Теорія оптимальних рішень

Описывается способ численного вычисления производной в точке для ее использования в методе оптимизации сложной системы, поведение которой изучается математическим аппаратом абстрактных пространств.

УДК 519.21

З.В. НЕКРЫЛОВА

О ЧИСЛЕННОМ ВЫЧИСЛЕНИИ ПРОИЗВОДНОЙ ОТОБРАЖЕНИЯ В ЗАДАННОЙ ТОЧКЕ

Введение. При оптимизации, управлении сложными системами, функционирование которых описывается математическим аппаратом бесконечномерных пространств, исследователь редко имеет дело с аналитическим выражением функционала цели. Чаще всего ему доступны наблюдения и измерения каких-то характеристик, описывающих поведение системы. Такие измерения могут представлять собой набор вектор-функций, заданных в фиксированном временном интервале, или даже их отдельные значения в фиксированные моменты времени.

В случае исследования стохастических систем такие наблюдения и измерения осуществляются над реализациями случайных параметров, включенных в описание поведения системы.

Одним из подходов нахождения оптимальных стратегий поведения систем в описанных ситуациях являются численные методы поиска. Остановимся на вопросах поиска, которые требуют умения вычислять производную в заданной точке исследуемого функционала, отображения.

В работе [1] описаны некоторые способы построения такой производной с помощью производной вспомогательного отображения, которое является заданным, дифференцируемым и обладает нужными свойствами в некоторой окрестности этой точки. Эти способы приводятся далее в виде утверждений.

Будем считать, что рассматриваемые пространства банаховы. Далее $C^n([t_0,t_1])$ обозначим пространство непрерывных отображений (вектор-функций) из $[t_0, t_1]$ в R^n с нормой вида $\|x(\circ)\|_{C^n} = \max\limits_{t\notin[t_0,t_1]} |x(t)|$, где $|x| = \left(\sum\limits_{i=1}^{n} \left(x^i\right)^2 \right)^{\frac{1}{2}}$. Пространство $C_m^n\left([t_0,t_1]\right)$ образовано m раз непрерывно дифференцируемыми отображениями отрезка $[t_0, t_1]$ в R^n. Норма в нем задается: $\|x(\circ)\|_{C_m^n} = \max\limits_{0\le i\le m} \|x^{(i)}(\circ)\|_{C^n}$.

Утверждение 1. Имеем отображение $h: R^n \to R^m$ вида $h(x) = \left(h_1(x),\ldots,h_m(x)\right)$, которое определено и непрерывно дифференцируемо в окрестности U точки x_0. Введем отображение $H_\tau\left(x(\circ)\right): C^n\left([t_0,t_1]\right) \to R^m$, определенное соотношением $H_\tau\left(x(\circ)\right) = h\left(x(\tau)\right)$, где τ – некоторая фиксированная точка отрезка $[t_0,t_1]$. Причем оно определено на множестве таких $x(\circ) \in C^n\left([t_0,t_1]\right)$, что $x(\tau) \in U$. Если при этом $x_0(\tau) = x_0$, то отображение H_τ дифференцируемо по Фреше в точке $x_0(\circ)$ и его производную можно вычислить следующим образом.

Из дифференцируемости $h(x)$ следует

$$h(x_0 + x) = h(x_0) + \left\langle h'(x_0), x \right\rangle + o\left(|x|\right),$$

(где символом $\langle \circ,\circ \rangle$ обозначено скалярное произведение в R^n), поэтому

$$H_\tau\left(x_0(\circ) + x(\circ)\right) = h\left(x_0(\tau)\right) + \left\langle h'\left(x_0(\tau)\right), x(\tau) \right\rangle + \mathrm{o}\left(|x(\tau)|\right).$$

Из введенных выше определений норм имеем $|x(\tau)| \le \|x(\circ)\|$, следовательно,

$$H_\tau\left(x_0(\circ) + x(\circ)\right) = H_\tau\left(x_0(\circ)\right) + \left\langle h'\left(x_0(\tau)\right), x(\tau) \right\rangle + \mathrm{o}\left(\|x(\circ)\|\right).$$

Это значит, что отображение H_τ дифференцируемо по Фреше в точке $x_0(\circ)$ и его производная равна $H'_\tau\left(x_0(\circ)\right) x(\circ) = \left\langle h'\left(x_0(\tau)\right), x(\tau) \right\rangle$.

Утверждение 2. Пусть $h_1(t,x),\ldots,h_m(t,x)$ действительные функции, определенные, непрерывные и непрерывно дифференцируемые по x в открытом множестве $U \in R \times R^n$. Обозначим $h(t,x) = \left(h_1(t,x),\ldots,h_m(t,x)\right)$ и предположим, что график непрерывной вектор-функции $x_0(t): [t_0,t_1] \to R^n$ принадлежит области U. Рассмотрим отображение $H: C_1^n\left([t_0,t_1]\right) \to C^m\left([t_0,t_1]\right)$, определенное соотношением $[H(x(\circ))](t) = h\left(t, x(t)\right),\ t_0 \le t \le t_1$. Покажем, что это отображение дифференцируемо по Фреше в точке $x_0(\circ)$ и вычислим производную.

Для этого напомним следующий результат. Пусть X и Y – банаховы пространства, а F непрерывное отображение окрестности U точки $x_0 \in X$ в Y. Предположим, что отображение F дифференцируемо по Гато в каждой точке множества U и при этом отображение $x \to F_{\tilde{A}}(x)$ из U в $L(X,Y)$ (пространство линейных непрерывных отображений из X в Y) непрерывно. Тогда отображение F дифференцируемо на U по Фреше и его производная равна производной по Гато.

Проверим дифференцируемость по Гато отображения H в некоторой окрестности точки $x_0(\circ)$ и покажем непрерывность этой производной. Множество U открыто, поэтому можно указать такое $\varepsilon > 0$, что из $|x_0(t) - x(t)| < \varepsilon$ следует, что $(t, x) \in U$. Если же $\|x(\circ) - x_0(\circ)\|_{C^n} < \varepsilon$, то, согласно определения дифференцируемости по Гато, имеем

$$\lim_{\lambda \to 0}\left[\frac{H\left(x(\circ) + \lambda z(\circ)\right) - H\left(x(\circ)\right)}{\lambda}\right](t) = h'_x\left(t, x(t)\right) z(t),$$

т. е. производная $H(x)$ по Гато имеет вид

$$\left[H'_{\Gamma}\left(x(\circ)\right) z(\circ)\right](t) = h'_x\left(t, x(t)\right) z(t).$$

Непрерывность отображения $x(\circ) \to H_{\tilde{A}}\left(x(\circ)\right)$ сразу же следует из непрерывности отображения $((t,x) \to h'_x(t,x),$ что отмечалось при задании $h(t,x)$. Таким образом, отображение H дифференцируемо по Фреше и его производная в точке $x_0(\circ)$ имеет вид $\left[H'\left(x_0(\circ)\right) z(\circ)\right](t) = h'_x\left(t, x_0(t)\right) z(t)$.

Для нахождения производной при решении задач оптимального управления будет полезным следующее утверждение.

Утверждение 3. Пусть $h_1(t,x,u), \ldots, h_m(t,x,u)$ – действительные функции, определенные, непрерывные и непрерывно дифференцируемые по x и u в некоторой области U пространства $R \times R^n \times R^r$. Положим

$$h(t,x,u) = \left(h_1(t,x,u), \ldots, h_m(t,x,u)\right)$$

и предполагается, что вектор-функции $x(\circ) \in C_1^n\left([t_0,t_1]\right)$ и $u(\circ) \in C_1^r\left([t_0,t_1]\right)$ таковы, что $\left(t, x_0(t), u_0(t)\right) \in U$ для всех $t \in [t_0, t_1]$. Для отображения $H:\ C_1^n\left([t_0,t_1]\right) \times C_1^r\left([t_0,t_1]\right) \to C^m\left([t_0,t_1]\right)$ вида

$$\left[H\left(x(\circ), u(\circ)\right)\right](t) = h\left(t, x(t), u(t)\right), \quad t_0 \le t \le t_1$$

рассуждения, аналогичные приведенным в утверждении 2, показывают, что H дифференцируемо по Фреше в точке $(x_0(\circ), u_0(\circ))$ и его производная имеет вид

$$\left[H'\left(x_0(\circ), u_0(\circ)\right)\left(z(\circ), y(\circ)\right)\right](t) = h'_x\left(t, x_0(t), u_0(t)\right) z(t) + h'_u\left(t, x_0(t), u_0(t)\right) y(t).$$

Из приведенных утверждений напрашивается вывод, что для вычисления производной в заданной точке важно уметь описать поведение исследуемого отображения в этой точке и ее окрестности, т. е. построить вспомогательное отображение h.

Таким образом, чтобы использовать приведенные утверждения для численного вычисления производной исследуемого отображения $H\left(x(\circ)\right)$, необходимо быть уверенным в том, что проведенные наблюдения и измерения над исследуемым объектом дают возможность утверждать, что в окрестности выбранной точки функционирование системы хорошо представимо отображением h, а значения измерений наблюдаемых характеристик можно считать его аргументами. Для утверждения 1 такими характеристиками будут величины, описывающие состояние системы в фиксированный момент τ временного отрезка $[t_0, t_1]$, для утверждения 2 и 3 – это кусок траектории в виде отрезка кривой при $t \in [t_0, t_1]$.

Рассмотрим подробнее вычисление производной вспомогательного отображения h. В утверждении 1 производная $M\left(H_\tau(x(\circ,\omega)) / \{\overline{x(\tau)}\}\right) = h\left(x(\tau)\right)$ – это вектор-функция в R^m, поэтому для ее вычисления можно использовать классический конечно-разностный подход. Для указанного $\delta > 0$, зависящего от имеющегося набора измерений $\left\{\overline{x(\tau)}\right\}$, на котором определена функция h в окрестности U точки $x_0(\tau)$, можно найти такое $x(\tau)$, что $\left|x_0(\tau) - x(\tau)\right| < \delta$. Это δ будет характеризовать степень приближения $M\left(H_\tau(x(\circ,\omega)) / \{\overline{x(\tau)}\}\right) =$ $= h\left(x(\tau)\right)$ и ее конечно-разностного аналога, который можно вычислить для каждой $h'_i\left(x_0(\tau)\right), i = 1, \ldots, m$ как

$$\tilde{h}'_i\left(x_0(\tau)\right) = h_i\left(x(\tau)\right) - h_i\left(x_0(\tau)\right).$$

В случае утверждения 2 для выбранного $\delta > 0$ из перечня измерений $\left\{\overline{x(t)},\ t_0 \le t \le t_1\right\}$ в окрестности точки $x_0(t)$ можно указать такое $x(t)$, что $\left\|x(t) - x_0(t)\right\| < \delta$. Тогда

$$\tilde{h}'_x\left(t, x_0(t)\right) = h\left(t, x(t)\right) - h\left(t, x_0(t)\right).$$

Для исследования системы, функционирующей в стохастических условиях, будем считать, что оптимизируется ожидаемое поведение системы, т. е. $E\left(H\left(x((\circ),\omega)\right)\right)$, где ω обозначает случайный элемент. Задав значение случайного элемента, воздействующего на поведение системы, проводим наблюдения и измерения изучаемых характеристик в избранной точке, чтобы иметь возможность воспользоваться отображением h. А затем вычисляем производную в этой точке отображения $E\left(H\left(x((\circ),\omega)\right)\right)$, используя полученные реализации $H\left(x(\circ)\right)$. Конкретнее это можно представить следующим образом. Рассмотрим это на примере утверждения 1.

Утверждение 4. Пусть h имеет те же свойства, что и в утверждении 1. Для вычисления производной отображения $E\left(H\left(x((\circ),\omega)\right)\right)$ применим подход, предложенный в [2], когда вместо значений отображения используются значения реализаций $H\left(x(\circ)\right)$. Для этого введем отображение вида $H_\tau\left(x(\circ)\right)=h\left(x(\tau)\right)$, которое выполняется с вероятностью 1 для значений $\left\{\overline{x(\tau)}\right\}$ и определено на множестве таких $x(\circ)$, что $x(\tau)\in U$. Рассуждения, аналогичные приведенным в утверждении 1, дают возможность утверждать, что $H'_\tau(x_0(\circ))\,x(\tau)=\left\langle h'\left(x_0(\tau)\right),x(\tau)\right\rangle$ с вероятностью 1, откуда $E\left(H'_\tau(x_0(\circ))x(\tau)\big/\overline{\{x(\tau)\}}\right)=\left\langle h'\left(x_0(\tau)\right),x(\tau)\right\rangle$.

Таким образом, вычисление производной отображения $EH\left(x((\circ),\omega)\right)$ с помощью отображения h фактически приводит к подходу, известному в стохастическом программировании как вычисление стохастического градиента [3]. Поэтому производную от h в этом случае можно назвать стохастическим дифференциалом отображения $E\left(H\left(x((\circ),\omega)\right)\right)$.

Заключение. Работа посвящена построению численных методов в абстрактных пространствах. В частности, приведены способы вычисления производной в точке при различных заданиях значений наблюдений и измерений над характеристиками изучаемой сложной системы.

З.В. Некрилова

ПРО ЧИСЕЛЬНЕ ОБЧИСЛЕННЯ ПОХІДНОЇ ВІДОБРАЖЕННЯ У ЗАДАНІЙ ТОЧЦІ

Описується спосіб чисельного обчислення похідної у точці з метою її використання в методі оптимізації складної системи, поведінка якої вивчається математичним апаратом абстрактних просторів.

Z.V. Nekrylova

ABOUT NUMERICAL COMPUTATION OF DERIVATIVE IN THE POINT FOR MAPPING

The method of numerical computation for derivative in the point is described to use it in optimization method for complex system. The system behaviour is described with abstract mathematics instrument.

1. *Иоффе А.Д., Тихомиров В.М.* Теория экстремальных задач. – М.: Наука, 1974. – 479 с.
2. *Вазан М.* Стохастическая аппроксимация. – М.: Мир, 1972. – 295 с.
3. *Ермольев Ю.М.* Методы стохастического программирования. – М.: Наука, 1976. – 239 с.

Получено 03.02.2015