1975 г. ОПТИМИЗАЦИЯ Выпуск 16(33)

Численные методы

УДК 518:517.948

О НЕКОТОРЫХ ОБОБЩЕНИЯХ МЕТОДА НЬЮТОНА-КАНТОРОВИЧА ПРИБЛИЖЕННОГО РЕШЕНИЯ НЕЛИНЕЙНЫХ ОПЕРАТОРНЫХ УРАВНЕНИЙ

Б.А.Вертгейм

В статье изучается множество допустимых векторов некоторых задач нелинейного программирования и с помощью метода Ньютона-Канторовича исследуется локальное строение этого множества, а также даётся способ нахождения допустимых векторов, являющийся более эффективным по сравнению с некоторыми предложенными недавно способами.

1. В общем случае речь идёт о множестве решений уравнения

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

где $f: \mathcal{U} \longrightarrow Y$ – нелинейная операция, определённая на открытом множестве $\mathcal{U}$ пространства Банаха X, принимающая значения в пространстве Банаха Y. Следующий частный случай представляет особый интерес в связи с задачами математического программирования: изучить множество решений системы уравнений

$$\begin{gathered} f_i(x_1, x_2, \ldots, x_n)=0, \quad i=\overline{1m}, \\ m<n, \quad x=(x_1, x_2, \ldots, x_n) \in R^n, \\ f_i: \mathcal{U} \longrightarrow R, \quad \mathcal{U} \subset R^n. \end{gathered} \tag{2}$$

В работах [2,5] изучены некоторые варианты обобщения метода Ньютона-Канторовича, приспособленные для приближенного отыскания допустимого вектора, то есть решения уравнений рассматриваемого вида. Эти варианты сходны между собой. Трудность здесь состоит в том, что, как правило, не выполнено обычное для метода Ньютона-Канторовича предположение о существовании непрерывного обратного оператора $[f'(x^0)]^{-1}$ (см. [I]), где x^0 - начальное приближение, $f'(x^0)$ - производная Фреше, которая в рассматриваемом сейчас случае обычно не является взаимно-однозначным отображением X в Y ; иначе говоря, ядро этого оператора - обозначим его буквой N - не тривиально: $N = \ker f'(x^0) \neq \{0\}$. В работе [2] применен переход к фактор-пространству, причем процедура отыскания решения на каждом итерационном шаге осложнена поиском в соответствующем классе смежности - элементе из пространства X/N - представителя x', имеющего норму в X , достаточно близкую к нижней грани норм элементов этого класса. В статье [5] взамен предположения о существовании ограниченного оператора $[f'(x^0)]^{-1}$ делается предположение о существовании правого обратного оператора для $f'(x^0)$ и доказываются аналоги известных теорем о сходимости основного и модифицированного процессов Ньютона - Канторовича, которые в рассматриваемом в [5] случае записываются с помощью упомянутых правых обратных операторов. При этом не рассматривается вопрос о локальном строении множества решений уравнения (I) вблизи исходного начального приближения, а процедура нахождения последовательных приближений, предложенная для системы (2), оказывается излишне трудоёмкой, что мы покажем далее; отметим также, что основной процесс Ньютона - Канторовича исследован в работе [5] лишь в случае, когда пространство X гильбертово.

2. Мы будем, как обычно, рассматривать некоторую начальную точку $x = a$ в качестве исходного приближения к решению уравнения (I). Пусть в точке a и в некоторой её окрестности $\mathcal{U}$ операция f дифференцируема в смысле Фреше. Обозначим $\ker f'(a) = N$ и предположим существование непрерывного проектора $P_N : X \longrightarrow N$, иначе говоря, предположим, что замкнутое подпространство N имеет в X топологическое дополнение $M = \ker P_N$.

Введём ещё одно банахово пространство Y_1 , гомеоморфное подпространству M и обозначим буквой φ линейное инъективное преобразование $\varphi : Y_1 \longrightarrow X$ такое, что $\varphi(Y_1) = M$. В частности, в качестве Y_1 можно взять само пространство M и φ - тождественное вложение M в X.

Обратим теперь внимание на то, что линейный оператор

$$A = f'(a)\circ\varphi : Y_1 \longrightarrow Y \qquad (3)$$

является взаимно-однозначным отображением Y_1 в Y. Предположим ещё, что $\operatorname{Im} f'(a) = Y$, тогда A, очевидно, также будет отображением на всё пространство Y, и в силу теоремы Банаха об обратном операторе будет существовать ограниченный обратный оператор A^{-1}.

Следующий вопрос является, по-видимому, наиболее интересным и естественным в данной ситуации: можно ли указать такие окрестности $\mathcal{U}_i$, $i = 1,2,3$, точки a, что, какова бы ни была точка $x^\circ \in \{a^0 + N\} \cap \mathcal{U}_1$, уравнение (I) имеет решение $x = x(x^\circ)$, причём это решение единственно в множестве $\mathcal{U}_3$; здесь $\mathcal{U}_2 \subset \mathcal{U}_3$. Естественность такой постановки видна, если вспомнить теорему Люстерника.

Для решения поставленного вопроса удобно сделать замену переменной в уравнении (I), полагая

$$x = x^\circ + \varphi(\eta) \qquad (4)$$

(элемент x° считаем заданным и фиксированным, элемент η переменным и искомым) и применить к полученному уравнению

$$F(x^\circ, \eta) \equiv f(x^\circ + \varphi(\eta)) = 0 \qquad (5)$$

метод Ньютона-Канторовича или близкие способы.

3. Переходим к описанию способа построения последовательных приближений. Нам потребуется производная

$$F'_\eta(x^\circ, \eta) = f'(x^\circ + \varphi(\eta))\circ\varphi : Y_1 \longrightarrow Y. \qquad (6)$$

При этом, как отмечалось, в силу теоремы об обратном операторе, при $x^\circ = a$ и $\eta = 0$ существует оператор

$$\left[F'_\eta(a, 0)\right]^{-1} = A^{-1} : Y \longrightarrow Y_1. \qquad (6')$$

Предположив, что зависимость f' от x непрерывна, мы можем гарантировать существование обратного оператора $[F'_\eta(x^\circ, \eta)]^{-1}$ для значений (x°, η), близких к $(a, 0)$.

Запишем формулы для основного процесса Ньютона-Канторовича, следуя работе [I] и выбирая в качестве начального приближения к решению уравнения (5) элемент $\eta^\circ = 0$:

$$\eta^{K+1} = \eta^K - \left[f'(x^\circ + \varphi(\eta^K)) \circ \varphi\right]^{-1} f(x^\circ + \varphi(\eta^K)), \qquad (7)$$
$$K = 0, 1, 2, \ldots .$$

Аналогичные формулы для модифицированного процесса используют $\left[f'(x^\circ) \circ \varphi\right]^{-1}$ вместо обратного оператора из формулы (7). Для фактических вычислений удобнее всего сделать дальнейшее упрощение, применяя следующий итеративный процесс с одним и тем же обратным оператором – на этот раз для всех начальных векторов, близких к a

$$\eta^{K+1} = \eta^K - \left[f'(a) \circ \varphi\right]^{-1} f(x + \varphi(\eta^K)) . \qquad (8)$$

Сформулируем результаты о существовании, единственности решения уравнения (I) и сходимости итераций.

ТЕОРЕМА I. Пусть дано уравнение (I) и точка $a \in X$, причем отображение f дифференцируемо по Фреше, $\mathcal{U}$ – открытое множество, $a \in \mathcal{U}$. Пусть выполнены указанные выше предположения о $N = \ker f'(a)$, M, φ, Y_1 и следующие условия:

1. $\|A^{-1}\| \le B_0$, $\quad A = f'(a) \circ \varphi$;
2. $B_0 \|f'(a)\| \le \gamma$;
3. $B_0 \|f(x^\circ) - f(a)\| \le \nu$, $\quad B_0 \|[f'(x^\circ) - f'(a)] \circ \varphi\| \le \delta < 1$;

здесь и далее $x^\circ \in \mathcal{U} \cap \{a + N\}$.

4. Для производной Фреше f' выполнено условие Гёльдера

$$\|[f'(x^\circ + \varphi(\eta)) - f'(x^\circ)] \circ \varphi\| \le K \|\eta\|^\alpha, \ 0 < \alpha \le 1, \ x^\circ + \varphi(\eta) \in \mathcal{U};$$

5. $h_0 = B_0 K (\gamma + \nu)^\alpha < \left(\frac{\alpha}{1+\alpha}\right)^\alpha (1 - \delta)^{1+\alpha}$;

6. $E_1 = \{x / x = x^\circ + \varphi(\eta), \ \eta \le R_1(h_0)(\gamma + \nu)\} \subset \mathcal{U}$,

где R_1' – меньший корень уравнения

$$\frac{h_0}{1+\alpha} R^{1+\alpha} - (1 - \delta) R - 1 = 0 . \qquad (9)$$

Тогда уравнение (I) имеет решение $x^* = a + \varphi(\eta^*) \in E_1$ и итерации (8) сходятся к η^*, $\|\eta^k - \eta^*\| \leq \frac{q^k}{1-q}(\gamma + \nu)$, $q = h_0 R_1^\alpha + \delta < 1$.

ТЕОРЕМА 2. В условиях теоремы I решение x^* единственно в множестве $E_2 = \{x / x = x^0 + \varphi(\eta), \|\eta\| \leq R_2(h_0)(\gamma + \nu)\}$, если $E_2 \subset \mathcal{U}$. Здесь $R_2 = R_2(h_0)$ - больший корень уравнения (9).

ТЕОРЕМА 3. Пусть выполнены условия, аналогичные условиям теоремы I, с тем отличием, что условие Гёльдера имеет вид

$$\|[f'(x^0 + \varphi(\eta')) - f'(x^0 + \varphi(\eta''))]\varphi\| \leq K \|\eta' - \eta''\|^\alpha, \quad 0 < \alpha \leq 1,$$

а условия 5 и 6 заменены на $5'$: $h_0 \leq Q(\alpha)(1-\delta)^{1+\alpha}$, где $Q(\alpha) \geq \frac{\alpha}{1+\alpha}$ - корень уравнения $\left(\frac{Q}{1+\alpha}\right)^\alpha = (1-Q)^{1+\alpha}$; $6'$: $E' = \{x = x^0 + \varphi(\eta), \|\eta\| \leq \frac{\gamma + \eta}{(1-s)(1-\delta)}\}$, $E' \subset \mathcal{U}$, где $s^\alpha = 1 - Q$. Тогда уравнение (I) имеет решение $x^* = x^0 + \varphi(\eta^*) \in E'$ и итерации (7) сходятся к η^*, $\|\eta^k - \eta^*\| \leq \frac{s^k}{1-s} \cdot \frac{\gamma + \nu}{1-\delta}$.

ЗАМЕЧАНИЕ. Если в условии $5'$ применить знак $<$, то скорость сходимости - сверхлинейная, порядка $1+\alpha$.

ДОКАЗАТЕЛЬСТВО теорем I-3. Часть рассуждений, необходимых для доказательства - обратимость оператора A, замена (4) - проведена выше. Вывод теорем I и 2 продолжается на основе теорем I.3 и I.4 работы [3], в которых изучен способ, близкий к модифицированному процессу, приводящий здесь к итерациям (8). Для доказательства теоремы 3 проверим, что существует обратный оператор $[f'(x^0)\varphi]^{-1}$. Введём оператор $T = A^{-1}[A - f'(x^0)\varphi]$. Так как $\|T\| \leq \delta < 1$, то оператор $I - T$ имеет обратный $G = [I - T]^{-1} = [A^{-1} f'(x^0)\varphi]^{-1}$, причем $\|G\| \leq (1-\delta)^{-1}$. Теперь проверяем, что существует обратный оператор $[f'(x^0)\varphi]^{-1} = G A^{-1}$

и $\|GA^{-1}\| \le B_0(1-\delta)^{-1}$. Наконец, проверяем выполнение для уравнения (5) всех условий теоремы 2 из статьи [4] о сходимости основного процесса Ньютона-Канторовича. Мы опускаем детали, в частности, связанные с упомянутой выше скоростью сходимости порядка $1+\alpha$.

Рассмотрим кратко примеры, поясняя лишь изложенную выше конструкцию, связанную с ядром N оператора $f'(a)$ и с его топологическим дополнением M ; на обычных иллюстрациях метода Ньютона мы не останавливаемся.

I. Пусть $X=R^n$, $Y=R^m$, $f'(a)$ определяется матрицей (p_{ij}), $i=\overline{1,m}$ $j=\overline{1,n}$; $n \ge m$; i-ю строку этой матрицы обозначим p_i. Пусть ранг этой матрицы равен m, так что $\operatorname{Im} f'(a)=R^m$. Ядро N определяется системой уравнений

$$\sum_{j=1}^{n} p_{ij}\, x_j = 0, \quad i=\overline{1,m} \quad . \tag{10}$$

В качестве M удобно взять $N^{\perp}$, при этом ясно, что

$$M=\{x\in R^n \mid x=\sum_{i=1}^{m} \lambda_i p_i,\ \lambda_i \in R\} = \operatorname{Im}\,[f'(a)]^*,$$

где $*$ означает переход к сопряжённому оператору (и к транспонированной матрице). Взяв $Y_1=R^m$, $\varphi=[f'(a)]^*$, $\varphi:\lambda \mapsto [f'(a)]^*\lambda$, $\lambda\in R^m$, мы удовлетворим соответствующим условиям теории, причем обратимый оператор $A=f'(a)$ будет иметь квадратную матрицу размером $m\times m$ (q_{ij}), $q_{ij}=\langle p_i, p_j\rangle$ (элементы матрицы - это скалярные произведения, определитель её, очевидно, ненулевой). При этом вычисление итераций по формулам (7,8) оказывается заметно более экономным, чем соответствующие вычисления по способу статьи [5], где предлагается находить правый обратный оператор для $f'(a)$, причём в построении участвует базис пространства N - его определение связано с нахождением всех нетривиальных решений однородной задачи (10) размером $m\times n$, - а затем решаются ещё системы уравнений для построения так называемого канонического правого обратного оператора.

2. Пусть $X=l^2$, $Y=R^m$. Предыдущие построения переносятся на этот случай: $f'(a)$ определяется m линейно-независимыми векторами $p_i\in(l^2)^*\cong l^2$; снова $M=N^{\perp}$, $Y_1=R^m$ и

оператор $A = f'(a)\circ\varphi : R^m \to R^m$ имеет матрицу ($\langle p_i, p_j \rangle$) $i, j = \overline{1,m}$.

3. $X = l^1$, $Y = R^m$, $f'_*(a)$ определяется с помощью независимых векторов $p_i \in (l^1)^* = m$ (напомним, что $\|x\|_m = \sup_i |x_i|$). На этот раз ряды $\sum_j p_{ij}^2$ могут и не сходиться, и конструкция примера 2 явно не подходит. Для построения подпространства M найдём номер $n \geqslant m$ такой, что векторы $[p_i]_n = (p_{i1}, \dots, p_{in}, 0, 0, \dots, 0)$ линейно-независимы. Теперь можно показать, что топологическим дополнением к ядру N является, например, множество

$$M = \{x \mid x = \sum_{i=1}^{m} \lambda_i [p_i]_n,\ \lambda_i \in R\}.$$

Литература

1. КАНТОРОВИЧ Л.В. Функциональный анализ и прикладная математика. -"Успехи матем. наук", т.3, вып.6, 1948, с.89-185.
2. ALTMAN M. On the generalization of Newton's method. Bull. Pol. Ac. Sc., cl.III, 1957, v.5, N 8, p. 789-795.
3. ВЕРТГЕЙМ Б.А. О некоторых методах приближенного решения нелинейных функциональных уравнений. Автореф. канд. дисс., 1957,
4. ВЕРТГЕЙМ Б.А. Об условиях применения метода Ньютона. - "Докл. АН СССР", 1956, т.110, № 5, с. 719-722.
5. РАКОВЩИК Л.С. О методе Ньютона-Канторовича.- ЖВМ и МФ, 1968, т. 8, № 6, с. 1208-1217.

Поступила в ред.-изд. отд.
10. II. 1975 г.