

УДК 519.3

ВЫЧИСЛЕНИЕ КВАЗИДИФФЕРЕНЦИАЛОВ И ЭКЗОСТЕРОВ НЕГЛАДКИХ ФУНКЦИЙ

М. Ю. Андрамонов¹

Предлагается общий метод вычисления квазидифференциалов и кодифференциалов, основанных на их аппроксимации многогранниками. Данный способ может применяться для решения задач негладкой оптимизации, в частности в методе Ньютона и методе наискорейшего спуска для кусочно-дифференцируемых функций. Точность аппроксимации зависит от количества векторов, для которых вычисляется производная по направлениям.

Ключевые слова: оптимизация, квазидифференциалы, метод Ньютона, метод наискорейшего спуска, квазидифференциальное исчисление, субдифференциалы.

Квазидифференциальное исчисление играет большую роль при решении задач с недифференцируемыми функциями. Такие задачи возникают в негладком анализе [1, 6], в частности, при решении задач оптимизации негладким методом Ньютона [1] и в невыпуклой оптимизации [3, 8].

В данной статье предлагается общий метод приближенного вычисления квазидифференциалов и кодифференциалов. Считается, что функция достаточно сложная и что можно вычислить ее значения и производную по направлениям, но затруднительно найти квазидифференциал аналитически (функция не обязательно задана аналитическим выражением).

1. Постановка задачи. Пусть функция f квазидифференцируема в точке x евклидова пространства $\mathbb{R}^n$. Это означает, что для любого x существует конечная производная $f'(x, s)$ по направлениям, причем

$$f'(x, s) = \max_{v \in \underline{df}(x)} (s, v) + \min_{w \in \overline{df}(x)} (s, w), \quad (1)$$

где s — произвольное направление из $\mathbb{R}^n$ и круглые скобки означают скалярное произведение. Пара множеств $Df(x) = [\underline{df}(x); \overline{df}(x)]$ называется квазидифференциалом функции $f(x)$ в точке x [1, 6], а сами множества — субдифференциалом и супердифференциалом соответственно.

Для вычисления квазидифференциалов функций типа максимума и минимума и их комбинаций удобно использовать квазидифференциальное исчисление [1, 6]. К сожалению, необходимость точного вычисления алгебраической суммы или выпуклой оболочки множеств и в этом случае затрудняет точное вычисление квазидифференциала негладкой функции. Поэтому разумно использовать другие подходы, позволяющие аппроксимировать квазидифференциал на основе значений функции и производных по направлениям.

Подобным образом возникает задача вычисления кодифференциала [6], или такой пары множеств $[df(x); \overline{df}(x)]$, каждое из которых принадлежит $\mathbb{R}^{n+1}$, что

$$f(x + \Delta) = f(x) + \max_{(a, b) \in \underline{df}(x)} (a + (b, \Delta)) + \min_{(g, h) \in \overline{df}(x)} (g + (h, \Delta)) + o(\Delta),$$

где $o(\Delta)/\|\Delta\| \rightarrow 0$ при $\|\Delta\| \rightarrow 0$ и $\Delta \in \mathbb{R}^n$. Данный подход имеет то преимущество, что можно определять непрерывно кодифференцируемые функции, хотя квазидифференциальное отображение является разрывным.

2. Вычисление квазидифференциала по вершинам. Будем искать аппроксимацию квазидифференциала в виде пары выпуклых многогранников, которые заданы своими вершинами. А именно, пусть $V = \text{co}\{v_1, v_2, \dots, v_k\}$, $W = \text{co}\{w_1, w_2, \dots, w_k\}$. Здесь co означает выпуклую оболочку векторов. Для простоты считаем, что количество вершин у данных многогранников совпадает (неизвестные вершины v_i, w_j могут совпадать).

Пусть $s_1, s_2, \dots, s_k$ — совокупность направлений из единичной сферы S пространства $\mathbb{R}^n$. Пусть v_k — вершина многогранника V , в которой достигается максимум функции (s_k, v) , а w_k — вершина многогранника W , в которой достигается минимум функции (s_k, w) . Направления s_k выбираются так, чтобы

¹ Казанский государственный университет, НИИ математики и механики им. Н. Г. Чеботарева, Университетская ул., 17, 420008, Казань; e-mail: myu21@mail.ru

равномерно заполнялась единичная сфера, а искомые вершины v_i, w_j вычисляются как решения следующей системы линейных уравнений и неравенств:

$$\begin{aligned} f'(x, s_i) &= (s_i, v_i + w_i), \quad i = 1, 2, \dots, k; \\ (s_1, v_1) &\geq (s_1, v_i) \quad \forall i \neq 1; \quad (s_2, v_2) \geq (s_2, v_i) \quad \forall i \neq 2; \quad (s_k, v_k) \geq (s_k, v_i) \quad \forall i \neq k; \\ (s_1, w_1) &\leq (s_1, w_i) \quad \forall i \neq 1; \quad (s_2, w_2) \leq (s_2, w_i) \quad \forall i \neq 2; \quad (s_k, w_k) \leq (s_k, w_i) \quad \forall i \neq k; \\ |v_{ir}| &\leq C, \quad i = 1, 2, \dots, k, \quad r = 1, 2, \dots, n; \quad |w_{ir}| \leq C, \quad i = 1, 2, \dots, k, \quad r = 1, 2, \dots, n. \end{aligned} \quad (2)$$

Здесь $C > 0$ — положительная константа, ограничивающая по норме суб- и супердифференциал и v_{ir}, w_{ir} — r -е компоненты векторов v_i, w_i соответственно.

Если известно, что функция суб- или супердифференцируема, то система (2) упрощается, поскольку можно положить все переменные v_i или w_j равными нулю. Например, для субдифференцируемой функции выполнено $w_j \equiv 0$, и поэтому имеем:

$$\begin{aligned} f'(x, s_i) &= (s_i, v_i), \quad i = 1, 2, \dots, k; \\ (s_1, v_1) &\geq (s_1, v_i) \quad \forall i \neq 1; \quad (s_2, v_2) \geq (s_2, v_i) \quad \forall i \neq 2; \\ (s_k, v_k) &\geq (s_k, v_i) \quad \forall i \neq k; \quad |v_{ir}| \leq C, \quad i = 1, 2, \dots, k, \quad \forall r = 1, 2, \dots, n. \end{aligned} \quad (3)$$

При стремлении C к плюс бесконечности система (2) или (3) имеет бесконечно много решений в предположении квазидифференцируемости или субдифференцируемости функции. Действительно, выберем любой квазидифференциал $[Q_1; Q_2]$; пусть

$$v_i \in \operatorname{Arg} \max_{v \in Q_1} (s_i, v), \quad i = 1, 2, \dots, k; \quad w_j \in \operatorname{Arg} \min_{w \in Q_2} (s_j, w), \quad j = 1, 2, \dots, k.$$

Тогда из условия (1) следует, что такая совокупность вершин является решением системы (2). При этом можно выбрать $C = \max(\max_{x \in Q_1} \|x\|, \max_{x \in Q_2} \|x\|)$. Применение данного подхода особенно удобно для задач малой размерности и для функций типа максимума (минимума) с небольшим количеством активных функций в каждой точке. При увеличении размерности лучше вычислять экзостер функции [2] либо использовать нелинейную аппроксимацию [6]. Заметим, что к системе (2) можно добавить любую целевую функцию. Такой прием следует применять, если желательно найти субдифференциал функции, но неизвестно заранее, является ли функция субдифференцируемой. Тогда следует добавить условие неотрицательности w_{ir} и минимизировать на допустимом множестве линейную функцию $\sum_{i=1}^k \sum_{r=1}^n w_{ir}$. В случае субдифференцируемости функции все величины w_{ir} окажутся нулевыми.

Множества V, W зависят от выбора направлений s_i . Докажем следующую теорему об аппроксимации квазидифференциала.

Теорема 1. Пусть $k \rightarrow +\infty$ и $\max_{s \in S} \min_{i=1,2,\dots,k} \|s - s_i\| \rightarrow 0$, где S — единичная сфера в $\mathbb{R}^n$. Пусть также $C > 0$ выбрано таким образом, что система (2) имеет решение. Тогда для соответствующего решения $[V(k); W(k)]$ системы (2) при количестве направлений k выполняется соотношение

$$\max_{s \in S} \left| f'(x, s) - \max_{v \in V(k)} (s, v) - \min_{w \in W(k)} (s, w) \right| \rightarrow 0.$$

Доказательство. По условию система имеет решение для любого k и множества $V(k), W(k)$ ограничены по норме в совокупности. Перепишем соотношение в следующем виде:

$$\begin{aligned} u &= \left| f'(x, s) - \max_{v \in V(k)} (s, v) - \min_{w \in W(k)} (s, w) \right| = \left| f'(x, s) - f'(x, s_j) + f'(x, s_j) - \right. \\ &\quad \left. - \max_{v \in V(k)} (s, v) - \min_{w \in W(k)} (s, w) + \max_{v \in V(k)} (s_j, v) - \max_{w \in W(k)} (s_j, w) + \min_{w \in W(k)} (s_j, w) - \min_{w \in W(k)} (s_j, w) \right|, \end{aligned}$$

где s_j — ближайшее к s направление из заданного набора векторов единичной сферы. Очевидно, что

$$\begin{aligned} u &\leq \left| f'(x, s) - f'(x, s_j) \right| + \left| f'(x, s_j) - \max_{v \in V(k)} (s_j, v) - \min_{w \in W(k)} (s_j, w) \right| + \\ &\quad + \left| \max_{v \in V(k)} (s, v) - \max_{v \in V(k)} (s_j, v) \right| + \left| \min_{w \in W(k)} (s, w) - \min_{w \in W(k)} (s_j, w) \right|. \end{aligned}$$

Первое слагаемое стремится к нулю в силу квазидифференцируемости f (тогда производная непрерывна относительно направления). Второе равно нулю по построению системы (2), поскольку направление s_j в нее входит.

Поскольку множества $V(k)$, $W(k)$ ограничены в совокупности, соответствующие маргинальные функции $\phi_1(x) = \max_{v \in V(k)} (x, v)$ и $\phi_2(x) = \max_{w \in W(k)} (x, w)$ непрерывны. Отсюда и последние два слагаемых стремятся к нулю. Теорема доказана.

Преимуществом данного подхода является простота реализации. Система (2) может, в частности, решаться симплексным методом. Устойчивость по отношению к ошибкам округления достигается выбором направлений s_i . Меняя набор направлений, можно добиться хорошей обусловленности матрицы ограничений.

3. Вычисление экзостеров. В ряде случаев удобнее аппроксимировать производные по направлениям не квазидифференциалом, а так называемым экзостером [2].

Определение. Семейство множеств E^* называется нижним экзостером функции f в точке x , если выполняется равенство $f'(x, s) = \min_{U \in E^*} \max_{v \in U} (s, v)$.

Будем искать данное семейство в виде $E^* = \{E_1, E_2, \dots, E_k\}$ для пучка направлений s_i при значениях $i = 1, 2, \dots, k$. Запишем систему вида

$$f'(x, s_i) = \min_{j=1,2,\dots,k} \max_{v \in E_j} (s_i, v). \quad (4)$$

Пусть E_j — многогранник $E_j = \{v_{j1}, v_{j2}, \dots, v_{jn_j}\}$ и n_j — число его вершин. Мы можем переписать (4) в форме

$$f'(x, s_i) = \min_{j=1,2,\dots,k} \max_{v_{jr} \in E_j} (s_i, v_{jr}), \quad i = 1, 2, \dots, k. \quad (5)$$

Здесь v_{ji} — вершина многогранника E_j , в которой достигается максимум линейной формы (s_i, v) . Кроме того, следует добавить ограничение по норме $\|v_{jr}\| \leq C \forall j, r$. В результате получим задачу минимаксной алгебры (дизъюнктивного программирования), которая может решаться методами из [4, 5, 7].

Применяя технику доказательства, аналогичную использованной в теореме 1, можно доказать, что для данного вида аппроксимации получается приближение производной по направлениям с любой наперед заданной точностью.

Теорема 2. Пусть $k \rightarrow +\infty$ и $\max_{s \in B} \min_{i=1,2,\dots,k} \|s - s_i\| \rightarrow 0$, где S — единичная сфера в $\mathbb{R}^n$. Пусть также система (5) имеет решение. Тогда для соответствующих решений v_{jr} системы (5) при количестве направлений k выполняется соотношение $|f'(x, s) - \min_{j=1,2,\dots,k} \max_{v_{jr} \in E_j} (s, v_{jr})| \rightarrow 0$.

Доказательство. Аналогично доказательству теоремы 1.

4. Приближенное вычисление кодифференциала по вершинам. Для вычисления кодифференциала необходимо использовать более сложную модель, выбирая не направления единичной сферы, а точки из окрестности x .

Вначале дадим следующее определение.

Определение. Множество $\underline{df}(x); \overline{df}(x)$ называется строгим (α, μ) -кодифференциалом, если выполняется

$$\left| f(x + \Delta) - f(x) - \max_{(a,b) \in \underline{df}(x)} (a + (b, \Delta)) - \min_{(p,q) \in \overline{df}(x)} (p + (q, \Delta)) \right| \leq \mu \|\Delta\|^\alpha,$$

где $\mu > 0$, $\alpha > 1$ для всех $\Delta \in \mathbb{R}^n$, меньших по норме некоторого заданного $\varepsilon > 0$.

Для получения результатов о сходимости приближений нужно предполагать, например, что функция обладает строгим (α, μ) -кодифференциалом, поскольку в явном виде включить $o(\Delta)$ в численную модель невозможно.

Выбираются вектора малой нормы Δ_i , $i = 1, 2, \dots, k$. На практике k может быть очень большим, поэтому кодифференциал вычислять сложнее, чем квазидифференциал. Незвестный кодифференциал будем искать в виде $[U; W]$, где

$$U = \text{co}((u_1, v_1), (u_2, v_2), \dots, (u_k, v_k)), \quad W = \text{co}((y_1, w_1), (y_2, w_2), \dots, (y_k, w_k)), \quad u_i, y_i \in \mathbb{R}, \quad v_i, w_i \in \mathbb{R}^n.$$

Система уравнений и неравенств имеет следующий вид:

$$\begin{aligned} f(x + \Delta_i) - f(x) - u_i - (v_i, \Delta_i) - y_i - (w_i, \Delta_i) &\leq \mu \|\Delta_i\|^\alpha, \\ f(x + \Delta_i) - f(x) - u_i - (v_i, \Delta_i) - y_i - (w_i, \Delta_i) &\geq -\mu \|\Delta_i\|^\alpha, \\ u_i - (v_i, \Delta_i) &\geq u_j - (v_j, \Delta_j), \quad i \neq j, \quad y_i - (w_i, \Delta_i) \leq y_j - (w_j, \Delta_j), \quad i \neq j, \\ |u_i| &\leq C, \quad |y_i| \leq C, \quad \|v_i\| \leq C, \quad \|w_i\| \leq C, \end{aligned} \quad (6)$$

где константа $C > 0$ ограничивает сверху норму элементов кодифференциала.

Пусть система (6) имеет решение. Тогда ее решения стремятся к строгому (α, μ) -кодифференциалу в некотором смысле.

Теорема 3. *Предположим, что функция $f(x)$ имеет строгий (α, μ) -кодифференциал в точке x . Пусть $[A; B]$ — некоторая пара множеств из $\mathbb{R}^{n+1}$, такая, что*

$$\left| f(x + \Delta) - f(x) - \max_{(a,b) \in A} (a + (b, \Delta)) - \min_{(p,q) \in B} (p + (q, \Delta)) \right| > \mu \|\Delta\|^\alpha \quad (7)$$

для некоторого Δ , $\|\Delta\| \leq \varepsilon$. Пусть C достаточно велико, система (6) разрешима и векторы Δ_i равномерно покрывают шар радиуса ε с центром в нуле при $k \rightarrow +\infty$. Тогда $[A; B]$ не является решением системы (6).

Более того, для достаточно мелкого покрытия все пары множеств $[A; B]$, удовлетворяющие условию

$$\left| f(x + \Delta) - f(x) - \max_{(a,b) \in A} (a + (b, \Delta)) - \min_{(p,q) \in B} (p + (q, \Delta)) \right| > (\mu + \mu_1) \|\Delta\|^\alpha \quad (8)$$

для некоторого Δ , $\|\Delta\| \leq \varepsilon$ при $\mu_1 > 0$, не являются решениями системы (6).

Замечание. Теорема означает, что достаточно плохие аппроксимации отбрасываются при измельчении покрытия окрестности x .

Доказательство вытекает из того факта, что множество Δ , для которого выполняются (7) или (8) и, следовательно, нарушается (6), открыто ввиду строгого неравенства и непрерывности маргинальных функций в левой части. Тогда условие (7) или (8) выполняется на некотором сколь угодно малом конечном подшаре Q из ε -окрестности x . Отсюда при измельчении покрытия некоторое Δ_i попадет в данный подшар Q , и в нем выполнится условие (6), противоречащее (7) или (8). Поэтому множества $[A; B]$ не будут более решениями системы. Теорема доказана.

Предположим, что функция $f(x)$ непрерывно кодифференцируема, т.е. кодифференциальное отображение непрерывно по Хаусдорфу, и при этом производная по направлениям допускает представление с помощью непрерывного кодифференциала. При этом не предполагается, что свободные члены равны нулю внутри максимума и минимума. Данные условия выполняются для непрерывно дифференцируемой функции.

Чтобы вычислить аппроксимацию непрерывного кодифференциала, мы выберем некоторые точки $x_1, x_2, \dots, x_N$ и в каждой точке x_i набор Δ_i из ее малой окрестности. Будем искать кодифференциал в каждой точке x_j в виде пары многогранников $V^{(j)}, W^{(j)}$ из $\mathbb{R}^{n+1}$, где при $u_i^j, y_i^j \in \mathbb{R}, w_i^j \in \mathbb{R}^n$ выполнено

$$V^{(j)} = \text{co}((u_1^j, v_1^j), (u_2^j, v_2^j), \dots, (u_k^j, v_k^j)), \quad W^{(j)} = \text{co}((y_1^j, w_1^j), (y_2^j, w_2^j), \dots, (y_k^j, w_k^j)). \quad (9)$$

Из условия кодифференцируемости при $\Delta = 0$ вытекает соотношение

$$\max_i u_i^j + \min_i y_i^j = 0, \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, N. \quad (10)$$

Условие (10) может быть сведено к совокупности линейных равенств и неравенств, связанных с помощью конъюнкций и дизъюнкций. Получается задача квазилинейной алгебры специального вида простой структуры. Для небольшого числа направлений ее можно решить точно перебором активных индексов. Использование данного условия в явном виде как ограничения усложняет решение задачи.

Условие непрерывности кодифференциального отображения по Хаусдорфу представить в виде, пригодном для численной реализации, сложно. Для произвольных точек x_1, x_2 мы можем требовать, чтобы расстояние между соответствующими аппроксимациями кодифференциалов было меньше заданного параметра $\gamma > 0$, который может при необходимости уменьшаться для проверки непрерывности кодифференциала.

Предположим для простоты, что функция $f(x)$ непрерывно гиподифференцируема, тогда все w_r^1, w_r^2 равны нулю. Получаем условия

$$\max_r |u_r^1 - u_r^2| \leq \gamma, \quad r = 1, 2, \dots, k, \quad \max_r \|v_r^1 - v_r^2\| \leq \gamma, \quad r = 1, 2, \dots, k. \quad (11)$$

Используя норму $\|\cdot\|_{+\infty}$ и раскрывая максимум в норме в (11), мы вновь получаем задачу линейного дизъюнктивного программирования (линейное программирование со связками ИЛИ между ограничениями), которая имеет сложность задачи целочисленного программирования, но в большинстве случаев эффективнее решается на практике методами релаксации и ветвей и границ [4,7].

Подставляя любые элементы непрерывного кодифференциального отображения (значения данного многозначного отображения в разных точках), получаем, что они дают решение системы (9). Если кодифференциальное решение разрывно, то при достаточно малом γ решения системы (9) не существует. Аналогичным способом можно вычислять кодифференциалы второго и более высоких порядков.

Автор выражает искреннюю благодарность В. Ф. Демьянову и А. М. Рубинову за постоянную поддержку и ряд ценных пожеланий.

СПИСОК ЛИТЕРАТУРЫ

1. Демьянов В.Ф. Теорема о неподвижной точке в негладком анализе и ее применение. СПб: Изд-во Санкт-Петербургского ун-та, 1996.
2. Demyanov V.F. Exhausters and convexifiers: new tools in nonsmooth analysis. Quasi-differentiability and related topics. Dordrecht: Kluwer Academic Publishers, 2001.
3. Andramonov M. Yu., Rubinov A.M., Glover B.M. Cutting angle methods in global optimization // Applied Mathematics Letters. May 1999.
4. Balas E. Disjunctive programming // Annals of Discrete Mathematics. 1979. **5**. 3–51.
5. Beaumont F. Algorithm for disjunctive programming problems // European Journal of Operational Research. 1990. **42**, N 3. 362–371.
6. Demyanov V.F., Rubinov A.M. Constructive non-smooth analysis. Frankfurt: Peter Lang, 1995.
7. Grossmann I., Lee S. New algorithms for generalized nonlinear disjunctive programming // Computers and Chemical Engineering. 2000. **24**. 2125–2141.
8. Rubinov A.M. Abstract convexity and global optimization. Dordrecht: Kluwer Academic Publishers, 2000.

Поступила в редакцию
13.06.2006