N1

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ДАЛЬНЕВОСТОЧНОЕ ОТДЕЛЕНИЕ
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ

РОССИЙСКАЯ АКАДЕМИЯ НАУК

ДАЛЬНЕВОСТОЧНЫЙ МАТЕМАТИЧЕСКИЙ СБОРНИК

1
1995

УДК 519.853.3

©1994 Е.А. Нурминский

A – СУБГРАДИЕНТ ВЫПУКЛЫХ ФУНКЦИЙ. ОПРЕДЕЛЕНИЕ И ОСНОВНЫЕ СВОЙСТВА

Предложено новое понятие субдифференциала, позволяющее дифференцировать вне области определения собственной выпуклой функции. Новый субдифференциал наследует основные свойства классического субдифференциала, позволяет сформулировать условия оптимальности и вводит новое правило дифференцирования для распространенного класса частично определенных выпуклых функций.

§1. Обозначения

E^n — n-мерное евклидовое пространство векторов со скалярным произведением xy. Норма определена обычным образом: $\|x\| = (xx)^{1/2}$.

Область определения функции f: $\operatorname{dom} f = \{x: f(x) < \infty\}$.

Индикаторная функция $\operatorname{Ind}_X(\cdot)$ множества X: $\operatorname{dom} \operatorname{Ind}_X = X, \operatorname{Ind}_X(x) \equiv 0$ для $x \in X$.

$\operatorname{epi} f$ — надграфик функции $f(\cdot)$: $\operatorname{epi} f = \{(\mu, x): \mu \geq f(x)\}$, $\operatorname{Epi} f$ обозначает замыкание $\operatorname{epi} f$.

$(X)_g$ — опорная функция множества X:

$$(X)_g = \sup_{x \in X} xg.$$

Операция инфинитезимальной конволюции функций f и g обозначается

$$\left(f \overset{+}{\vee} g\right)(x) = \inf_{x'+x''=x} \{f(x') + g(x'')\}.$$

§2. Кратчайшая история негладкого анализа и мотивации для последующего развития

Одним из побудительных мотивов для развития методов негладкой оптимизации является сведение задачи математического программирования

$$\min f(x, y),$$
$$g(x, y) \leq 0,$$

где неизвестные разбиты на две группы $x \in E^{n_x}$ и $y \in E^{n_y}$, к последовательной минимизации

$$\min_x \min_{y \in Y(x)} f(x, y) \equiv \min_x \psi(x), \tag{1}$$

где

$$Y(x) = \{y : g(x,y) \le 0\}.$$

Минимизация функции $\psi(x)$, определяемой (1) как

$$\psi(x) = \min_{y \in Y(x)} f(x,y),$$

может проводится так называемыми многоэтапными методами, учитывающими ее неявное определение.

Существенным усложнением при этом является отсутствие у $\psi(x)$ производных в обычном смысле слова, однако еще более существенным является то, что она определена лишь для тех x, для которых $Y(x)$ непусто. Формальное определение $\operatorname{dom}\psi$ как множества таких x, что $Y(x)$ непусто, неконструктивно и на практике вынуждает либо к его аппроксимации, либо к введению в модели предположений типа полной рекурсии для задач двухэтапного стохастического программирования, ведущих к непустоте $Y(x)$ для всех x.

В этих предположениях хорошо развито субдифференциальное исчисление, определяющее субградиент $\psi(x)$ в случае ее выпуклости и различные типы обобщенных градиентов [2 4], если функция ψ невыпукла. Развитое субдифференциальное исчисление используется для формулировки необходимых условий экстремума и построения в конечном итоге численных методов его нахождения.

Значительный интерес представляет ϵ-субдифференциальное исчисление [5, 11] представляющее собой регуляризованное расширение классического субдифференциала выпуклого анализа. E-субдифференциал непрерывен в метрике Хаусдорфа [6], даже Липшицего непрерывен [7] и в определенном смысле дифференцируем [8], что может быть использовано для развития методов типа Ньютона для задач негладкой выпуклой оптимизации [9, 10].

Эти соображения послужили исходной точкой для обобщения понятия ϵ-субдифференциала на случай частично определенных выпуклых функций. В результате удалось определить аналог этого понятия, названный α-субдифференциалом, определенный во всем пространстве и позволяющий развить соответетствующее дифференциальное исчисление и сформулировать условия оптимальности.

В дальнейшем можно надеяться на использование α-субдифференциала и для построения численных методов.

§3. Определение и основные свойства

Определение 1. Будем называть α-субдифференциалом $D_\alpha f(x)$ выпуклой функции f в точке x множество векторов g таких, что

$$yg - f(y) \le xg - \alpha \tag{2}$$

для всех $y \in \operatorname{dom} f$.

Заметим, что если формально положить $f(y) = \infty$ для $y \notin \operatorname{dom} f$, то неравенство (2) можно распостранить и вне области определения f, где оно тривиально выполняется.

Следующая теорема демонстрирует, что при весьма общих предположениях α-субдифференциал $D_\alpha f(x)$ существует и определен везде.

ТЕОРЕМА 1. *Пусть f – собственная выпуклая функция. Тогда для любого x существует α_x такое, что $D_\alpha f(x)$ непусто для всех $\alpha \le \alpha_x$.*

Доказательство. В условиях теоремы для произвольной точки x и ее некоторой окрестности U существует $F_x > -\infty$ такое, что $f(z) > F_x$ для $z \in U$. Положим $\alpha_x < F_x$ и покажем, что $D_\alpha f(x) \neq$ для $\alpha < \alpha_x$.

Действительно, так как $(\alpha_x, x) \notin \operatorname{epi} f$ и даже $(\alpha_x, x) \notin \operatorname{Epi} f$ по построению, то существует гиперплоскость $(-\pi, g)$ такая, что

$$-\pi\mu + gy \leq -\pi\alpha_x + gx \tag{3}$$

для произвольной пары $(\mu, y) \in \operatorname{epi} f$. Отсюда немедленно следует, что $\pi \geq 0$.

Если $\pi > 0$, то можно считать без потери общности, что $\pi = 1$ и $g \in D_\alpha f(x)$ по определению. Покажем, что такой вектор $(-\pi, g)$ с $\pi > 0$, удовлетворяющий (3), всегда существует.

В качестве подходящего кандидата можно рассмотреть, например, вектор, порождаемый проекцией точки $\tilde{x}_\alpha = (\alpha, x)$ на epi f:

$$\tilde{g}_\alpha = \tilde{x}_\alpha - P_{\operatorname{Epi} f}(\tilde{x}_\alpha) = (-\pi_\alpha, g_\alpha).$$

Здесь $P_{\operatorname{Epi} f}(\cdot)$ — оператор проекции на Epi f, то есть $P_{\operatorname{Epi} f}(\tilde{x}_\alpha) \in \operatorname{Epi} f$ и

$$\|P_{\operatorname{Epi} f}(\tilde{x}_\alpha) - \tilde{x}_\alpha\|^2 = \min_{\tilde{y} \in \operatorname{Epi} f} \|\tilde{y} - \tilde{x}_\alpha\|^2.$$

Так как $\tilde{x}_\alpha$ строго отделимо от Epi f, то

$$P_{\operatorname{Epi} f}(\tilde{x}_\alpha) - \tilde{x}_\alpha \neq 0. \tag{4}$$

Покажем, что для $\alpha < \alpha_x$, справедливо неравенство $\pi_\alpha > 0$.

Положим, для упрощения обозначений, $x = 0$, то есть, $\tilde{x}_\alpha = (\alpha, 0)$. Если $\pi_\alpha = 0$, то есть $\tilde{g}_\alpha = (0, -g_\alpha)$, то

$$P_{\operatorname{Epi} f}(\tilde{x}_\alpha) = (\alpha, -g_\alpha) \in \operatorname{Epi} f$$

по построению и

$$-yg_\alpha \geq \|g_\alpha\|^2 \tag{5}$$

для всех $(\mu, y) \in \operatorname{Epi} f$ по свойству проекции.

Для $\alpha' < \alpha$ таким же образом

$$-y'g_{\alpha'} \geq \|g_{\alpha'}\|^2 \tag{6}$$

для всех $(\mu', y') \in \operatorname{Epi} f$. Полагая в (5) $y = -g_{\alpha'}$, и в (6) $y' = -g_\alpha$ получаем, что

$$-g_{\alpha'}g_\alpha \geq \|g_\alpha\|^2 > 0, \qquad -g_{\alpha'}g_\alpha \geq \|g_\alpha\|^2 > 0.$$

Перемножив эти неравенства, получим

$$\|g_\alpha\|^2\|g_{\alpha'}\|^2 \leq (g_{\alpha'}g_\alpha)^2 \leq \|g_\alpha\|^2\|g_{\alpha'}\|^2,$$

откуда следует, что вектора g_α и $g_{\alpha'}$ коллинеарны, то есть $g_{\alpha'} = \lambda g_\alpha$. Подставляя это равенство в (6) для $y = g_\alpha$, получаем

$$\lambda\|g_\alpha\|^2 = \lambda^2\|g_\alpha\|^2$$

или $\lambda = 1$ так как $\lambda = 0$ исключено в силу (4).

Таким образом $g_{\alpha'} = g_\alpha$ для всех $\alpha' < \alpha < \alpha_x$ или $(\mu, -g_\alpha) \in \operatorname{Epi} f$ для всех $\mu < \alpha_x$, откуда следует, что $f(-g_\alpha) < C$ для произвольного C, что противоречит условиям теоремы. Теорема доказана.

Легко проверить следующие элементарные свойства $D_\alpha f(x)$.

1. $D_\alpha f(x)$ — выпуклое замкнутое множество.

2. Если для $x \in \operatorname{dom} f$ существует ϵ-субдифференциал $\partial_\epsilon f(x)$, то

$$D_{f(x)-\epsilon} f(x) = \partial_\epsilon f(x).$$

3. Если $x \in \operatorname{dom} f$, то для непустоты $D_\alpha f(x)$ достаточно, чтобы $\alpha < f(x)$. Пример $f : E^1 \to E^1$ такой, что $\operatorname{dom} f = \{x \in E^1, x \geq 0\}$ с $f(x) = -\sqrt{x}$ для $x \in \operatorname{dom} f$ показывает, что от строго неравенства (при $x = 0$) нельзя отказаться.

4. Если $0 \in D_\alpha f(x)$, то $\alpha \leq \inf_x f(x)$.

Свойство 4 играет роль условий оптимальности, выраженных в терминах $D_\alpha f(x)$.

Наиболее существенным следствием этих свойств является следующая лемма, аналогичная фейеровскому свойству классического субдифференциала:

ЛЕММА 3.1. *Если $0 \notin D_\alpha f(x)$, то $g(x^* - x) < 0$ для $g \in D_\alpha f(x)$ и произвольного x^* такого, что $f(x^*) = \inf_x f(x)$.*

Д о к а з а т е л ь с т в о. Из свойства 4 и предположений леммы следует, что $\alpha > \inf_x f(x) = f(x^*)$. По определению $gy - f(y) \leq gx - \alpha < gx - f(x^*)$. Подставляя в это неравенство $y = x^*$, получаем утверждение леммы.

Для развития α-субдифференциального исчисления наиболее существенную роль играет правило дифференцирорвания функций вида

$$f(x) = \phi(x) + \operatorname{Ind}_X(x), \tag{7}$$

где $\phi(x)$ — конечная выпуклая функция, а Ind_X — индикаторная функция некоторого выпуклого множества X.

ТЕОРЕМА 2. *Для $x \notin X$ и произвольного α существует λ_x такое, что для любого $\lambda > \lambda_x$*

$$g_\phi + \lambda g_X \in D_\alpha f(x) \tag{8}$$

для всех $g_\phi \in \partial\phi(x)$ и g_X, которое строго отделяет x и множество X. Множители $\lambda_x \geq 0$ таковы, что

$$\lambda_x(g_X x - (X)_{g_X}) \geq \alpha - \phi(x)$$

Д о к а з а т е л ь с т в о. Из определения $f(x)$ следует, что f^* имеет следующий вид [1]:

$$f^*(g) = \phi^*(g) \overset{+}{\vee} (\operatorname{Ind}_X)^*(g) = \inf_{g'+g''=g} \{\phi^*(g') + (\operatorname{Ind}_X)^*(g'')\}.$$

Принимая во внимание, что $(\operatorname{Ind}_X)^*(g'') = (X)_{g''}$ получаем

$$D_\alpha f(x) = \{g = g' + g'' : xg' - \phi^*(g') + xg'' - (X)_{g''} \geq \alpha\}.$$

Пусть $g' \in \partial\phi(x)$, тогда $xg' - \phi^*(g') = \phi(x)$ и, следовательно,

$$\partial\phi(x) + \{g'' : xg'' - (X)_{g''} \geq \alpha - \phi(x)\} \subset D_\alpha f(x).$$

Достаточно удовлетворить неравенству

$$xg'' - (X)_{g''} \geq \alpha - \phi(x),$$

выбрав g'' таким образом, что $xg'' - (X)_{g''} > 0$ и соответствующим образом нормализовав этот вектор, то есть определив $g_X = \lambda g''$, где $\lambda \geq \lambda_* \geq (\alpha - \phi(x))/(xg'' - (X)_{g''})$. Теорема доказана.

Принимая во внимание, что для $x \in \mathrm{dom}\, f = X$ α-субдифференциал $D_\alpha f(x)$ содержит $\partial_{\phi(x)-\alpha}\phi$, теорема 2 позволяет вычислить некоторые элементы $D_\alpha f(x)$ во всем пространстве.

В заключение автор выражает благодарность обсудившим изложенные выше результаты участникам международной конференции по методам декомпозиции и параллельным вычислениям для задач большой размерности, происходившей в Международном институте прикладного системного анализа, Лаксенберг, Австрия, 13-23 июня 1994 г.

Работа по этой теме также была частично поддержана грантом РФФИ 94-01-01771.

Список литературы

1. *Рокафеллар Р.Т.* Выпуклый анализ. М.: Мир, 1975.

2. *Кларк Ф.* Оптимизация и негладкий анализ. М.: Наука, 1988.

3. *Нурминский Е.А.* Численные методы решения детерминированных и стохастических минимаксных задач. Киев: Наукова Думка, 1979.

4. *Демьянов В.Ф., Рубинов А.М.* Основы негладкого анализа и квазидифференциальное исчисление. М.: Наука, 1990.

5. **Hiriart-Urruty J.-B.** ϵ-subdifferential calculus. Convex Analysis and Optimization / (J.-P. Auben and R. Vinter, eds.). Pitman. 1982. P. 43-92.

6. **Asplund E., Rockafellar R.T.** Gradients of convex functions // Trans. Amer. Math. Soc. 1969. Vol. 139. P. 443–467.

7. **Nurminski E.A.** On ϵ-subgradient mappings and their applications in nondifferentiable optimization. Working paper 78-58. 1978, IIASA, 2361 Laxenburg, Austria.

8. **Lemarechal C., Nurminski E.A.** Sur la differentiabilite de la fonction d'apui de sous-differentiel approche // C.r.Acad. Sci. A. 1980. Vol. 290. P. 855–858.

9. **Auslender A.** On the differential properties of the support function of the ϵ-subdifferential of a convex function // Mathematical Programming. 1982. Vol. 24, 3. P. 257–268.

10. **Nurminski E.A.** Toward Newton method in nondifferentiable optimization // Journal of Math. Reasearch and Exposition. 1992. Vol. 12, N 2. P. 159–163.

11. *Кусраев А.Г., Кутателадзе С.С.* Субдифференциальное исчисление. Новосибирск: Наука, 1987.

Ин-т прикладной математики
ДВО РАН

Поступило в редакцию
1.XII.1994