УДК 519.615.5

© **В. М. Вержбицкий, Н. А. Макурочкина**
vervm@udm.ru, pulenkov@udm.net

СХОДИМОСТЬ МЕТОДА НЬЮТОНА С АППРОКСИМАЦИЕЙ ОБРАТНЫХ МАТРИЦ ПРОЦЕССОМ ШУЛЬЦА–ЗЕЙДЕЛЯ

Ключевые слова: система нелинейных уравнений, метод Ньютона–Шульца–Зейделя, сходимость, оценка погрешности.

Abstract. For solving systems of non-linear finite-dimensional equations, a modification of the Newton method is proposed which uses consecutive approximation of inverse matrices by the Schulz–Seidel process. Conditions for a quadratic convergence of this method are specified.

Для решения систем нелинейных уравнений

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

где $F : M \subseteq \mathbb{R}_n \to \mathbb{R}_n$ — дифференцируемая по Фреше векторная функция векторного аргумента, широко используется метод Ньютона и его различные модификации. В частности, в [1] обоснована квадратичная сходимость метода Ньютона с аппроксимацией обратных матриц по методу Шульца, определяемая формулами

$$\begin{cases} \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mathbf{A}_k F\big(\mathbf{x}^{(k)}\big), \qquad k = 0, 1, 2, \dots, \\ \quad \boldsymbol{\Psi}_k = \mathbf{E} - F'\big(\mathbf{x}^{(k+1)}\big)\mathbf{A}_k, \quad \mathbf{A}_{k+1} = \mathbf{A}_k + \mathbf{A}_k \boldsymbol{\Psi}_k, \end{cases} \tag{2}$$

где $\mathbf{A}_0$ — матрица, близкая к матрице $[F'\big(\mathbf{x}^{(0)}\big)]^{-1}$. Там же, а также в [2] можно найти предложения по модификации итерационного процесса Шульца привлечением к нему идеи, заложенной в известном методе Зейделя решения систем линейных алгебраических уравнений. Применение процесса Шульца–Зейделя в методе Ньютона трансформирует совокупность формул (2) к виду

$$\begin{cases} \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mathbf{A}_k F\big(\mathbf{x}^{(k)}\big), \quad k = 0, 1, 2, \ldots, \\ \mathbf{\Psi}_k = \mathbf{E} - F'\big(\mathbf{x}^{(k+1)}\big)\mathbf{A}_k, \quad \mathbf{A}_{k+1} = \mathbf{A}_k + \mathbf{A}_k \underline{\mathbf{\Psi}_k} + \mathbf{A}_{k+1}\overline{\mathbf{\Psi}_k}, \end{cases} \tag{3}$$

где $\underline{\mathbf{\Psi}_k}$ и $\overline{\mathbf{\Psi}}_k$ — результаты аддитивного разложения матрицы-невязки $\mathbf{\Psi}_k$ на нижнюю и строго верхнюю треугольные матрицы, соответственно.

Как известно [1, 2], в случае, если обращаемая матрица является верхней треугольной с диагональю, не содержащей нулей, и если в качестве начального приближения берется диагональная матрица с диагональю из обратных величин диагонали исходной матрицы, то итерационный процесс Шульца–Зейделя выдает точную обратную матрицу за один шаг. Следовательно, если матрица Якоби $F'\big(\mathbf{x}\big)$ на каждом итерационном шаге метода (3) сохраняет верхнетреугольную структуру и за начальную матрицу $\mathbf{A}_0$ принимается матрица $[F'\big(\mathbf{x}^{(0)}\big)]^{-1}$, то метод (3) эквивалентен классическому методу Ньютона–Канторовича. Такое заведомо имеет место для систем (1), например, с $F(\mathbf{x}) = \big(f_1(x_1, x_2, \ldots, x_n); f_2(x_2, \ldots, x_n); \ldots; f_n(x_n)\big)^T$.

Изучение метода (3) на основе общих теорем сходимости из [1] (теоремы 9.3, 9.4) по аналогии с получением утверждения о сходимости метода (2) (теорема 9.7 в [1]) приводит к следующему результату.

Т е о р е м а 1. *Пусть функция* $F(\mathbf{x})$ *определена и дифференцируема в* $M \subseteq \mathbb{R}_n$, *причем* $\exists\, L > 0 : \|F'(\mathbf{x}) - F'(\widetilde{\mathbf{x}})\| \leqslant L\|\mathbf{x} - \widetilde{\mathbf{x}}\|$ $\forall\, \mathbf{x}, \widetilde{\mathbf{x}} \in M$.

Тогда, если вектор $\mathbf{x}^{(0)}$ *и матрица* $\mathbf{A}_0$ *таковы, что при* *некотором* $\lambda > 0$ *выполняются неравенства*

$$\|\mathbf{E} - F'\big(\mathbf{x}^{(0)}\big)\mathbf{A}_0\| \leqslant L\lambda^2 \|F\big(\mathbf{x}^{(0)}\big)\|,$$

$$\nu := 4L\lambda^2 \|F\big(\mathbf{x}^{(0)}\big)\| \leqslant 1 - \frac{\|\mathbf{A}_0\|}{\lambda} \tag{4}$$

$$\text{и} \quad S := \left\{ \mathbf{x} \in \mathbb{R}_n : \|\mathbf{x} - \mathbf{x}^{(0)}\| \leqslant \lambda \|F\big(\mathbf{x}^{(0)}\big)\| \cdot \sum_{i=0}^{\infty} \nu^{2^i - 1} \right\} \subseteq M, \tag{5}$$

то начатый с данных $\mathbf{x}^{(0)}$, $\mathbf{A}_0$ *метод* (3) *сходится в* S *к решению* $\mathbf{x}^*$ *уравнения* (1) *и имеет место оценка погрешности*

$$\|\mathbf{x}^* - \mathbf{x}^{(k)}\| \leqslant \frac{\lambda \|F\big(\mathbf{x}^{(0)}\big)\|}{1 - \nu^{2^k}} \nu^{2^k}.$$

Доказательство теоремы сводится к построению последовательностей невозрастающих скалярных величин p_k, β_k, b_k, определяемых при $k = 0, 1, 2, \ldots$ формулами $p_{k+1} = 4L\lambda^2 p_k^2$, ($p_0 := \|F\big(\mathbf{x}^{(0)}\big)\|$), $\beta_k = 2L\lambda^2 p_k$, $b_{k+1} = \dfrac{\beta_k^2}{1 - \beta_k}$ ($b_0 := L\lambda^2 p_0$), и показу того, что они одновременно мажорируют соответственно $\|F\big(\mathbf{x}^{(k)}\big)\|$, $\|\mathbf{\Psi}_k\|$, $\|\mathbf{E} - F'\big(\mathbf{x}^{(k)}\big)\mathbf{A}_k\|$ и при этом $\|\mathbf{A}_k\| \leqslant \lambda$.

Фигурирующее в теореме неравенство (4) можно рассматривать как основное требование, предъявляемое к параметру λ. Чтобы проанализировать это требование, преобразуем неравенство (4) к виду

$$4L\|F\big(\mathbf{x}^{(0)}\big)\|\lambda^3 - \lambda + \|\mathbf{A}_0\| \leqslant 0. \tag{6}$$

Соответствующее ему кубическое уравнение всегда имеет вещественный отрицательный корень λ_1, не представляющий интереса. Найдя дискриминант, выясняем, что существование других вещественных (заведомо положительных) корней λ_2, λ_3 обеспечивается выполнением условия $\eta := L\|\mathbf{A}_0\|^2 \|F\big(\mathbf{x}^{(0)}\big)\| \leqslant \frac{1}{27}$. Наиболее простое выражение этих корней имеет место в случае их совпадения (при $\eta = \frac{1}{27}$). Подставив соответствующее значение

$$\lambda = \frac{1}{2} \sqrt[3]{\frac{\|\mathbf{A}_0\|}{L\|F\big(\mathbf{x}^{(0)}\big)\|}} \tag{7}$$

в вышеприведенную теорему, приходим к следующему утверждению.

С л е д с т в и е 1. *Если начальный вектор* $\mathbf{x}^{(0)}$ *и начальная матрица* $\mathbf{A}_0$ *в итерационном процессе* (3) *таковы, что:*

$$\eta := L\|\mathbf{A}_0\|^2\|F\bigl(\mathbf{x}^{(0)}\bigr)\| \leqslant \frac{1}{27}, \quad \|\mathbf{E} - F'\bigl(\mathbf{x}^{(0)}\bigr)\mathbf{A}_0\| \leqslant \frac{\sqrt[3]{\eta}}{4}$$
$$\text{и} \qquad S\bigl(\mathbf{x}^{(0)}, r\bigr) \subseteq M,$$

то справедливо заключение теоремы (*с подстановкой значения* λ *из* (7) *и* $\nu := 4\eta$ *в формулы* (4) *и* (5)) .

Другие следствия теоремы, ориентированные на расширение области выбора начальных приближений или на ускорение сходимости $\bigl(\mathbf{x}^{(k)}\bigr)$ к $\mathbf{x}^{(*)}$ (аналоги см. в [3], где исследовался метод (2) применительно к операторным уравнениям в банаховых пространствах) получаются изучением случая, когда имеется отрезок $[\lambda_2, \lambda_3]$ положительных решений рассматриваемого кубического неравенства (6).

Список литературы

1. Вержбицкий В. М. Численные методы (линейная алгебра и нелинейные уравнения). 2-е изд. М.: ОНИКС 21 век, 2005. 432 с.
2. Вержбицкий В. М. Об итерационном процессе Шульца-Зейделя // Алгоритмический анализ неустойчивых задач. Екатеринбург: Изд-во Урал. ун-та, 2004. С. 108–109.
3. Вержбицкий В. М. Выбор параметров в теоремах сходимости одного аппроксимационного аналога метода Ньютона // Журн. вычислит. математики и мат. физики. 1975. Т. 15, № 6. С. 1594–1597.