k)\|)\|J(\beta_k)\|.$$

Согласно определению разделенной разности первого порядка [4] справедливы оценки

$$\|(I-J'(\beta_k)A_k)J(\mu_k)\|\leqslant\|(J(2\mu_k-\beta_k;\ \beta_k)-J(\beta_k;\ \beta_k))J(2\mu_k-\beta_k;\ \beta_k)^{-1}\times$$
$$\times J(\mu_k)\|\leqslant\|J(2\mu_k-\beta_k;\ \beta_k)-J(\beta_k;\ \beta_k)\|\ \|J(2\mu_k-\beta_k;\ \beta_k)^{-1}J(\mu_k)\|\leqslant$$
$$\leqslant L/2\|2\mu_k-2\beta_k\|\lambda\|J(\mu_k)\|\leqslant\lambda^2L\|J(\beta_k)\|\ \|J(\mu_k)\|;$$
$$\|(I-J'(\beta_k)A_{k-1})J(\beta_k)\|\leqslant\|(J(2\mu_{k-1}-\beta_{k-1};\ \beta_{k-1})-J'(\beta_k))J(2\mu_{k-1}-$$
$$-\beta_{k-1};\ \beta_{k-1})^{-1}J(\beta_k)\|\leqslant(\|J(2\mu_{k-1}-\beta_{k-1};\ \beta_{k-1})-J(\beta_{k-1};\ \beta_k)\|+$$
$$+\|J(\beta_{k-1};\ \beta_k)-J'(\beta_k)\|)\|J(2\mu_{k-1}-\beta_{k-1};\ \beta_{k-1})^{-1}J(\beta_k)\|\leqslant L(\|\mu_{k-1}-\beta_{k-1}\|+$$
$$+\|\beta_{k-1}-\beta_k\|)\lambda\|J(\beta_k)\|\leqslant\lambda^2L(2\|J(\beta_{k-1})\|+\|J(\mu_{k-1})\|)\|J(\beta_k)\|.$$

Поэтому можно принять

$$\gamma_k=\lambda^2L\|J(\beta_k)\|, \tag{7}$$

$$\tilde{\gamma}_k=\lambda^2L(2\|J(\beta_{k-1})\|+\|J(\mu_{k-1})\|). \tag{8}$$

Используя (7) и (8), из неравенств (5) и (6) получаем

$$\|J(\beta_{k+1})\|\leqslant\lambda^4L^2[2\|J(\beta_k)\|+1/2\|J(\mu_k)\|]\ [2\|J(\beta_{k-1})\|+\|J(\mu_{k-1})\|+$$
$$+1/2\|J(\beta_k)\|]\|J(\beta_k)\|\leqslant\lambda^4L^2[2+1/2\lambda^2L(2\|J(\beta_{k-1})\|+\|J(\mu_{k-1})\|+$$
$$+1/2\|J(\beta_k)\|)]\ [2\|J(\beta_{k-1})\|+\|J(\mu_{k-1})\|+1/2\|J(\beta_k)\|]\|J(\beta_k)\|^2. \tag{9}$$

Поскольку по предположению 4° $\delta<1$, то из (5) получаем

$$\|J(\beta_1)\|\leqslant(\lambda^2L\|J(\beta_0)\|+\lambda^2L\|J(\beta_0)\|+1/2\lambda^2L\alpha\|J(\beta_0)\|)\alpha\|J(\beta_0)\|=$$
$$=\lambda^2L(2+1/2\alpha)\|J(\beta_0)\|\alpha\|J(\beta_0)\|=\delta\|J(\beta_0)\|\leqslant\|J(\beta_0)\|.$$

Отсюда, используя (5) — (8), получаем

$$\|J(\mu_1)\|\leqslant[\lambda^2L(2\|J(\beta_0)\|+\|J(\mu_0)\|)+1/2\lambda^2L\|J(\beta_1)\|]\|J(\beta_1)\|\leqslant\alpha\|J(\beta_1)\|\leqslant$$
$$\leqslant\alpha\lambda^2L(2\|J(\beta_0)\|+1/2\alpha\|J(\beta_0)\|)\|J(\mu_0)\|=\delta\|J(\mu_0)\|\leqslant\|J(\mu_0)\|.$$

Так, продолжая, получаем оценки

$$\|J(\beta_k)\|\leqslant\|J(\beta_{k-1})\|,\quad \|J(\mu_k)\|\leqslant\|J(\mu_{k-1})\|,\quad \|J(\mu_k)\|\leqslant\alpha\|J(\beta_k)\|$$

для всех k.

Учитывая это и обозначая $d_k=\lambda^2L\alpha[2+1/2\lambda^2L(2\|J(\beta_{k-1})\|+\|J(\mu_{k-1})\|+1/2\|J(\beta_k)\|)]$, можно заметить, что $d_k\leqslant d_{k-1}\leqslant\ldots\leqslant d_0$. Поэтому из неравенства (9) получаем

$$\|J(\beta_{k+1})\|\leqslant d_k\|J(\beta_k)\|^2\leqslant\frac{1}{2\lambda^2L\alpha}(d_k\|J(\beta_k)\|)^2\leqslant$$

$$\leqslant\frac{1}{2\lambda^2L\alpha}(d_0\|J(\beta_0)\|)^{2^{k+1}}=d^{-1}\delta^{2^{k+1}}, \tag{10}$$

где $d=2\lambda^2L\alpha$. Далее,

$$\|\beta_{k+1}-\beta_k\|\leqslant\lambda(\|J(\beta_k)\|+\|J(\mu_k)\|)\leqslant\lambda(1+\alpha)\|J(\beta_k)\|\leqslant\lambda(1+\alpha)d^{-1}\delta^{2^k},$$

следовательно, для $m\geqslant k$

$$\|\beta_m-\beta_k\|\leqslant\sum_{i=k}^{m-1}\|\beta_{i+1}-\beta_i\|\leqslant\lambda d^{-1}(1+\alpha)(T_k(\delta)-T_m(\delta)):$$

Таким образом, последовательность $\{\beta_k\}$ фундаментальная, и существует $\beta^* = \lim_{k\to\infty} \beta_k$, $J(\beta^*) = 0$, причем для любого k

$$\|\beta_k - \beta^*\| \leqslant \lambda d^{-1}(1+\alpha) T_k(\delta).$$

Наконец, мы имели право пользоваться условиями теоремы, поскольку

$$\|\beta_k - \beta_0\| \leqslant \lambda d^{-1}(1+\alpha)\,(T_0(\delta) - T_k(\delta)) \leqslant r \leqslant \varrho;$$

$$\|\mu_k - \beta_0\| \leqslant \|\mu_k - \beta_k\| + \|\beta_k - \beta_0\| \leqslant \lambda \|J(\beta_k)\| + \|\beta_k - \beta_0\| \leqslant \lambda d^{-1}\delta^{2^k} +$$
$$+ r\left(1 - \frac{T_k(\delta)}{T_0(\delta)}\right) \leqslant r + T_k(\delta)\left(\lambda d^{-1} - \frac{r}{T_0(\delta)}\right) \leqslant r \leqslant \varrho;$$

т. е. все β_k, μ_k принадлежат множеству S.

Поскольку $\varrho \geqslant 2\lambda d^{-1} T_0(\delta)$, то

$$\|\mu_k - \beta_k\| \leqslant \lambda \|J(\beta_k)\| + 1/2\|\beta_k - \beta_0\| - 1/2\|\beta_k - \beta_0\| \leqslant \lambda d^{-1}\delta^{2^k} +$$
$$+ 1/2 r\left(1 - \frac{T_k(\delta)}{T_0(\delta)}\right) - 1/2\|\beta_k - \beta_0\| \leqslant 1/2\,(r - \|\beta_k - \beta_0\|) +$$
$$+ T_k(\delta)\left(\lambda d^{-1} - 1/2\,\frac{r}{T_0(\delta)}\right) \leqslant 1/2\,(\varrho - \|\beta_k - \beta_0\|),$$

поэтому

$$\|(2\mu_k - \beta_k) - \beta_0\| \leqslant 2\|\mu_k - \beta_k\| + \|\beta_k - \beta_0\| \leqslant \varrho.$$

Следовательно, все $(2\mu_k - \beta_k)$ также принадлежат множеству S.

Для оценивания асимптотического порядка сходимости метода (4) заметим, что доминирующим членом в оценке (9), учитывая (6), является $O(\|J(\beta_k)\|^2 \|J(\beta_{k-1})\|)$. Следуя [6], асимптотическим порядком сходимости получим $\tau = 1 + \sqrt{2}$. Теорема доказана.

Замечание 1. Если у оператора $J(\beta)$ существуют производные до третьего порядка, ограниченные по норме, а третья производная удовлетворяет условию Липшица в области S, то асимптотический порядок сходимости равен 3 [5].

Замечание 2. Существенной особенностью рассматриваемого метода является обстоятельство, что он не требует вычисления $J'(\beta)$. Существование $J'(\beta)$ использовалось нами лишь для доказательства сходимости и получения оценок ее скорости.

3. Чтобы получить условия применения метода (4) для решения системы (1), используем доказанную теорему, принимая в ней $J(\beta) = H(x(\beta), \beta)$. Очевидно, основным требованием в данном случае является дифференцируемость функции $H(x(\beta), \beta)$. Это равносильно существованию частных производных функций $H_j(x, \beta)$ $(j = 1, \dots, m)$ и дифференцируемости функций $x_i(\beta)$ $(i = 1, \dots, n)$. Существование дифференцируемого решения у задачи параметрического нелинейного программирования изучалось несколькими авторами и получены достаточные условия для довольно общего класса задач [8]. Учитывая эти условия, доказанную теорему можно переформулировать для системы уравнений (1). Например, в частном случае, когда в задачах (2) $\Gamma_i(\beta) = R^{n_i}$ $(i = 1, \dots, n)$, легко получить из теоремы следующее

Следствие. Пусть $\beta_0 \in R^m$. Предположим, что при $\beta = \beta_0$ существуют решения $x_i(\beta_0)$ задач (2) $(i = 1, \dots, n)$ и выполнены условия

1) функции $F_i(x_i, \beta)$ выпуклы по x_i и дважды непрерывно дифференцируемы по x_i, β в окрестности точек $(x_i(\beta_0), \beta_0)$;

2) матрицы $\nabla_{xx}^2 F_i(x_i(\beta_0), \beta_0)$ невырожденные.

Отсюда по теореме о неявной функции следует, что существует окрестность $U(\beta_0)$ точки β_0 такая, что имеются единственные непрерывно дифференцируемые на $U(\beta_0)$ решения $x_i(\beta)$ задач (2) $(i = 1, \ldots, n)$. Поэтому, если дифференцируемы также функции $H_j(x, \beta)$ $(j = 1, \ldots, m)$ (x и β здесь как независимые переменные) на соответствующих множествах и на множестве $S = \{\beta \in R^m \mid \|\beta - \beta_0\| \leqslant \varrho\} \subset U(\beta_0)$ выполнены условия 2°, 3°, 4° теоремы при $J(\beta) = H(x(\beta), \beta)$, то справедливы все утверждения теоремы о сходимости метода (4) относительно системы уравнений (1).

Замечание 3. Здесь предполагалось, что при фиксированных β вычисляются точные значения функций $H_j(x(\beta), \beta)$, но в общем случае это невозможно, поскольку вспомогательные задачи (2) обычно решаются приближенно. Поэтому эффективность численного нахождения решения β^*, наряду с выбранным методом решения системы (1), определяется также методом минимизации, используемым в вспомогательных задачах (2), и точностью найденных там приближенных решений [2].

ЛИТЕРАТУРА

1. *Ульм С.* Методы декомпозиции для решения задач оптимизации. Таллин, «Валгус», 1979.
2. *Евтушенко Ю. Г.* Методы решения экстремальных задач и их применение в системах оптимизации. М., «Наука», 1982.
3. *Ортега Дж., Рейнболдт В.* Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М., «Мир», 1975.
4. *Ульм С. Ю.* Изв. АН ЭССР, Физ. Матем., **16**, № 1, 13—27 (1967).
5. *Ваарманн О., Полль В.* Изв. АН ЭССР, Физ. Матем., **26**, № 2, 123—127 (1977).
6. *Schmidt, J. W.* Z. angew. Math. und Mech., **43**, № 3, 97—110 (1963).
7. *Поляк Б. Т.* Ж. вычисл. мат. и мат. физ., **4**, № 6, 995—1005 (1964).
8. *Fiacco, A.* Math. Progr., **10**, № 3, 287—311 (1976).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
18/II 1983

T. MERESSOO

ÜHEST MITTELINEAARSETE VORRANDITE LAHENDAMISE MEETODIST JÄRGUGA $1+\sqrt{2}$

Artiklis on vaadeldud mittelineaarsete võrrandite süsteemi (1) lahendamist. Sel eesmärgil on uuritud tuletistevaba meetodit (4) ja tõestatud küllalt üldistel eeldustel meetodi koonduvus operaatorvõrrandi $J(\beta)=0$ lahendiks asümptootilise järguga $1+\sqrt{2}$. On esitatud tingimused süsteemi (1) lahendamiseks meetodil (4).

T. MERESSOO

ON A $1+\sqrt{2}$ ORDER METHOD FOR THE SOLUTION OF NONLINEAR EQUATIONS

A nonlinear equation $J(\beta)=0$, where $J(\beta)$ is a differentiable operator from a Banach space to another, is considered. To solve this equation, the convergence theorem concerning the derivative-free iterative method (4) is proved. It is shown that the method (4) converges to a solution of the equation with the order $1+\sqrt{2}$ at least. This result is reformulated for the system of nonlinear equations $H_j(x, \beta)=0$ $(j=1, \ldots, m)$, where $x=(x_1, \ldots, x_n)$, and for each i $x_i=x_i(\beta)$ is the solution of the nonlinear programming problem (2), depending parametrically on $\beta=(\beta_1, \ldots, \beta_m)$.