

Домашняя работа 4: Субдифференциалы

Срок сдачи: 8 декабря 2018 (суббота), 23:59

- 1 (Характеризация субдифференциала через сопряженную функцию) Пусть $f : E \rightarrow \mathbb{R}$ — функция, определенная на множестве E в евклидовом пространстве V . Пусть $x \in E$, и пусть $f^* : E_* \rightarrow \mathbb{R}$ — сопряженная функция с естественной областью определения E_* . Покажите, что

$$\partial f(x) = \{s \in E_* : \langle s, x \rangle = f^*(s) + f(x)\}.$$

Другими словами, вектор $s \in E_*$ является субградиентом функции f в точке x , если и только если неравенство Фенхеля–Юнга $\langle s, x \rangle \leq f^*(s) + f(x)$ переходит в равенство.

- 2 Для каждой из следующих функций $f : \mathbb{R}^n \rightarrow \mathbb{R}$ найдите субдифференциал ∂f . Изобразите схематично множество $\partial f(x)$ для $n = 2$ и заданных вариантов точек x .

(a) $f(x) := \|x\|_\infty$. Точки для рисунка: а) $(0, 0)$; б) $(1, 1)$; в) $(-1, -1)$; д) $(-1, 1)$.

(b) $f(x) := \sum_{1 \leq i < j \leq n} |x_i - x_j|$. Точки для рисунка: а) $(0, 0)$; б) $(1, 1)$; в) $(1, 0)$; д) $(0, 1)$.

(c) $f(x) := \sum_{i=1}^m \|x - a_i\|_2$, где a_1, \dots, a_m — попарно различные точки в \mathbb{R}^n . Параметры для рисунка: $m = 3$, $a_1 := (0, 1)$, $a_2 := (-1, 0)$, $a_3 := (1, 0)$. Точки: а) a_1 ; б) a_2 ; в) a_3 ; д) 0 .

- 3 Пусть $\lambda_{\max} : \mathbb{S}^n \rightarrow \mathbb{R}$ — функция максимального собственного значения, заданная в пространстве \mathbb{S}^n .

(a) Найдите субдифференциал $\partial \lambda_{\max}$. (*Подсказка:* Воспользуйтесь вариационным представлением λ_{\max} и формулой для субдифференциала максимума.)

(b) Покажите, что функция λ_{\max} дифференцируема в точке $X \in \mathbb{S}^n$ тогда и только тогда, когда максимальное собственное значение матрицы X является простым (т. е. имеет кратность 1). Чему равен градиент $\nabla \lambda_{\max}(X)$?

- 4 (Полярный конус) Пусть C — множество в евклидовом пространстве V . *Полярным конусом* C называется множество $C^\circ := \{s \in V : \langle s, x \rangle \leq 0 \text{ для всех } x \in C\}$.

(a) Покажите, что C° , действительно, является конусом, причем выпуклым и замкнутым.

(b) Выведите в качестве следствия из теоремы Фенхеля–Моро *теорему о биполярном конусе*: $C^{\circ\circ} = C$, если и только если C является выпуклым замкнутым конусом. (*Подсказка:* Рассмотрите индикаторную функцию C .)

- 5 (Нормальный конус) Пусть C — множество в евклидовом пространстве V . *Нормальным конусом* множества C в точке $x \in C$ называется множество $N_C(x) := (C - x)^\circ = \{s \in V : \langle s, y - x \rangle \leq 0 \text{ для всех } y \in C\}$. Покажите, что:

(a) Если $x \in \text{int}(C)$, то $N_C(x) = \{0\}$ (нормальный конус является нетривиальным только на границе).

(b) $\partial \delta_C(x) = N_C(x)$ (субдифференциал индикаторной функции равен нормальному конусу).

- 6 (Характеризация субдифференциала через нормальный конус) Пусть $f : E \rightarrow \mathbb{R}$ — функция, заданная на множестве E в евклидовом пространстве V , и пусть $x \in E$. Покажите, что

$$\partial f(x) = \{s \in V : (s, -1) \in N_{\text{Epi } f}(x, f(x))\}.$$

- 7 (Условие минимума функции на множестве) Пусть $f : E \rightarrow \mathbb{R}$ — выпуклая функция, заданная на множестве E в евклидовом пространстве. Пусть C — выпуклое подмножество множества E , такое, что $C \cap \text{int}(E) \neq \emptyset$, и пусть $x^* \in C$. Докажите следующее условие оптимальности:

$$x^* \in \underset{C}{\text{Argmin}} f \iff (-\partial f(x^*)) \cap N_C(x^*) \neq \emptyset.$$

В частности, если $x^* \in \text{int}(E)$, и f дифференцируема в x^* , то последнее условие эквивалентно вложению

$$-\nabla f(x^*) \in N_C(x^*).$$

(*Подсказка:* Примените теорему Моро–Рокафеллара для функции $f|_C = f + \delta_C$.)

8 (Примеры полярных конусов) Покажите, что:

- (a) Если H — линейное подпространство, то $H^\circ = H^\perp$ (полярный конус равен ортогональному дополнению).
- (b) $(\mathbb{R}_+^n)^\circ = \mathbb{R}_-^n$;
- (c) $(\mathbb{S}_+^n)^\circ = \mathbb{S}_-^n$;

9 (Примеры нормальных конусов) Пусть K — конус в евклидовом пространстве V , $a \in V$, $x \in a + K$. Покажите, что

$$N_{a+K}(x) = \{s \in K^\circ : \langle s, x - a \rangle = 0\}.$$

- (a) Если H — линейное подпространство V , то $N_{a+H}(x) = H^\perp$ для всех $x \in a + H$.
- (b) $N_{\mathbb{R}_+^n}(x) = \{s \in \mathbb{R}_-^n : \langle s, x \rangle = 0\}$ для всех $x \in \mathbb{R}_+^n$.
- (c) $N_{\mathbb{S}_+^n}(X) = \{S \in \mathbb{S}_-^n : \langle S, X \rangle = 0\}$ для всех $X \in \mathbb{S}_+^n$.

(Подсказка: Используйте характеризацию субдифференциала через сопряженную функцию.)

10 Пусть $X \in \mathbb{S}^n$ — симметричная матрица с собственными значениями $\lambda \in \mathbb{R}^n$. Пусть $X = Q \text{Diag}(\lambda) Q^T$ — спектральное разложение для некоторой ортогональной $n \times n$ матрицы Q . Докажите следующую формулу для проекции на конус \mathbb{S}_+^n :

$$P_{\mathbb{S}_+^n}(X) = Q \text{Diag}(\lambda^+) Q^T,$$

где λ^+ — поэлементная положительная срезка собственных значений. (Подсказка: Используйте условие минимума функции на множестве и формулу для нормального конуса к \mathbb{S}_+^n .)

11 (Проксимальный оператор) Пусть $f : E \rightarrow \mathbb{R}$ — функция, заданная на множестве E в евклидовом пространстве V . Проксимальным оператором функции f называется отображение $\text{Prox}_f : V \rightarrow E$, заданное как $\text{Prox}_f(x) := \text{Argmin}_{y \in E} \{f(y) + \frac{1}{2}\|y - x\|^2\}$. Известно, что если функция f выпуклая и замкнутая, то в каждой точке Prox_f состоит из одного элемента (как множество минимумов сильно выпуклой замкнутой функции). Покажите, что для $x \in V$, $u \in E$ и выпуклой функции f следующие утверждения эквивалентны:

- (a) $u \in \text{Prox}_f(x)$;
- (b) $x - u \in \partial f(u)$;
- (c) $f(v) \geq f(u) + \langle x - u, v - u \rangle$ для всех $v \in E$.

(Подсказка: Используйте теорему Моро–Рокафеллара.)

12 (Разложение Моро) Пусть $f : E \rightarrow \mathbb{R}$ — выпуклая замкнутая функция, заданная на множестве E в евклидовом пространстве V , и пусть $x \in V$. Покажите, что

$$x = \text{Prox}_f(x) + \text{Prox}_{f^*}(x).$$

(Подсказка: Используйте второй пункт предыдущей задачи и свойство $s \in \partial f(x) \Leftrightarrow x \in \partial f^*(s)$.)

13 Пусть $f : E \rightarrow \mathbb{R}$ — выпуклая замкнутая функция, заданная на множестве E в евклидовом пространстве V , и пусть $x, y \in V$. Покажите, что

$$\|\text{Prox}_f(x) - \text{Prox}_f(y)\|^2 \leq \langle \text{Prox}_f(x) - \text{Prox}_f(y), x - y \rangle.$$

В частности, проксимальный оператор является сжимающим:

$$\|\text{Prox}_f(x) - \text{Prox}_f(y)\| \leq \|x - y\|.$$

(Подсказка: Второй пункт является простым следствием первого. Для доказательства первого воспользуйтесь третьим пунктом из задачи про проксимальный оператор.)

Бонусные задачи

- 14 (Уточненная формула для $N_{\mathbb{S}_+^n}$) Пусть $A, B \in \mathbb{S}_+^n$. Покажите, что $\langle A, B \rangle = 0$, если и только если $AB = BA = 0$. Таким образом,

$$N_{\mathbb{S}_+^n}(X) = \{S \in \mathbb{S}_+^n : SX = 0\}$$

для всех $X \in \mathbb{S}_+^n$. (*Подсказка:* Покажите, что спектры матриц AB и $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ совпадают; затем воспользуйтесь тем, что $A^{\frac{1}{2}}BA^{\frac{1}{2}} \in \mathbb{S}_+^n$.)

- 15 (Коническое разложение Моро) Пусть K — выпуклый замкнутый конус в евклидовом пространстве V . Покажите, что для любого $x \in V$ справедливо однозначное представление в виде

$$x = x_K + x_{K^\circ},$$

где

$$x_K \in K, \quad x_{K^\circ} \in K^\circ, \quad \langle x_K, x_{K^\circ} \rangle = 0,$$

которое переходит в известное из линейной алгебры ортогональное разложение в случае, когда K является линейным подпространством. (*Подсказка:* Примените разложение Моро для индикаторной функции K .)

- 16 (Субдифференциал маргинальной функции) Пусть V и W — евклидовы пространства, E — множество в пространстве $V \oplus W$. Пусть $E_V := \{x : (x, y) \in E\}$ — тень множества E на пространство V , и пусть $E_x := \{y \in W : (x, y) \in E\}$ вертикальный срез множества E для $x \in E_V$ (таким образом, $E = \cup_{x \in E_V} (\{x\} \times E_x)$). Пусть $f : E \rightarrow \mathbb{R}$ — функция, такая, что $E_x \rightarrow \mathbb{R} : f(x, \cdot)$ ограничено снизу для всех $x \in E_V$, и пусть $g : E_V \rightarrow \mathbb{R}$ — маргинальная функция

$$g(x) := \inf_{y \in E_x} f(x, y),$$

полученная путем минимизации функции f по всевозможным допустимым значениям второго аргумента при фиксированном первом. Пусть $\bar{x} \in E_V$, и пусть $Y(\bar{x}) := \{y \in E_{\bar{x}} : g(\bar{x}) = f(\bar{x}, y)\}$ — множество точек, на которых достигается минимум (возможно пустое). Покажите, что:

- (а) Для любого $\bar{y} \in Y(\bar{x})$ выполнено

$$\partial g(\bar{x}) = \{s \in V : (s, 0) \in \partial f(\bar{x}, \bar{y})\}.$$

(В частности, последнее множество не зависит от выбора конкретного \bar{y} .)

- (б) Если функция f выпуклая, то маргинальная функция g также выпуклая. При этом если $\bar{x} \in \text{int}(E_V)$ и найдется $\bar{y} \in Y(\bar{x})$, такое, что $(\bar{x}, \bar{y}) \in \text{int}(E)$ и функция f дифференцируема в точке (\bar{x}, \bar{y}) , то маргинальная функция g дифференцируема в точке \bar{x} с градиентом

$$\nabla g(\bar{x}) = \nabla_x f(\bar{x}, \bar{y}),$$

где $\nabla_x f$ обозначает градиент f по первому аргументу.

- 17 (Производная расстояния до множества) Опираясь на результат предыдущей задачи, покажите, если C — выпуклое замкнутое множество в евклидовом пространстве V , и $d_C : V \rightarrow \mathbb{R}$ — функция $d_C(x) := \min_{y \in C} \|y - x\|$ расстояния до множества C , то d_C дифференцируема в каждой точке $x \notin C$, причем $d_C(x) > 0$ и

$$\nabla d_C(\bar{x}) = d_C(\bar{x})^{-1}(\bar{x} - P_C(\bar{x})),$$

где $P_C(x) := \operatorname{argmin}_{y \in C} \|y - x\|$ — проекция точки на множество.

- 18 (Сопряженная норма к операторной норме) Пусть $\|\cdot\| : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ — операторная норма, $A \in \mathbb{R}^{m \times n}$, $q := \min\{m, n\}$. Покажите, что

$$\|A\|_* = \sum_{i=1}^q \sigma_i(A).$$

Таким образом, сопряженная норма к операторной норме равна сумме сингулярных чисел (эта норма называется *ядерной нормой*). (*Подсказка:* Воспользуйтесь сингулярным разложением $A = \sum_{i=1}^q \sigma_i u_i v_i^T$ и оцените $|\langle A, X \rangle|$ сверху; затем покажите, что неравенство достигается для некоторого X .)

- 19 * (Субдифференциал ядерной нормы) Пусть $A \in \mathbb{R}^{m \times n}$ — матрица ранга r , и пусть $A = U \Sigma V^T$ — сингулярное разложение матрицы A , где $U \in \mathbb{R}^{m \times r}$, $V \in \mathbb{R}^{n \times r}$ — матрицы с ортонормированными столбцами, $\Sigma \in \mathbb{R}^{r \times r}$ — диагональная матрица с положительными элементами на диагонали. Покажите, что

$$\partial \|\cdot\|_*(A) = \{UV^T + W : W \in \mathbb{R}^{m \times n}; U^T W = 0; W V = 0; \|W\| \leq 1\}.$$

(*Подсказка:* Используйте характеристику субдифференциала через сопряженную функцию. Далее покажите, что неравенство Фенхеля–Юнга переходит в равенство, если и только если $S = UV^T + W$, где матрица W удовлетворяет указанным выше условиям.)