

М.Я.Бартін, Ю.М.Щербіна

ІТЕРАЦІЙНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ
НЕЛІНІЙНИХ ФУНКЦІОНАЛЬНИХ РІВНЯНЬ
І ЕКСТРЕМАЛЬНИХ ЗАДАЧ

/Огляд/

1. Постановка задач.

Розроблено ітераційні методи розв'язування таких трьох важливих для практики нелінійних задач.

Перша задача – розв'язування нелінійного функціонального рівняння – формулюється таким чином:

задано $P: X \rightarrow Y$,
знайти $x_* \in X$, для якого $P(x_*) = 0 \in Y$,
де X та Y – банахові простори. Зокрема, при $X = Y = R^n$,
де R^n означає n -вимірний евклідов простір, отримаємо систему нелінійних рівнянь з дійсними змінними. /1.1/

Друга задача – мінімізація функцій без обмежень:

задано $f: R^n \rightarrow R$,
знайти $x_* \in R^n$, для якого $f(x_*) \leq f(x)$
при всіх $x \in R$. /1.2/

В обчислювальній практиці часто обмежуються пошуком локального мінімуму, тобто вимагають виконання умови /1.2/ лише для деякого околу $S(x_*)$ точки x_* . Задача /1.2/ звичайно записується у формі

$$\min f(x), \quad x \in R^n. \quad /1.2/$$

Третя задача називається задачею мінімізації з обмеженнями:

$$\left. \begin{aligned} f_0: R^n \rightarrow R, \quad \min f(x), \\ \text{при обмеженнях} \quad \left. \begin{aligned} f_i(x) &\leq 0, \quad i \in J, \\ f_i(x) &= 0, \quad i \in J^0, \\ x &\in D \subseteq R^n, \end{aligned} \right\} \end{aligned} \right\} \quad /1.3/$$

де D - опукла множина /простої будови/. J^-, J^0 - скінченні множини індексів, $J^- \cap J^0 = \emptyset$.

2. Загальна схема дослідження методів типу Ньютон-Канторовича для функціональних рівнянь. Ефективність методів.

Нехай x_k - достатньо добре наближення до розв'язку x_* рівняння /1.1/. Запишемо розклад $P(x)$ в околі точки x_k за формулою Тейлора /далі вважаємо, що оператор $P(x)$ потрібну кількість разів диференційований за Фреше/

$$P(x) = P(x_k) + P'(x_k)(x - x_k) + \dots + \frac{1}{m!} P^{(m)}(x_k)(x - x_k)^m + o(\|x - x_k\|^m). \quad /2.1/$$

Лінеаризувачи записаний розклад в околі точки x_k , отримуємо

$$P(x) = P(x_k) + P_m(x_k)(x - x_k) + o(\|x - x_k\|^{\tau}), \quad \tau \leq m,$$

де

$$P_m(x_k) = P'(x_k) + \frac{1}{2} P''(x_k) \psi_2(x_k) + \dots + \frac{1}{m!} P^{(m)}(x_k) \psi_m(x_k),$$

$\psi_i(x_k), i=1, \dots, m$ - деякі лінійні оператори.

Для визначення $(k+1)$ -го наближення x_{k+1} дістаємо лінійне рівняння

$$P(x_k) + P_m(x_k)(x - x_k) = 0.$$

Звідси, якщо існує обернений оператор $P_m^{-1}(x_k)$, знаходимо

$$x_{k+1} = x_k - P_m^{-1}(x_k) P(x_k), \quad k = 0, 1, \dots \quad /2.2/$$

Показано [1], що при відповідному виборі операторів $\psi_i(x_k)$ для послідовності $\{x_k\}$, що отримується за формулою /2.2/, дійсна оцінка

$$\|x_{k+1} - x_*\| \leq R \|x_k - x_*\|^{m+1}, \quad /2.3/$$

де $R = \text{const}$.

Таким чином, оператор $\Omega(x) = x - P_m^{-1}(x) P(x)$ породжує ітераційну формулу $x_{k+1} = \Omega(x_k)$ порядку $\tau = m+1$.

Частинними варіантами формули /2.2/ є метод Ньютон / $m=1$ /, метод дотичних гіпербол / $m=2$; $\psi_2(x) = -[P'(x)]^{-1} P(x)$ / та ін.

Замість лінеаризації на основі формули Тейлора, що приводить до високх похідних, можна скористатись ідеєю Рунге про апроксима-

цію відрізка ряду Тейлора лінійною комбінацією перших похідних, обчислених, правда, у різних точках. Так приходимо до методів "типу Рунге" [2]. Розглянемо лінійну комбінацію

$$F(x) = P(x_k) + [\alpha_1 F_1(x_k) + \dots + \alpha_m F_m(x_k)](x - x_k), \quad /2.4/$$

де $F_i(x_k)$, $i = 1, \dots, m$ - лінійні оператори:

$$F_1(x_k) = P'(x_k)$$

$$F_m(x_k) = P'(x_k) + \sum_{i=1}^{m-1} \beta^{(m,i)} [F_i(x_k)]^{-1} P(x_k),$$

$\alpha_i, \beta^{(i,j)}$ ($i = 1, \dots, m$; $j = 1, \dots, i-1$) - дійсні числа.

Коефіцієнти $\alpha_i, \beta^{(i,j)}$ підбираються так, щоб лінійна комбінація /2.4/ збігалася з розкладом оператора $P(x)$ за формулою Тейлора в околі точки x_k з максимальною точністю. Нехай такі $\alpha_i, \beta^{(i,j)}$ вже знайдені. Тоді наступне наближення розв'язку рівняння /1.1/ знаходимо з умови

$$F(x_{k+1}) = 0.$$

Якщо існує лінійний оператор $V_k^{-1} = [\alpha_1 F_1(x_k) + \dots + \alpha_m F_m(x_k)]^{-1}$,

то для обчислення x_{k+1} отримаємо формулу

$$x_{k+1} = x_k - V_k^{-1} P(x_k).$$

Наприклад, при $m = 2$ дістаємо однопараметричну сім'ю методів

$$V_k = (1 - \beta) P'(x_k) + \beta P'(x_k - 0.5\beta^{-1} \Gamma_k P(x_k)),$$

$$\Gamma_k = [P'(x_k)]^{-1}, \quad \beta \neq 0 \quad - \text{довільне число.}$$

При $\beta = 1$ отримаємо запропонований Т.І.Коган метод, який має третій порядок збіжності /див. /2.3/ при $m = 2$ /. Деякі застосування методів типу Рунге висвітлено у [15]. Деякі інші методи з використанням лише першої похідної запропоновано в [23].

У рамках описаних методик розглядаються також різнищеві аналоги методів, у цьому випадку похідні в обчислювальних формулах замінюються поділеними різницями [1, 2].

Із викладеного видно, що можна будувати дуже багато методів розв'язування рівняння /1.1/. Тому виникає питання про вибір кращого у певному смислі методу для розв'язування конкретної задачі. У [3] запропоновано один із варіантів асимптотичної оцінки ефективності методу в сенсі кількості обчислень, витрачених на отримання

мання кінцевого результату. Метод з порядком збіжності $\tau_1 > 1$ є ефективнішим від методу з порядком збіжності $\tau_2 > 1$, якщо виконується нерівність

$$\log_{\tau_1} \tau_2 < \frac{N_2}{N_1},$$

де N_i / $i = 1, 2$ / характеризує складність однієї ітерації за першим методом /може бути кількість елементарних операцій, час тощо/.

З метою отримання більш ефективних методів можна використувати різні модифікації основних методів. Один із таких шляхів отримання нових ітераційних методів – це застосування рекурсії ітераційних формул. У [4] показано, що коли неперервні оператори $\Phi: X \rightarrow X$, $\Omega: X \rightarrow X$, $\Omega(x) = x - R(x)P(x)$ –

породжують ітераційні формули $x_{k+1} = \Phi(x_k)$, $x_{k+1} = \Omega(x_k)$ порядку φ та ω відповідно, то оператор $Q: X \rightarrow X$,

$$Q(x) = \Phi(x) - H(x)P[\Phi(x)],$$

де $\|R(\Phi(x_k)) - H(x_k)\| \leq K \|x_k - x_*\|^L$,

$$K = 0, 1, 2, \dots, K = \text{const},$$

породжує ітераційну формулу $x_{k+1} = Q(x_k)$ з порядком збіжності послідовності $\{x_k\}$ до розв'язку x_* рівняння /1.1/, який дорівнює $\varphi = \min\{\omega\varphi, \ell + \varphi\}$.

На основі t -кратного застосування цього результату можна отримати метод Ньютона, у якого оператор $\Gamma(x) = [P'(x)]^t$ залишається постійним протягом кількох кроків і ряд інших методів, наприклад, такий [16]:

$$x_{k+1} = A_t(x_k),$$

де $A_j(x) = A_{j-1}(x) - [P'(x - \Gamma(x)P(x))]^{-1} P(A_{j-1}(x))$,
 $j = 2, \dots, t$

$$A_1(x) = x - \frac{1}{2} \Gamma(x) P(x) - \frac{1}{2} [P'(x - \Gamma(x)P(x))]^{-1} P(x),$$

$$\Gamma(x) = [P'(x)]^t.$$

Порядок збіжності цього методу дорівнює $2t + 1$ і залежить від глибини рекурсії t .

В іншому варіанті побудови комбінацій основних методів застосовується така модифікація методу Ньютона [5]:

$$x_{k+1} = x_k - [P'(\bar{x}_k)]^{-1} P(x_k), \quad k = 0, 1, 2, \dots \quad /2.5/$$

де $\bar{x}_k = \alpha x_k + (1-\alpha)\Phi(x_k)$, $0 \leq \alpha \leq 1$,

$\Phi(x)$ - оператор, що породжує ітераційну формулу порядку τ ($1 \leq \tau \leq 2$). Застосування цього методу та його подальші дослідження є в [11, 12].

Докладно досліджено вперше запропонований у [3] ітераційний метод з пам'яттю для розв'язування рівняння /1.1/

$$x_{k+1} = x_k - [P'(\bar{x}_k)]^{-1} P(x_k) \quad /2.6/$$

$$\bar{x}_k = \begin{cases} x_0 & \text{при } k=0 \\ x_k - \frac{1}{2} [P'(\bar{x}_{k-1})]^{-1} P(x_k), & k=1, 2, \dots \end{cases}$$

Швидкість збіжності його дорівнює $1 + \sqrt{2}$, а за ефективністю він не поступається і навіть у багатьох випадках переважає метод Ньютона. Дослідження цього методу з урахуванням похибок, а також його різнищового аналога продовжено в [6-8], рекурсивний варіант розглянуто в [9].

3. Безумовна мінімізація.

Для розв'язування задачі /1.2/ у першу чергу було досліджено метод, аналогічний /2.6/

$$x_{k+1} = x_k - [f''(\bar{x}_k)]^{-1} f'(x_k), \quad /3.1/$$

де

$$\bar{x}_k = \begin{cases} x_0 & \text{при } k=0 \\ x_k - \frac{1}{2} [f''(\bar{x}_{k-1})]^{-1} f'(x_k), & k=1, 2, \dots \end{cases}$$

Для сильно опуклих функцій доведена локальна збіжність методу з порядком $1 + \sqrt{2}$ [17].

У [18] розглянуто подальше вдосконалення цього алгоритму з відмовою від властивості опуклості та доведено теорему про глобальну збіжність.

При цьому застосовується LDL^T - розклад Холеського матриці Гессе.

Нехай функція $f(x)$ двічі неперервно диференційовна і $f''(x)$ - матриця Гессе. Побудуємо матрицю $F(x)$, пов'язану з $f''(x)$ наступним чином:

$$F(x) = LDL^T = f''(x) + E(x), \quad /3.2/$$

де L і D - фактори розкладу Холевського матриці $F(x)$;
 L - нижня одинична трикутна матриця; D - додатна діагональна матриця; $E(x)$ - діагональна матриця поправок. Матриця $F(x)$ - додатно визначена і відрізняється від $f''(x)$ лише діагональними елементами. Зауважимо, що коли $f''(x)$ додатно визначена, то

$$F(x) = f''(x).$$

Зараз метод /3.1/ запишемо так:

$$x_{k+1} = x_k - \alpha_k F^{-1}(\bar{x}_k) f'(x_k),$$

$$x_k = \begin{cases} x_0, & k=0 \\ x_k - \frac{1}{2} \alpha_{k-1} F^{-1}(\bar{x}_{k-1}) f'(x_k), & k=1,2,3,\dots, \end{cases} \quad /3.3/$$

де $F(x)$ визначається за формулою /3.2/, а крок α_k визначається шляхом дроблення, починаючи з $\alpha = 1$ до виконання нерівності

$$f(x_k - \alpha p_k) \leq f(x_k) + \varepsilon \alpha < f'(x_k), p_k >,$$

де $p_k = -F^{-1}(\bar{x}_k) f'(x_k)$, $0 < \varepsilon < \frac{1}{2}$,

$< \cdot, \cdot >$ - скалярний добуток. Глобальна збіжність методу /3.3/ має місце для двічі неперервно диференційованої обмеженої знизу функції.

При певних додаткових умовах порядок збіжності $1 + \sqrt{2}$.

У праці [10] розглянуто іншу модифікацію методу Ньютона з прискоренням збіжності.

Нехай деяка послідовність $\{x_k\}$ визначена за формулою

$$x_{k+1} = \Phi(x_k), \quad k=0,1,\dots,$$

збігається до x_* - розв'язку задачі /1.2/ із швидкістю $\tilde{\gamma} > 1$.
 Пропонується метод

$$x_{k+1} = x_k - \left[f''\left(\frac{x_k + \Phi(x_k)}{2}\right) \right]^{-1} f'(x_k); \quad k=0,1,\dots \quad /3.4/$$

Для сильно опуклої функції цей метод локально збігається з порядком $\tilde{\gamma} + 1$. Розглянуто також варіант методу з регулюванням кроку, який забезпечує глобальну збіжність.

Деякі методи, засновані на ідеї рекурсії, розглянуто в [14, 19].

4. Нелінійне програмування.

Основні результати з нелінійного програмування опубліковано в [20-22].

В основі лежить квазіньютонівська апроксимація та факторизація за Холеським матриці других похідних за компонентами вектора x функції Лагранжа задачі /1.3/. Ми використовуємо цей підхід для прискорення збіжності методу лінеаризації Б.М.Пшеничного [13]. Не зменшуючи загальності для теоретичних досліджень, задачу /1.3/ перепишемо у такій формі:

$$\min \{ f_0(x) : f_i(x) \leq 0, i \in J = \{1, 2, \dots, m\} \}. \quad /4.1/$$

Тут $f_i(x), i \in \{0\} \cup J$ - неперервно диференційовні функції.

Поставимо у відповідність точці x задачу квадратичного програмування

$$\min_p \{ \langle f'_0(x), p \rangle + \frac{1}{2} \langle p, Ap \rangle : \langle f'_i(x), p \rangle + f_i(x) \leq 0, i \in J_\delta(x) \} \quad /4.2/$$

де A - симетрична додатно визначена матриця,

$$J_\delta(x) = \{ i \in J : f_i(x) \geq F(x) - \delta \}, \quad \delta > 0,$$

$$F(x) = \max \{ 0, f_1(x), \dots, f_m(x) \},$$

$\langle \cdot, \cdot \rangle$ - скалярний добуток векторів.

Розв'язок задачі /4.2/ та її множники Лагранжа позначимо відповідно через $p(x)$ та $u^i(x), i \in J_\delta(x)$.

Опишемо алгоритм.

Матрицю A в задачі /4.2/ перераховуємо під час роботи алгоритму таким чином. Нехай відома додатно визначена симетрична матриця A_k . Використовуючи довільну квазіньютонівську формулу [20], на k -й ітерації перераховуємо матрицю A_{k+1}

$$\bar{A}_{k+1} = A_k + B_k. \quad /4.3/$$

Симетрична матриця B_k малого рангу визначає конкретний тип квазіньютонівського перерахунку.

За допомогою модифікованого LDL^T - розкладу Холеського [20] будемо додатно визначену матрицю

$$A_{k+1} = L_{k+1} D_{k+1} L_{k+1}^T = \bar{A}_{k+1} + E_{k+1}, \quad /4.4/$$

де L_{k+1} - одинична нижня трикутна матриця; D_{k+1} - додатна діагональна матриця; E_{k+1} - деяка матриця, яка дорівнює нулю-

вій матриці у випадку, якщо $\bar{A}_{k+1}$ - істотно додатно визначена. Через T позначено транспонування матриці.

Нехай обрані числа $N > 0$, $S > 0$, $\delta > 0$, $0 < \varepsilon < \frac{1}{2}$, $C_0 > 0$, $0 < \gamma < 1$; початкове наближення x_0 . Прийmemo $A_0 = I_n$ / I_n - одинична матриця порядку n /.

Загальний крок алгоритму.

Нехай точка x_k , матриця A_k і число C_k вже побудовані.

1. Розв'язуючи задачу /4.2/ при $x = x_k$, $\bar{A} = A_k$, обчислити $p_k = p(x_k)$, $u^i(x_k)$, $i \in J_s(x_k)$. Приймати $u^i(x_k) = 0$ для $i \in J_g(x_k)$.

2. Якщо $\|p_k\| > C_k$ або $f_0(x_k + p_k) + NF(x_k + p_k) > f_0(x_0) + NF(x_0)$, то прийняти $C_{k+1} = C_k$ і перейти до кроку 3. Інакше прийняти $x_{k+1} = x_k + p_k$, $C_{k+1} = \gamma \|p_k\|$ і перейти до кроку 4.

3. Починаючи з $\alpha = 1$, дробити його шляхом ділення навпіл до першого виконання нерівності

$$f_0(x_k + \alpha p_k) + NF(x_k + \alpha p_k) \leq f_0(x_k) + NF(x_k) - \alpha \varepsilon < p_k, A_k p_k >.$$

Прийняти $x_{k+1} = x_k + \alpha p_k$.

4. Перерахувати матрицю A_{k+1} за формулами /4.3/-/4.4/.

Якщо

$$\max_i a_{k+1}^i / \min_i d_{k+1}^i \leq S,$$

то перейти до кроку 1 / a_{k+1}^i та d_{k+1}^i - i -ті діагональні елементи матриць A_{k+1} , D_{k+1} /.

Інакше прийняти $A_{k+1} = I_n$ і перейти до кроку 1.

Процес обчислень припинити, якщо $\|p_k\| \leq \varepsilon \rho_s$, де $\varepsilon \rho_s$ задана точність. Керування параметрами N та δ здійснюється, як у [13].

Докладно вивчено умови збіжності алгоритму [20].

Розглянуто модифікації для простих обмежень типу "паралелепіпеду" [21] та застосування до розв'язування задач дискретного міні-максу [22]. Складено комплекс програм для розв'язування задачі /4.1/ на мові Фортран для ЕОМ ЕС 1045.

І. Б а р т и ш М.Я. О методах типа Ньютона-Канторовича // Сиб. мат. журн. 1973. Т.14. № 4. С.904. 2. Б а р т и ш М.Я., С е н ь о П.С. Про метода типу Рунге розв'язування нелінійних операторних рівнянь // Доп. АН УРСР. Сер. А. 1972. № 9. С.771-775.

3. Бартиш М.Я. Про один ітераційний метод розв'язування функціональних рівнянь // Доп. АН УРСР. Сер. А. 1968. № 5. С.387-391. 4. Бартиш М.Я., Щербина Ю.Н. Итерационные формулы, получаемые с помощью рекурсии // Математический сборник. К., 1976. С.50-53. 5. Бартиш М.Я. Об одном классе методов типа Ньютона // Вестн. МГУ. Сер. 15. Вычислит. математика и кибернетика. 1987. № 2. С.16-20. 6. Бартиш М.Я., Щербина Ю.Н. Об одном итерационном процессе решения нелинейного операторного уравнения // Вычислит. и прикл. математика. 1972. Вып. 16. С.115-121. 7. Бартиш М.Я., Щербина Ю.М. Про один різницевий метод розв'язування нелінійних операторних рівнянь // Доп. АН УРСР. Сер. А., 1972. № 7. С.579-582. 8. Бартиш М.Я., Щербина Ю.Н. Исследование условий сходимости и оценка полной погрешности одного итерационно-разностного метода для решения нелинейных операторных уравнений // Вычислит. и прикл. математика. 1976. Вып. 28. С.3-9. 9. Бартиш М.Я., Роман Л.Д. Про один рекурсивный метод розв'язування нелінійних функціональних рівнянь // Вісн. Львів. ун-ту. Сер. мех.-мат. 1986. Вып. 26. С.3-7. 10. Бартиш М.Я. Про один клас методів типу Ньютона розв'язування задач на екстремум // Вісн. Львів. ун-ту. Сер. мех.-мат. 1988. Вып. 29. С.3-5. 11. Бартиш М.Я., Шахно С.М. О методе Ньютона с ускоренной сходимостью // Вестн. Киев. ун-та. Моделирование и оптимизация слож. систем. 1987. Вып. 6. С.62-66. 12. Бартиш М.Я., Шахно С.М. Метод Ньютона с ускоренной сходимостью и его применение к задачам газовой динамики // Распараллеливание обработки информации: Тез. докл. VI Всесоюз. школы-семинара. Львов, 1987. Ч.1. С.109-110. 13. Пшеничный Б.Н. Метод линеаризации. М., 1983. 14. Роман Л.Д. Збіжність рекурсивного методу з пам'яттю для мінімізації функцій // Вісн. Львів. ун-ту. Сер. мех.-мат. 1990. Вып.33. С.51-55. 15. Сенъо П.С., Шахно С.М. Розв'язування різницевих рівнянь газової динаміки ітераційним методом типу Рунге // Вісн. Львів. ун-ту. Сер. мех.-мат. 1981. Вып. 17. С.7-14. 16. Щербина Ю.М. Про один рекурсивний ітераційний метод із збуреннями для операторних рівнянь // Вісн. Львів. ун-ту. Сер. мех.-мат. 1981. Вып. 17. С.16-21. 17. Щербина Ю.М., Голуб Б.М. Збіжність ітераційного методу з пам'яттю для мінімізації функцій // Вісн. Львів. ун-ту. Сер. мех.-мат. 1984. Вып. 22. С.3-7. 18. Щербина Ю.М., Голуб Б.М. Збіжність ітераційного методу з пам'яттю з використанням LDL^T - розкладу Холецкого для мінімізації функцій // Вісн. Львів. ун-ту. Сер. мех.-мат. 1987. Вып. 27. С.23-25. 19. Щербина Ю.М., Остапчук Л.А. Один ітераційний метод для мінімізації функцій // Вісн. Львів. ун-ту. Сер. мех.-мат. 1986. Вып. 26. С.20-22. 20. Щербина Ю.Н., Голуб Б.М. Квазиньютоновская модификация метода линеаризации // Кибернетика. 1988. № 6. С.66-71. 21. Щербина Ю.Н., Голуб Б.М. Модификация метода линеаризации для решения задачи математического программирования на простом множестве типа "параллелепипеда" // Мат. методы и физ.-механ. поля. 1989. № 20. С.24-28. 22. Щербина Ю.Н., Голуб Б.М. Квазиньютоновская модификация метода линеаризации для решения задачи дискретного минимакса. Львов, 1987. 17 с. Рукопись деп. в УкрНИИТИ, № 2305-ук87. 23. Щербина Ю.М. Один клас ітераційних методів для розв'язування нелінійних операторних рівнянь у банаховому просторі // Вісн. Львів. ун-ту. Сер. мех.-мат. 1977. Вып. 12. С.57-60.

Стаття надійшла до редколегії 02.10.90