Методы оптимизации
Лекция 5: Выпуклые функции: непрерывность, дифференцируемость, L-гладкость

Александр Катруца

Физтех-школа прикладной математики и информатики
Московский физико-технический институт

5 октября 2020 г.

На прошлой лекции

- Выпуклые функции и их свойства

На прошлой лекции

- Выпуклые функции и их свойства
- Критерии выпуклости

На прошлой лекции

- Выпуклые функции и их свойства
- Критерии выпуклости
- Операции, сохраняющие выпуклость

На прошлой лекции

- Выпуклые функции и их свойства
- Критерии выпуклости
- Операции, сохраняющие выпуклость
- Неравенство Йенсена

План на эту лекцию

- В критериях выпуклости фигурируют градиент и гессиан

План на эту лекцию

- В критериях выпуклости фигурируют градиент и гессиан
- Однако вопрос их существования для выпуклой функции пока остался за кадром

План на эту лекцию

- В критериях выпуклости фигурируют градиент и гессиан
- Однако вопрос их существования для выпуклой функции пока остался за кадром
- Но сначала надо изучить непрерывность выпуклой функции

План на эту лекцию

- В критериях выпуклости фигурируют градиент и гессиан
- Однако вопрос их существования для выпуклой функции пока остался за кадром
- Но сначала надо изучить непрерывность выпуклой функции
- После чего рассмотрим дифференцируемость

План на эту лекцию

- В критериях выпуклости фигурируют градиент и гессиан
- Однако вопрос их существования для выпуклой функции пока остался за кадром
- Но сначала надо изучить непрерывность выпуклой функции
- После чего рассмотрим дифференцируемость
- И в конце обсудим L-гладкость: одно из ключевых понятий для анализа численных методов

Важное свойства строго выпуклой функции

Единственность точки минимума

Если функция f строго выпукла, то локальный минимум достигается в единственной точке.

Важное свойства строго выпуклой функции

Единственность точки минимума

Если функция f строго выпукла, то локальный минимум достигается в единственной точке.

Доказательство

Важное свойства строго выпуклой функции

Единственность точки минимума

Если функция f строго выпукла, то локальный минимум достигается в единственной точке.

Доказательство

- Пусть есть две точки $\mathbf{x}_1 \neq \mathbf{x}_2$ таких что $f(\mathbf{x}_1) = f(\mathbf{x}_2)$ является минимумом

Важное свойства строго выпуклой функции

Единственность точки минимума

Если функция f строго выпукла, то локальный минимум достигается в единственной точке.

Доказательство

- Пусть есть две точки $\mathbf{x}_1 \neq \mathbf{x}_2$ таких что $f(\mathbf{x}_1) = f(\mathbf{x}_2)$ является минимумом
- Рассмотрим точку $\mathbf{z}$ из отрезка $[\mathbf{x}_1, \mathbf{x}_2]$ и значение $f(\mathbf{z}) < \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2) = f(\mathbf{x}_2)$

Важное свойства строго выпуклой функции

Единственность точки минимума

Если функция f строго выпукла, то локальный минимум достигается в единственной точке.

Доказательство

- Пусть есть две точки $\mathbf{x}_1 \neq \mathbf{x}_2$ таких что $f(\mathbf{x}_1) = f(\mathbf{x}_2)$ является минимумом
- Рассмотрим точку $\mathbf{z}$ из отрезка $[\mathbf{x}_1, \mathbf{x}_2]$ и значение $f(\mathbf{z}) < \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2) = f(\mathbf{x}_2)$
- Тогда взяв λ достаточно близкое к 1 получим противоречие, с тем что $\mathbf{x}_2$ точка минимума

Непрерывность выпуклой функции

Теорема

Пусть $f : \mathcal{X} \to \mathbb{R}$ выпуклая функция и произвольная точка $\mathbf{a} \in \operatorname{relint}(\mathcal{X})$. Тогда существует относительно открытая окрестность $\mathcal{O}(\mathbf{a}) \subset \operatorname{dom}(f)$ и $M > 0$ такие что

$$|f(\mathbf{x}) - f(\mathbf{a})| \le M\|\mathbf{x} - \mathbf{a}\|_2,$$

для любой точки $\mathbf{x} \in \mathcal{O}(\mathbf{a})$. Из этого неравенства следует непрерывность выпуклой функции в точках относительной внутренности.

Непрерывность выпуклой функции

Теорема

Пусть $f : \mathcal{X} \to \mathbb{R}$ выпуклая функция и произвольная точка $\mathbf{a} \in \operatorname{relint}(\mathcal{X})$. Тогда существует относительно открытая окрестность $\mathcal{O}(\mathbf{a}) \subset \operatorname{dom}(f)$ и $M > 0$ такие что

$$|f(\mathbf{x}) - f(\mathbf{a})| \leq M\|\mathbf{x} - \mathbf{a}\|_2,$$

для любой точки $\mathbf{x} \in \mathcal{O}(\mathbf{a})$. Из этого неравенства следует непрерывность выпуклой функции в точках относительной внутренности.

Доказательство

Непрерывность выпуклой функции

Теорема

Пусть $f : \mathcal{X} \to \mathbb{R}$ выпуклая функция и произвольная точка $\mathbf{a} \in \mathrm{relint}\,(\mathcal{X})$. Тогда существует относительно открытая окрестность $\mathcal{O}(\mathbf{a}) \subset \mathrm{dom}\,(f)$ и $M > 0$ такие что

$$|f(\mathbf{x}) - f(\mathbf{a})| \leq M\|\mathbf{x} - \mathbf{a}\|_2,$$

для любой точки $\mathbf{x} \in \mathcal{O}(\mathbf{a})$. Из этого неравенства следует непрерывность выпуклой функции в точках относительной внутренности.

Доказательство

- Рассмотрим частный случай

Непрерывность выпуклой функции

Теорема

Пусть $f : \mathcal{X} \to \mathbb{R}$ выпуклая функция и произвольная точка $\mathbf{a} \in \operatorname{relint}(\mathcal{X})$. Тогда существует относительно открытая окрестность $\mathcal{O}(\mathbf{a}) \subset \operatorname{dom}(f)$ и $M > 0$ такие что

$$|f(\mathbf{x}) - f(\mathbf{a})| \leq M\|\mathbf{x} - \mathbf{a}\|_2,$$

для любой точки $\mathbf{x} \in \mathcal{O}(\mathbf{a})$. Из этого неравенства следует непрерывность выпуклой функции в точках относительной внутренности.

Доказательство

- Рассмотрим частный случай
  - $\mathcal{X}$ — открытое подмножество $\mathbb{R}^n$

Непрерывность выпуклой функции

Теорема

Пусть $f : \mathcal{X} \to \mathbb{R}$ выпуклая функция и произвольная точка $\mathbf{a} \in \mathrm{relint}\,(\mathcal{X})$. Тогда существует относительно открытая окрестность $\mathcal{O}(\mathbf{a}) \subset \mathrm{dom}\,(f)$ и $M > 0$ такие что

$$|f(\mathbf{x}) - f(\mathbf{a})| \leq M\|\mathbf{x} - \mathbf{a}\|_2,$$

для любой точки $\mathbf{x} \in \mathcal{O}(\mathbf{a})$. Из этого неравенства следует непрерывность выпуклой функции в точках относительной внутренности.

Доказательство

- Рассмотрим частный случай
  - $\mathcal{X}$ — открытое подмножество $\mathbb{R}^n$
  - $\mathbf{a} = 0$ и $f(0) = 0$

- Покажем, что $|f(\mathbf{x})| \le M\|\mathbf{x}\|_2$ для любого $\mathbf{x}$ из окрестности 0.

- Покажем, что $|f(\mathbf{x})| \leq M\|\mathbf{x}\|_2$ для любого $\mathbf{x}$ из окрестности 0.
- Рассмотрим множество $B_\infty(r) = \{\mathbf{x} \mid \|\mathbf{x}\|_\infty \leq r\}$ для такого $r > 0$, что $B_\infty(r) \subset \mathcal{X}$

- Покажем, что $|f(\mathbf{x})| \leq M\|\mathbf{x}\|_2$ для любого $\mathbf{x}$ из окрестности 0.
- Рассмотрим множество $B_\infty(r) = \{\mathbf{x} \mid \|\mathbf{x}\|_\infty \leq r\}$ для такого $r > 0$, что $B_\infty(r) \subset \mathcal{X}$
- $B_\infty(r)$ есть выпуклая оболочка 2^n точек (каких?) $\mathbf{y}_1, \ldots, \mathbf{y}_{2^n}$

- Покажем, что $|f(\mathbf{x})| \le M\|\mathbf{x}\|_2$ для любого $\mathbf{x}$ из окрестности 0.
- Рассмотрим множество $B_\infty(r) = \{\mathbf{x} \mid \|\mathbf{x}\|_\infty \le r\}$ для такого $r > 0$, что $B_\infty(r) \subset \mathcal{X}$
- $B_\infty(r)$ есть выпуклая оболочка 2^n точек (каких?) $\mathbf{y}_1, \ldots, \mathbf{y}_{2^n}$
- Пусть $\mathbf{z} = \sum_{i=1}^k \alpha_i \mathbf{y}_i$ произвольная точка из $B_\infty(r)$

- Покажем, что $|f(\mathbf{x})| \leq M\|\mathbf{x}\|_2$ для любого $\mathbf{x}$ из окрестности 0.
- Рассмотрим множество $B_\infty(r) = \{\mathbf{x} \mid \|\mathbf{x}\|_\infty \leq r\}$ для такого $r > 0$, что $B_\infty(r) \subset \mathcal{X}$
- $B_\infty(r)$ есть выпуклая оболочка 2^n точек (каких?) $\mathbf{y}_1, \dots, \mathbf{y}_{2^n}$
- Пусть $\mathbf{z} = \sum_{i=1}^k \alpha_i \mathbf{y}_i$ произвольная точка из $B_\infty(r)$
- Тогда $f(\mathbf{z}) \leq \sum_{i=1}^k \alpha_i f(\mathbf{y}_i) \leq \max_i f(\mathbf{y}_i)$

- Покажем, что $|f(\mathbf{x})| \le M\|\mathbf{x}\|_2$ для любого $\mathbf{x}$ из окрестности 0.
- Рассмотрим множество $B_\infty(r) = \{\mathbf{x} \mid \|\mathbf{x}\|_\infty \le r\}$ для такого $r > 0$, что $B_\infty(r) \subset \mathcal{X}$
- $B_\infty(r)$ есть выпуклая оболочка 2^n точек (каких?) $\mathbf{y}_1, \dots, \mathbf{y}_{2^n}$
- Пусть $\mathbf{z} = \sum_{i=1}^k \alpha_i \mathbf{y}_i$ произвольная точка из $B_\infty(r)$
- Тогда $f(\mathbf{z}) \le \sum_{i=1}^k \alpha_i f(\mathbf{y}_i) \le \max_i f(\mathbf{y}_i)$
- С другой стороны $\{\mathbf{y}_1, \dots, \mathbf{y}_{2^n}\} \subset B_\infty(r)$, а значит $\max_{\mathbf{z} \in B_\infty(r)} f(\mathbf{z}) \ge \max_i f(\mathbf{y}_i)$

- Покажем, что $|f(\mathbf{x})| \leq M\|\mathbf{x}\|_2$ для любого $\mathbf{x}$ из окрестности 0.
- Рассмотрим множество $B_\infty(r) = \{\mathbf{x} \mid \|\mathbf{x}\|_\infty \leq r\}$ для такого $r > 0$, что $B_\infty(r) \subset \mathcal{X}$
- $B_\infty(r)$ есть выпуклая оболочка 2^n точек (каких?) $\mathbf{y}_1, \ldots, \mathbf{y}_{2^n}$
- Пусть $\mathbf{z} = \sum_{i=1}^k \alpha_i \mathbf{y}_i$ произвольная точка из $B_\infty(r)$
- Тогда $f(\mathbf{z}) \leq \sum_{i=1}^k \alpha_i f(\mathbf{y}_i) \leq \max_i f(\mathbf{y}_i)$
- С другой стороны $\{\mathbf{y}_1, \ldots, \mathbf{y}_{2^n}\} \subset B_\infty(r)$, а значит $\max_{\mathbf{z} \in B_\infty(r)} f(\mathbf{z}) \geq \max_i f(\mathbf{y}_i)$
- Значит $\max_{\mathbf{z} \in B_\infty(r)} f(\mathbf{z}) = \max_i f(\mathbf{y}_i) = L$

- Покажем, что $|f(\mathbf{x})| \leq M\|\mathbf{x}\|_2$ для любого $\mathbf{x}$ из окрестности 0.
- Рассмотрим множество $B_\infty(r) = \{\mathbf{x} \mid \|\mathbf{x}\|_\infty \leq r\}$ для такого $r > 0$, что $B_\infty(r) \subset \mathcal{X}$
- $B_\infty(r)$ есть выпуклая оболочка 2^n точек (каких?) $\mathbf{y}_1, \dots, \mathbf{y}_{2^n}$
- Пусть $\mathbf{z} = \sum_{i=1}^k \alpha_i \mathbf{y}_i$ произвольная точка из $B_\infty(r)$
- Тогда $f(\mathbf{z}) \leq \sum_{i=1}^k \alpha_i f(\mathbf{y}_i) \leq \max_i f(\mathbf{y}_i)$
- С другой стороны $\{\mathbf{y}_1, \dots, \mathbf{y}_{2^n}\} \subset B_\infty(r)$, а значит $\max_{\mathbf{z} \in B_\infty(r)} f(\mathbf{z}) \geq \max_i f(\mathbf{y}_i)$
- Значит $\max_{\mathbf{z} \in B_\infty(r)} f(\mathbf{z}) = \max_i f(\mathbf{y}_i) = L$
- Следовательно, $f(\mathbf{x}) \leq L$ для всех точек $\mathbf{x} \in B_\infty(r)$, а значит и для всех $\mathbf{x} \in B_2(r) = \{\mathbf{x} \mid \|\mathbf{x}\|_2 \leq r\} \subset B_\infty(r)$

- Рассмотрим $\mathbf{x} \neq 0 \in B_2(r)$ и $\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_2} r \in \partial B_2(r)$

- Рассмотрим $\mathbf{x} \neq 0 \in B_2(r)$ и $\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_2} r \in \partial B_2(r)$
- Тогда $\mathbf{x} \in [0, \mathbf{y}]$ и выполнено $f(\mathbf{x}) \leq \lambda f(\mathbf{y}) + (1 - \lambda) \underbrace{f(0)}_{=0} \leq \lambda L = \frac{L}{r} \|\mathbf{x}\|_2 = M \|\mathbf{x}\|_2$

- Рассмотрим $\mathbf{x} \neq 0 \in B_2(r)$ и $\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_2} r \in \partial B_2(r)$
- Тогда $\mathbf{x} \in [0, \mathbf{y}]$ и выполнено $f(\mathbf{x}) \leq \lambda f(\mathbf{y}) + (1 - \lambda) \underbrace{f(0)}_{=0} \leq \lambda L = \frac{L}{r} \|\mathbf{x}\|_2 = M \|\mathbf{x}\|_2$
- Теперь покажем, что $f(\mathbf{x}) \geq -M \|\mathbf{x}\|_2$

- Рассмотрим $\mathbf{x} \neq 0 \in B_2(r)$ и $\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_2} r \in \partial B_2(r)$
- Тогда $\mathbf{x} \in [0, \mathbf{y}]$ и выполнено
  $f(\mathbf{x}) \leq \lambda f(\mathbf{y}) + (1 - \lambda) \underbrace{f(0)}_{=0} \leq \lambda L = \frac{L}{r} \|\mathbf{x}\|_2 = M \|\mathbf{x}\|_2$
- Теперь покажем, что $f(\mathbf{x}) \geq -M \|\mathbf{x}\|_2$
- Для этого заметим, что если $\mathbf{x} \in B_2(r)$, то и $-\mathbf{x} \in B_2(r)$

- Рассмотрим $\mathbf{x} \neq 0 \in B_2(r)$ и $\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_2} r \in \partial B_2(r)$
- Тогда $\mathbf{x} \in [0, \mathbf{y}]$ и выполнено $f(\mathbf{x}) \leq \lambda f(\mathbf{y}) + (1-\lambda) \underbrace{f(0)}_{=0} \leq \lambda L = \frac{L}{r}\|\mathbf{x}\|_2 = M\|\mathbf{x}\|_2$
- Теперь покажем, что $f(\mathbf{x}) \geq -M\|\mathbf{x}\|_2$
- Для этого заметим, что если $\mathbf{x} \in B_2(r)$, то и $-\mathbf{x} \in B_2(r)$
- Тогда $0 = f(0) = f\left(\frac{1}{2}\mathbf{x} - \frac{1}{2}\mathbf{x}\right) \leq \frac{1}{2}f(\mathbf{x}) + \frac{1}{2}f(-\mathbf{x})$

- Рассмотрим $\mathbf{x} \neq 0 \in B_2(r)$ и $\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_2} r \in \partial B_2(r)$
- Тогда $\mathbf{x} \in [0, \mathbf{y}]$ и выполнено
  $f(\mathbf{x}) \leq \lambda f(\mathbf{y}) + (1 - \lambda) \underbrace{f(0)}_{=0} \leq \lambda L = \frac{L}{r} \|\mathbf{x}\|_2 = M \|\mathbf{x}\|_2$
- Теперь покажем, что $f(\mathbf{x}) \geq -M\|\mathbf{x}\|_2$
- Для этого заметим, что если $\mathbf{x} \in B_2(r)$, то и $-\mathbf{x} \in B_2(r)$
- Тогда $0 = f(0) = f\left(\frac{1}{2}\mathbf{x} - \frac{1}{2}\mathbf{x}\right) \leq \frac{1}{2} f(\mathbf{x}) + \frac{1}{2} f(-\mathbf{x})$
- И $f(\mathbf{x}) \geq -f(-\mathbf{x}) \geq -M\|\mathbf{x}\|_2$

- Рассмотрим $\mathbf{x} \neq 0 \in B_2(r)$ и $\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_2} r \in \partial B_2(r)$
- Тогда $\mathbf{x} \in [0, \mathbf{y}]$ и выполнено
$f(\mathbf{x}) \leq \lambda f(\mathbf{y}) + (1 - \lambda) \underbrace{f(0)}_{=0} \leq \lambda L = \frac{L}{r} \|\mathbf{x}\|_2 = M \|\mathbf{x}\|_2$
- Теперь покажем, что $f(\mathbf{x}) \geq -M \|\mathbf{x}\|_2$
- Для этого заметим, что если $\mathbf{x} \in B_2(r)$, то и $-\mathbf{x} \in B_2(r)$
- Тогда $0 = f(0) = f\left(\frac{1}{2}\mathbf{x} - \frac{1}{2}\mathbf{x}\right) \leq \frac{1}{2} f(\mathbf{x}) + \frac{1}{2} f(-\mathbf{x})$
- И $f(\mathbf{x}) \geq -f(-\mathbf{x}) \geq -M \|\mathbf{x}\|_2$
- Таким образом, $|f(\mathbf{x})| \leq M \|\mathbf{x}\|_2$

Доказательство: общий случай

- Пусть n размерность $\operatorname{dom}(f)$

Доказательство: общий случай

- Пусть n размерность $\operatorname{dom}(f)$
- Тогда $\operatorname{aff}(\operatorname{dom}(f)) = \mathbf{a} + \mathcal{V}$, где $\mathcal{V}$ — линейное подпространство размерности n

Доказательство: общий случай

- Пусть n размерность $\mathrm{dom}\,(f)$
- Тогда $\mathrm{aff}\,(\mathrm{dom}\,(f)) = \mathbf{a} + \mathcal{V}$, где $\mathcal{V}$ — линейное подпространство размерности n
- $\mathcal{V}$ изоморфно $\mathbb{R}^n$, поэтому можно построить линейное биективное отображение $v : \mathbb{R}^n \to \mathcal{V}$ выбором некоторого базиса в $\mathcal{V}$

Доказательство: общий случай

- Пусть n размерность $\operatorname{dom}(f)$
- Тогда $\operatorname{aff}(\operatorname{dom}(f)) = \mathbf{a} + \mathcal{V}$, где $\mathcal{V}$ — линейное подпространство размерности n
- $\mathcal{V}$ изоморфно $\mathbb{R}^n$, поэтому можно построить линейное биективное отображение $v : \mathbb{R}^n \to \mathcal{V}$ выбором некоторого базиса в $\mathcal{V}$
- Рассмотрим отображение $h : \mathbb{R}^n \to \operatorname{aff}(\operatorname{dom}(f))$ вида $h(\mathbf{y}) = \mathbf{a} + v(\mathbf{y})$

Доказательство: общий случай

- Пусть n размерность $\mathrm{dom}\,(f)$
- Тогда $\mathrm{aff}\,(\mathrm{dom}\,(f)) = \mathbf{a} + \mathcal{V}$, где $\mathcal{V}$ — линейное подпространство размерности n
- $\mathcal{V}$ изоморфно $\mathbb{R}^n$, поэтому можно построить линейное биективное отображение $v : \mathbb{R}^n \to \mathcal{V}$ выбором некоторого базиса в $\mathcal{V}$
- Рассмотрим отображение $h : \mathbb{R}^n \to \mathrm{aff}\,(\mathrm{dom}\,(f))$ вида $h(\mathbf{y}) = \mathbf{a} + v(\mathbf{y})$
- Применение отображения h^{-1} к множеству $\mathrm{relint}\,(\mathrm{dom}\,(f))$ даст некоторое открытое подмножество $\mathcal{Y}$ в $\mathbb{R}^n$

Доказательство: общий случай

- Пусть n размерность $\mathrm{dom}\,(f)$
- Тогда $\mathrm{aff}\,(\mathrm{dom}\,(f)) = \mathbf{a} + \mathcal{V}$, где $\mathcal{V}$ — линейное подпространство размерности n
- $\mathcal{V}$ изоморфно $\mathbb{R}^n$, поэтому можно построить линейное биективное отображение $v : \mathbb{R}^n \to \mathcal{V}$ выбором некоторого базиса в $\mathcal{V}$
- Рассмотрим отображение $h : \mathbb{R}^n \to \mathrm{aff}\,(\mathrm{dom}\,(f))$ вида $h(\mathbf{y}) = \mathbf{a} + v(\mathbf{y})$
- Применение отображения h^{-1} к множеству $\mathrm{relint}\,(\mathrm{dom}\,(f))$ даст некоторое открытое подмножество $\mathcal{Y}$ в $\mathbb{R}^n$
- Более того, $h(0) = \mathbf{a} \in \mathrm{relint}\,(\mathrm{dom}\,(f))$, а значит $h^{-1}(\mathbf{a}) = 0 \ni \mathcal{Y}$

Доказательство: общий случай

- Пусть n размерность $\mathrm{dom}\,(f)$
- Тогда $\mathrm{aff}\,(\mathrm{dom}\,(f)) = \mathbf{a} + \mathcal{V}$, где $\mathcal{V}$ — линейное подпространство размерности n
- $\mathcal{V}$ изоморфно $\mathbb{R}^n$, поэтому можно построить линейное биективное отображение $v : \mathbb{R}^n \to \mathcal{V}$ выбором некоторого базиса в $\mathcal{V}$
- Рассмотрим отображение $h : \mathbb{R}^n \to \mathrm{aff}\,(\mathrm{dom}\,(f))$ вида $h(\mathbf{y}) = \mathbf{a} + v(\mathbf{y})$
- Применение отображения h^{-1} к множеству $\mathrm{relint}\,(\mathrm{dom}\,(f))$ даст некоторое открытое подмножество $\mathcal{Y}$ в $\mathbb{R}^n$
- Более того, $h(0) = \mathbf{a} \in \mathrm{relint}\,(\mathrm{dom}\,(f))$, а значит $h^{-1}(\mathbf{a}) = 0 \ni \mathcal{Y}$
- Определим функцию $g : \mathcal{Y} \to \mathbb{R}$ в виде $g(\mathbf{y}) = f(\mathbf{a} + v(\mathbf{y})) - f(\mathbf{a})$

Доказательство: общий случай

- Пусть n размерность $\operatorname{dom}(f)$
- Тогда $\operatorname{aff}(\operatorname{dom}(f)) = \mathbf{a} + \mathcal{V}$, где $\mathcal{V}$ — линейное подпространство размерности n
- $\mathcal{V}$ изоморфно $\mathbb{R}^n$, поэтому можно построить линейное биективное отображение $v : \mathbb{R}^n \to \mathcal{V}$ выбором некоторого базиса в $\mathcal{V}$
- Рассмотрим отображение $h : \mathbb{R}^n \to \operatorname{aff}(\operatorname{dom}(f))$ вида $h(\mathbf{y}) = \mathbf{a} + v(\mathbf{y})$
- Применение отображения h^{-1} к множеству $\operatorname{relint}(\operatorname{dom}(f))$ даст некоторое открытое подмножество $\mathcal{Y}$ в $\mathbb{R}^n$
- Более того, $h(0) = \mathbf{a} \in \operatorname{relint}(\operatorname{dom}(f))$, а значит $h^{-1}(\mathbf{a}) = 0 \ni \mathcal{Y}$
- Определим функцию $g : \mathcal{Y} \to \mathbb{R}$ в виде $g(\mathbf{y}) = f(\mathbf{a} + v(\mathbf{y})) - f(\mathbf{a})$
- Заметим, что $g(0) = 0$ и g выпукла

- Обозначим $\mathbf{x} = \mathbf{a} + v(\mathbf{y}) \in \operatorname{relint}(\operatorname{dom}(f))$, тогда $g(\mathbf{y}) = f(\mathbf{x}) - f(\mathbf{a})$

- Обозначим $\mathbf{x} = \mathbf{a} + v(\mathbf{y}) \in \operatorname{relint}(\operatorname{dom}(f))$, тогда $g(\mathbf{y}) = f(\mathbf{x}) - f(\mathbf{a})$
- Теперь для завершения доказательства нужно показать, что для некоторого $M > 0$

$$|g(\mathbf{y})| \le M \|v(\mathbf{y})\|_2$$

  выполнено для любого $\mathbf{y} \in \mathcal{O}(0)$

- Обозначим $\mathbf{x} = \mathbf{a} + v(\mathbf{y}) \in \mathrm{relint}\,(\mathrm{dom}\,(f))$, тогда $g(\mathbf{y}) = f(\mathbf{x}) - f(\mathbf{a})$
- Теперь для завершения доказательства нужно показать, что для некоторого $M > 0$

$$|g(\mathbf{y})| \le M \|v(\mathbf{y})\|_2$$

  выполнено для любого $\mathbf{y} \in \mathcal{O}(0)$
- Заметим, что отображение $\|v(\mathbf{y})\|$ есть некоторая норма в силу линейности v

- Обозначим $\mathbf{x} = \mathbf{a} + v(\mathbf{y}) \in \operatorname{relint}(\operatorname{dom}(f))$, тогда $g(\mathbf{y}) = f(\mathbf{x}) - f(\mathbf{a})$
- Теперь для завершения доказательства нужно показать, что для некоторого $M > 0$

$$|g(\mathbf{y})| \leq M \|v(\mathbf{y})\|_2$$

  выполнено для любого $\mathbf{y} \in \mathcal{O}(0)$
- Заметим, что отображение $\|v(\mathbf{y})\|$ есть некоторая норма в силу линейности v
- Значит в силу эквивалентности норм в $\mathbb{R}^n$ достаточно показать, что найдётся $M > 0$, что

$$|g(\mathbf{y})| \leq M \|\mathbf{y}\|_2$$

  для любого $\mathbf{y} \in \mathcal{O}(0)$.

- Обозначим $\mathbf{x} = \mathbf{a} + v(\mathbf{y}) \in \mathrm{relint}\,(\mathrm{dom}\,(f))$, тогда $g(\mathbf{y}) = f(\mathbf{x}) - f(\mathbf{a})$
- Теперь для завершения доказательства нужно показать, что для некоторого $M > 0$

$$|g(\mathbf{y})| \leq M \|v(\mathbf{y})\|_2$$

  выполнено для любого $\mathbf{y} \in \mathcal{O}(0)$
- Заметим, что отображение $\|v(\mathbf{y})\|$ есть некоторая норма в силу линейности v
- Значит в силу эквивалентности норм в $\mathbb{R}^n$ достаточно показать, что найдётся $M > 0$, что

$$|g(\mathbf{y})| \leq M \|\mathbf{y}\|_2$$

  для любого $\mathbf{y} \in \mathcal{O}(0)$.
- Но мы уже показали это в самом начале доказательства!

А что же в граничных точках?

- В граничных точках выпуклая функция может иметь разрыв

$$f(x) = \begin{cases} x^2, & -1 < x < 1 \\ 2, & x = 1, -1 \end{cases}$$

А что же в граничных точках?

- В граничных точках выпуклая функция может иметь разрыв
$$f(x) = \begin{cases} x^2, & -1 < x < 1 \\ 2, & x = 1, -1 \end{cases}$$
- Функция f называется полунепрерывной снизу в точке $\mathbf{x}_0$, если для любого $y < f(\mathbf{x}_0)$ найдётся такая окрестность $\mathcal{O}(\mathbf{x}_0)$, что $y < f(\mathbf{x})$ для любого $\mathbf{x} \in \mathcal{O}(\mathbf{x}_0)$

А что же в граничных точках?

- В граничных точках выпуклая функция может иметь разрыв

$$f(x) = \begin{cases} x^2, & -1 < x < 1 \\ 2, & x = 1, -1 \end{cases}$$

- Функция f называется полунепрерывной снизу в точке $\mathbf{x}_0$, если для любого $y < f(\mathbf{x}_0)$ найдётся такая окрестность $\mathcal{O}(\mathbf{x}_0)$, что $y < f(\mathbf{x})$ для любого $\mathbf{x} \in \mathcal{O}(\mathbf{x}_0)$

Q: является ли функция выше полунепрерывной снизу?

А что же в граничных точках?

- В граничных точках выпуклая функция может иметь разрыв
$$f(x) = \begin{cases} x^2, & -1 < x < 1 \\ 2, & x = 1, -1 \end{cases}$$
- Функция f называется полунепрерывной снизу в точке $\mathbf{x}_0$, если для любого $y < f(\mathbf{x}_0)$ найдётся такая окрестность $\mathcal{O}(\mathbf{x}_0)$, что $y < f(\mathbf{x})$ для любого $\mathbf{x} \in \mathcal{O}(\mathbf{x}_0)$

Q: является ли функция выше полунепрерывной снизу?

Определение

Выпуклая функция называется замкнутой, если её надграфик замкнутое множество.

А что же в граничных точках?

- В граничных точках выпуклая функция может иметь разрыв
$$f(x) = \begin{cases} x^2, & -1 < x < 1 \\ 2, & x = 1, -1 \end{cases}$$
- Функция f называется полунепрерывной снизу в точке $\mathbf{x}_0$, если для любого $y < f(\mathbf{x}_0)$ найдётся такая окрестность $\mathcal{O}(\mathbf{x}_0)$, что $y < f(\mathbf{x})$ для любого $\mathbf{x} \in \mathcal{O}(\mathbf{x}_0)$

Q: является ли функция выше полунепрерывной снизу?

Определение

Выпуклая функция называется замкнутой, если её надграфик замкнутое множество.

Теорема

Выпуклая функция замкнута iff она полунепрерывна снизу

А что же в граничных точках?

- В граничных точках выпуклая функция может иметь разрыв
$$f(x) = \begin{cases} x^2, & -1 < x < 1 \\ 2, & x = 1, -1 \end{cases}$$
- Функция f называется полунепрерывной снизу в точке $\mathbf{x}_0$, если для любого $y < f(\mathbf{x}_0)$ найдётся такая окрестность $\mathcal{O}(\mathbf{x}_0)$, что $y < f(\mathbf{x})$ для любого $\mathbf{x} \in \mathcal{O}(\mathbf{x}_0)$

Q: является ли функция выше полунепрерывной снизу?

Определение

Выпуклая функция называется замкнутой, если её надграфик замкнутое множество.

Теорема

Выпуклая функция замкнута iff она полунепрерывна снизу

Замечание

Замкнутая выпуклая функция может иметь незамкнутую область определения, например $f(x) = 1/x, x > 0$

Ещё один контрпример

Разрывная замкнутая выпуклая функция

Ещё один контрпример

Разрывная замкнутая выпуклая функция

- Пусть $f(x,y) = \begin{cases} \frac{x^2}{y}, & (x,y) \in \mathcal{X} \\ 0, & (x,y) = (0,0) \end{cases}$

Ещё один контрпример

Разрывная замкнутая выпуклая функция

- Пусть $f(x,y) = \begin{cases} \frac{x^2}{y}, & (x,y) \in \mathcal{X} \\ 0, & (x,y) = (0,0) \end{cases}$
- $\mathcal{X} = \{(x,y) \mid y \geq x^2\}$

Ещё один контрпример

Разрывная замкнутая выпуклая функция

- Пусть $f(x, y) = \begin{cases} \frac{x^2}{y}, & (x, y) \in \mathcal{X} \\ 0, & (x, y) = (0, 0) \end{cases}$
- $\mathcal{X} = \{(x, y) \mid y \geq x^2\}$

Свойства

Ещё один контрпример

Разрывная замкнутая выпуклая функция

- Пусть $f(x, y) = \begin{cases} \frac{x^2}{y}, & (x, y) \in \mathcal{X} \\ 0, & (x, y) = (0, 0) \end{cases}$
- $\mathcal{X} = \{(x, y) \mid y \geq x^2\}$

Свойства

- Разрывна в $(0, 0)$ так как $(x_n, x_n^2) \in \mathcal{X}$ и $(x_n, 2x_n^2) \in \mathcal{X}$ такие что $x_n \to 0$ дают два различных значения f_n

Ещё один контрпример

Разрывная замкнутая выпуклая функция

- Пусть $f(x,y) = \begin{cases} \frac{x^2}{y}, & (x,y) \in \mathcal{X} \\ 0, & (x,y) = (0,0) \end{cases}$
- $\mathcal{X} = \{(x,y) \mid y \geq x^2\}$

Свойства

- Разрывна в $(0,0)$ так как $(x_n, x_n^2) \in \mathcal{X}$ и $(x_n, 2x_n^2) \in \mathcal{X}$ такие что $x_n \to 0$ дают два различных значения f_n
- Выпукла при $y > 0$ по критерию второго порядка (проверьте!)

Ещё один контрпример

Разрывная замкнутая выпуклая функция

- Пусть $f(x, y) = \begin{cases} \frac{x^2}{y}, & (x, y) \in \mathcal{X} \\ 0, & (x, y) = (0, 0) \end{cases}$
- $\mathcal{X} = \{(x, y) \mid y \geq x^2\}$

Свойства

- Разрывна в $(0, 0)$ так как $(x_n, x_n^2) \in \mathcal{X}$ и $(x_n, 2x_n^2) \in \mathcal{X}$ такие что $x_n \to 0$ дают два различных значения f_n
- Выпукла при $y > 0$ по критерию второго порядка (проверьте!)
- Выпукла при $y = 0$ по определению (проверьте!)

Ещё один контрпример

Разрывная замкнутая выпуклая функция

- Пусть $f(x, y) = \begin{cases} \frac{x^2}{y}, & (x, y) \in \mathcal{X} \\ 0, & (x, y) = (0, 0) \end{cases}$
- $\mathcal{X} = \{(x, y) \mid y \geq x^2\}$

Свойства

- Разрывна в $(0, 0)$ так как $(x_n, x_n^2) \in \mathcal{X}$ и $(x_n, 2x_n^2) \in \mathcal{X}$ такие что $x_n \to 0$ дают два различных значения f_n
- Выпукла при $y > 0$ по критерию второго порядка (проверьте!)
- Выпукла при $y = 0$ по определению (проверьте!)
- Непрерывна на $\mathcal{X}$, а значит и полунепрерывна снизу на $\mathcal{X}$

Ещё один контрпример

Разрывная замкнутая выпуклая функция

- Пусть $f(x, y) = \begin{cases} \frac{x^2}{y}, & (x, y) \in \mathcal{X} \\ 0, & (x, y) = (0, 0) \end{cases}$
- $\mathcal{X} = \{(x, y) \mid y \geq x^2\}$

Свойства

- Разрывна в $(0, 0)$ так как $(x_n, x_n^2) \in \mathcal{X}$ и $(x_n, 2x_n^2) \in \mathcal{X}$ такие что $x_n \to 0$ дают два различных значения f_n
- Выпукла при $y > 0$ по критерию второго порядка (проверьте!)
- Выпукла при $y = 0$ по определению (проверьте!)
- Непрерывна на $\mathcal{X}$, а значит и полунепрерывна снизу на $\mathcal{X}$
- Рассмотрим точку $(0, 0)$. Проверим, что для любого $u < f(x_0, y_0) = 0$ найдётся окрестность $\mathcal{O}((0, 0)) \subset \mathcal{X}$, что $u < f(x, y) = \frac{x^2}{y}$, $(x, y) \in \mathcal{O}((0, 0))$. Так как $\frac{x^2}{y} > 0$, то неравенство выполнено

Производная по направлению

Определение

Производная функции f по направлению $\mathbf{d}$ в точке $\mathbf{x}$ определяется как

$$f'(\mathbf{x}, \mathbf{d}) = \lim_{\alpha \to +0} \frac{f(\mathbf{x} + \alpha \mathbf{d}) - f(\mathbf{x})}{\alpha}$$

Производная по направлению

Определение

Производная функции f по направлению $\mathbf{d}$ в точке $\mathbf{x}$ определяется как

$$f'(\mathbf{x}, \mathbf{d}) = \lim_{\alpha \to +0} \frac{f(\mathbf{x} + \alpha \mathbf{d}) - f(\mathbf{x})}{\alpha}$$

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$ произвольная точка. Тогда для любого направления $\mathbf{d}$ существует производная по направлению $f'(\mathbf{x}, \mathbf{d})$.

Производная по направлению

Определение

Производная функции f по направлению $\mathbf{d}$ в точке $\mathbf{x}$ определяется как

$$f'(\mathbf{x}, \mathbf{d}) = \lim_{\alpha \to +0} \frac{f(\mathbf{x} + \alpha \mathbf{d}) - f(\mathbf{x})}{\alpha}$$

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$ произвольная точка. Тогда для любого направления $\mathbf{d}$ существует производная по направлению $f'(\mathbf{x}, \mathbf{d})$.

Доказательство

Производная по направлению

Определение

Производная функции f по направлению $\mathbf{d}$ в точке $\mathbf{x}$ определяется как

$$f'(\mathbf{x}, \mathbf{d}) = \lim_{\alpha \to +0} \frac{f(\mathbf{x} + \alpha \mathbf{d}) - f(\mathbf{x})}{\alpha}$$

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ произвольная точка. Тогда для любого направления $\mathbf{d}$ существует производная по направлению $f'(\mathbf{x}, \mathbf{d})$.

Доказательство

- Рассмотрим функцию $\varphi(\alpha) = \frac{f(\mathbf{x}+\alpha\mathbf{d})-f(\mathbf{x})}{\alpha}$

Производная по направлению

Определение

Производная функции f по направлению $\mathbf{d}$ в точке $\mathbf{x}$ определяется как

$$f'(\mathbf{x}, \mathbf{d}) = \lim_{\alpha \to +0} \frac{f(\mathbf{x} + \alpha \mathbf{d}) - f(\mathbf{x})}{\alpha}$$

Теорема

Пусть f выпуклая функция и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ произвольная точка. Тогда для любого направления $\mathbf{d}$ существует производная по направлению $f'(\mathbf{x}, \mathbf{d})$.

Доказательство

- Рассмотрим функцию $\varphi(\alpha) = \frac{f(\mathbf{x}+\alpha\mathbf{d}) - f(\mathbf{x})}{\alpha}$
- Пусть $0 < \alpha_1 \le \alpha_2$, обозначим $\lambda = \frac{\alpha_1}{\alpha_2}$ и в силу выпуклости f выполнено $f(\mathbf{x} + \alpha_1 \mathbf{d}) = f((1-\lambda)\mathbf{x} + \lambda(\mathbf{x} + \alpha_2 \mathbf{d})) \le (1-\lambda) f(\mathbf{x}) + \lambda f(\mathbf{x} + \alpha_2 \mathbf{d})$

- Тогда $\varphi(\alpha_1) = \frac{f(\mathbf{x}+\alpha_1\mathbf{d})-f(\mathbf{x})}{\alpha_1} \leq \frac{(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_1} = \frac{\lambda(f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x}))}{\alpha_1} = \frac{f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_2} = \varphi(\alpha_2)$

- Тогда $\varphi(\alpha_1) = \frac{f(\mathbf{x}+\alpha_1\mathbf{d})-f(\mathbf{x})}{\alpha_1} \leq \frac{(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_1} = \frac{\lambda(f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x}))}{\alpha_1} = \frac{f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_2} = \varphi(\alpha_2)$
- Таким образом, $\varphi(\alpha)$ неубывающая функция

- Тогда $\varphi(\alpha_1) = \frac{f(\mathbf{x}+\alpha_1\mathbf{d})-f(\mathbf{x})}{\alpha_1} \leq \frac{(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_1} = \frac{\lambda(f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x}))}{\alpha_1} = \frac{f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_2} = \varphi(\alpha_2)$
- Таким образом, $\varphi(\alpha)$ неубывающая функция
- Далее рассмотрим $\alpha_+ > 0$ и $\alpha_- < 0$ и обозначим $\lambda_1 = \frac{\alpha_+}{\alpha_+-\alpha_-}$, $\lambda_2 = \frac{-\alpha_-}{\alpha_+-\alpha_-}$ так что $\lambda_{1,2} \geq 0$ и $\lambda_1 + \lambda_2 = 1$

- Тогда $\varphi(\alpha_1) = \frac{f(\mathbf{x}+\alpha_1\mathbf{d})-f(\mathbf{x})}{\alpha_1} \leq \frac{(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_1} = \frac{\lambda(f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x}))}{\alpha_1} = \frac{f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_2} = \varphi(\alpha_2)$
- Таким образом, $\varphi(\alpha)$ неубывающая функция
- Далее рассмотрим $\alpha_+ > 0$ и $\alpha_- < 0$ и обозначим $\lambda_1 = \frac{\alpha_+}{\alpha_+-\alpha_-}$, $\lambda_2 = \frac{-\alpha_-}{\alpha_+-\alpha_-}$ так что $\lambda_{1,2} \geq 0$ и $\lambda_1 + \lambda_2 = 1$
- Поскольку $\mathbf{x} = \lambda_1(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2(\mathbf{x} + \alpha_+\mathbf{d})$ и функция f выпукла, то $f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$

- Тогда $\varphi(\alpha_1) = \frac{f(\mathbf{x}+\alpha_1\mathbf{d})-f(\mathbf{x})}{\alpha_1} \leq \frac{(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_1} = \frac{\lambda(f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x}))}{\alpha_1} = \frac{f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_2} = \varphi(\alpha_2)$
- Таким образом, $\varphi(\alpha)$ неубывающая функция
- Далее рассмотрим $\alpha_+ > 0$ и $\alpha_- < 0$ и обозначим $\lambda_1 = \frac{\alpha_+}{\alpha_+-\alpha_-}$, $\lambda_2 = \frac{-\alpha_-}{\alpha_+-\alpha_-}$ так что $\lambda_{1,2} \geq 0$ и $\lambda_1 + \lambda_2 = 1$
- Поскольку $\mathbf{x} = \lambda_1(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2(\mathbf{x} + \alpha_+\mathbf{d})$ и функция f выпукла, то $f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$
- Иначе $(\lambda_1 + \lambda_2)f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$

- Тогда $\varphi(\alpha_1) = \frac{f(\mathbf{x}+\alpha_1\mathbf{d})-f(\mathbf{x})}{\alpha_1} \leq \frac{(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_1} = \frac{\lambda(f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x}))}{\alpha_1} = \frac{f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_2} = \varphi(\alpha_2)$
- Таким образом, $\varphi(\alpha)$ неубывающая функция
- Далее рассмотрим $\alpha_+ > 0$ и $\alpha_- < 0$ и обозначим $\lambda_1 = \frac{\alpha_+}{\alpha_+-\alpha_-}$, $\lambda_2 = \frac{-\alpha_-}{\alpha_+-\alpha_-}$ так что $\lambda_{1,2} \geq 0$ и $\lambda_1 + \lambda_2 = 1$
- Поскольку $\mathbf{x} = \lambda_1(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2(\mathbf{x} + \alpha_+\mathbf{d})$ и функция f выпукла, то $f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$
- Иначе $(\lambda_1 + \lambda_2)f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$
- Или $\lambda_1(f(\mathbf{x} + \alpha_-\mathbf{d}) - f(\mathbf{x})) \geq -\lambda_2(f(\mathbf{x} + \alpha_+\mathbf{d}) - f(\mathbf{x}))$

- Тогда $\varphi(\alpha_1) = \frac{f(\mathbf{x}+\alpha_1\mathbf{d})-f(\mathbf{x})}{\alpha_1} \leq \frac{(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_1} = \frac{\lambda(f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x}))}{\alpha_1} = \frac{f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_2} = \varphi(\alpha_2)$
- Таким образом, $\varphi(\alpha)$ неубывающая функция
- Далее рассмотрим $\alpha_+ > 0$ и $\alpha_- < 0$ и обозначим $\lambda_1 = \frac{\alpha_+}{\alpha_+ - \alpha_-}$, $\lambda_2 = \frac{-\alpha_-}{\alpha_+ - \alpha_-}$ так что $\lambda_{1,2} \geq 0$ и $\lambda_1 + \lambda_2 = 1$
- Поскольку $\mathbf{x} = \lambda_1(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2(\mathbf{x} + \alpha_+\mathbf{d})$ и функция f выпукла, то $f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$
- Иначе $(\lambda_1 + \lambda_2)f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$
- Или $\lambda_1(f(\mathbf{x} + \alpha_-\mathbf{d}) - f(\mathbf{x})) \geq -\lambda_2(f(\mathbf{x} + \alpha_+\mathbf{d}) - f(\mathbf{x}))$
- Заменяя $\lambda_{1,2}$ на $\alpha_{+,-}$ получим $\alpha_+(f(\mathbf{x} + \alpha_-\mathbf{d}) - f(\mathbf{x})) \geq \alpha_-(f(\mathbf{x} + \alpha_+\mathbf{d}) - f(\mathbf{x}))$

- Тогда $\varphi(\alpha_1) = \frac{f(\mathbf{x}+\alpha_1\mathbf{d})-f(\mathbf{x})}{\alpha_1} \leq \frac{(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_1} = \frac{\lambda(f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x}))}{\alpha_1} = \frac{f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_2} = \varphi(\alpha_2)$
- Таким образом, $\varphi(\alpha)$ неубывающая функция
- Далее рассмотрим $\alpha_+ > 0$ и $\alpha_- < 0$ и обозначим $\lambda_1 = \frac{\alpha_+}{\alpha_+ - \alpha_-}$, $\lambda_2 = \frac{-\alpha_-}{\alpha_+ - \alpha_-}$ так что $\lambda_{1,2} \geq 0$ и $\lambda_1 + \lambda_2 = 1$
- Поскольку $\mathbf{x} = \lambda_1(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2(\mathbf{x} + \alpha_+\mathbf{d})$ и функция f выпукла, то $f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$
- Иначе $(\lambda_1 + \lambda_2)f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$
- Или $\lambda_1(f(\mathbf{x} + \alpha_-\mathbf{d}) - f(\mathbf{x})) \geq -\lambda_2(f(\mathbf{x} + \alpha_+\mathbf{d}) - f(\mathbf{x}))$
- Заменяя $\lambda_{1,2}$ на $\alpha_{+,-}$ получим $\alpha_+(f(\mathbf{x} + \alpha_-\mathbf{d}) - f(\mathbf{x})) \geq \alpha_-(f(\mathbf{x} + \alpha_+\mathbf{d}) - f(\mathbf{x}))$
- Откуда следует $\varphi(\alpha_-) \leq \varphi(\alpha_+)$

- Тогда $\varphi(\alpha_1) = \frac{f(\mathbf{x}+\alpha_1\mathbf{d})-f(\mathbf{x})}{\alpha_1} \leq \frac{(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_1} = \frac{\lambda(f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x}))}{\alpha_1} = \frac{f(\mathbf{x}+\alpha_2\mathbf{d})-f(\mathbf{x})}{\alpha_2} = \varphi(\alpha_2)$
- Таким образом, $\varphi(\alpha)$ неубывающая функция
- Далее рассмотрим $\alpha_+ > 0$ и $\alpha_- < 0$ и обозначим $\lambda_1 = \frac{\alpha_+}{\alpha_+-\alpha_-}$, $\lambda_2 = \frac{-\alpha_-}{\alpha_+-\alpha_-}$ так что $\lambda_{1,2} \geq 0$ и $\lambda_1 + \lambda_2 = 1$
- Поскольку $\mathbf{x} = \lambda_1(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2(\mathbf{x} + \alpha_+\mathbf{d})$ и функция f выпукла, то $f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$
- Иначе $(\lambda_1 + \lambda_2)f(\mathbf{x}) \leq \lambda_1 f(\mathbf{x} + \alpha_-\mathbf{d}) + \lambda_2 f(\mathbf{x} + \alpha_+\mathbf{d})$
- Или $\lambda_1(f(\mathbf{x} + \alpha_-\mathbf{d}) - f(\mathbf{x})) \geq -\lambda_2(f(\mathbf{x} + \alpha_+\mathbf{d}) - f(\mathbf{x}))$
- Заменяя $\lambda_{1,2}$ на $\alpha_{+,-}$ получим $\alpha_+(f(\mathbf{x} + \alpha_-\mathbf{d}) - f(\mathbf{x})) \geq \alpha_-(f(\mathbf{x} + \alpha_+\mathbf{d}) - f(\mathbf{x}))$
- Откуда следует $\varphi(\alpha_-) \leq \varphi(\alpha_+)$
- В итоге, $\varphi(\alpha)$ неубывающая функция для $\alpha > 0$ и ограничена снизу, таким образом существует предел $\lim\limits_{\alpha\to+0} \varphi(\alpha) = f'(\mathbf{x}, \mathbf{d})$

Свойства производной по направлению

Выпуклость и однородность

Пусть функция f выпукла и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$. Тогда

- $f'(\mathbf{x}, \mathbf{d})$ выпукла по $\mathbf{d}$
- для любого $\lambda \geq 0$ выполнено $f'(\mathbf{x}, \lambda\mathbf{d}) = \lambda f'(\mathbf{x}, \mathbf{d})$

Свойства производной по направлению

Выпуклость и однородность

Пусть функция f выпукла и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$. Тогда

- $f'(\mathbf{x}, \mathbf{d})$ выпукла по $\mathbf{d}$
- для любого $\lambda \geq 0$ выполнено $f'(\mathbf{x}, \lambda\mathbf{d}) = \lambda f'(\mathbf{x}, \mathbf{d})$

Доказательство

Свойства производной по направлению

Выпуклость и однородность

Пусть функция f выпукла и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Тогда

- $f'(\mathbf{x}, \mathbf{d})$ выпукла по $\mathbf{d}$
- для любого $\lambda \geq 0$ выполнено $f'(\mathbf{x}, \lambda\mathbf{d}) = \lambda f'(\mathbf{x}, \mathbf{d})$

Доказательство

- По определению $f'(\mathbf{x}, \lambda\mathbf{d}) = \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x}+\lambda\alpha\mathbf{d}) - f(\mathbf{x})}{\alpha} = \lambda \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x}+\lambda\alpha\mathbf{d}) - f(\mathbf{x})}{\alpha\lambda} = \lambda f'(\mathbf{x}, \mathbf{d})$

Свойства производной по направлению

Выпуклость и однородность

Пусть функция f выпукла и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$. Тогда

- $f'(\mathbf{x}, \mathbf{d})$ выпукла по $\mathbf{d}$
- для любого $\lambda \geq 0$ выполнено $f'(\mathbf{x}, \lambda\mathbf{d}) = \lambda f'(\mathbf{x}, \mathbf{d})$

Доказательство

- По определению $f'(\mathbf{x}, \lambda\mathbf{d}) = \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x}+\lambda\alpha\mathbf{d})-f(\mathbf{x})}{\alpha} =$
  $\lambda \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x}+\lambda\alpha\mathbf{d})-f(\mathbf{x})}{\alpha\lambda} = \lambda f'(\mathbf{x}, \mathbf{d})$
- Используем определение
  $f'(\mathbf{x}, \lambda\mathbf{d}_1 + (1-\lambda)\mathbf{d}_2) = \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x}+\alpha(\lambda\mathbf{d}_1+(1-\lambda)\mathbf{d}_2))-f(\mathbf{x})}{\alpha} =$
  $\lim\limits_{\alpha \to +0} \frac{f(\lambda(\mathbf{x}+\alpha\mathbf{d}_1)+(1-\lambda)(\mathbf{x}+\alpha\mathbf{d}_2))-f(\mathbf{x})}{\alpha} \leq$
  $\lim\limits_{\alpha \to +0} \frac{\lambda f(\mathbf{x}+\alpha\mathbf{d}_1)+(1-\lambda)f(\mathbf{x}+\alpha\mathbf{d}_2)-f(\mathbf{x})}{\alpha} =$
  $\lambda \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x}+\alpha\mathbf{d}_1)-f(\mathbf{x})}{\alpha} + (1-\lambda) \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x}+\alpha\mathbf{d}_2)-f(\mathbf{x})}{\alpha} =$
  $\lambda f'(\mathbf{x}, \mathbf{d}_1) + (1-\lambda) f'(\mathbf{x}, \mathbf{d}_2)$

Утверждение

Пусть f выпуклая функция и $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ произвольная точка. Тогда для любой точки $\mathbf{y} \in \mathrm{dom}\,(f)$ выполнено $f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x}, \mathbf{y} - \mathbf{x})$

Утверждение

Пусть f выпуклая функция и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$ произвольная точка. Тогда для любой точки $\mathbf{y} \in \operatorname{dom}(f)$ выполнено $f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x}, \mathbf{y} - \mathbf{x})$

Доказательство

Утверждение

Пусть f выпуклая функция и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$ произвольная точка. Тогда для любой точки $\mathbf{y} \in \operatorname{dom}(f)$ выполнено $f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x}, \mathbf{y} - \mathbf{x})$

Доказательство

- По определению $f'(\mathbf{x}, \mathbf{y} - \mathbf{x}) = \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x}+\alpha(\mathbf{y}-\mathbf{x}))-f(\mathbf{x})}{\alpha} = \lim\limits_{\alpha \to +0} \frac{f((1-\alpha)\mathbf{x}+\alpha\mathbf{y})-f(\mathbf{x})}{\alpha}$

Утверждение

Пусть f выпуклая функция и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$ произвольная точка. Тогда для любой точки $\mathbf{y} \in \operatorname{dom}(f)$ выполнено $f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x}, \mathbf{y} - \mathbf{x})$

Доказательство

- По определению $f'(\mathbf{x}, \mathbf{y} - \mathbf{x}) = \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x}+\alpha(\mathbf{y}-\mathbf{x}))-f(\mathbf{x})}{\alpha} = \lim\limits_{\alpha \to +0} \frac{f((1-\alpha)\mathbf{x}+\alpha\mathbf{y})-f(\mathbf{x})}{\alpha}$
- В силу выпуклости f: $\lim\limits_{\alpha \to +0} \frac{f((1-\alpha)\mathbf{x}+\alpha\mathbf{y})-f(\mathbf{x})}{\alpha} \leq \lim\limits_{\alpha \to +0} \frac{(1-\alpha)f(\mathbf{x})+\alpha f(\mathbf{y})-f(\mathbf{x})}{\alpha} = f(\mathbf{y}) - f(\mathbf{x})$

Утверждение

Пусть f выпуклая функция и $\mathbf{x} \in \operatorname{int}(\operatorname{dom}(f))$ произвольная точка. Тогда для любой точки $\mathbf{y} \in \operatorname{dom}(f)$ выполнено $f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x}, \mathbf{y} - \mathbf{x})$

Доказательство

- По определению $f'(\mathbf{x}, \mathbf{y} - \mathbf{x}) = \lim\limits_{\alpha \to +0} \frac{f(\mathbf{x} + \alpha(\mathbf{y} - \mathbf{x})) - f(\mathbf{x})}{\alpha} = \lim\limits_{\alpha \to +0} \frac{f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) - f(\mathbf{x})}{\alpha}$
- В силу выпуклости f: $\lim\limits_{\alpha \to +0} \frac{f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) - f(\mathbf{x})}{\alpha} \leq \lim\limits_{\alpha \to +0} \frac{(1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}) - f(\mathbf{x})}{\alpha} = f(\mathbf{y}) - f(\mathbf{x})$
- В итоге $f'(\mathbf{x}, \mathbf{y} - \mathbf{x}) \leq f(\mathbf{y}) - f(\mathbf{x})$

Градиент и производная по направлению

Определение

Функция f является дифференцируемой в точке $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ если найдётся вектор $\mathbf{g}$ такой что

$$\lim_{\mathbf{d}\to 0} \frac{f(\mathbf{x}+\mathbf{d}) - f(\mathbf{x}) - \langle \mathbf{g}, \mathbf{d} \rangle}{\|\mathbf{d}\|} = 0.$$

Этот вектор единственный (проверьте!) и называется градиентом.

Градиент и производная по направлению

Определение

Функция f является дифференцируемой в точке $\mathbf{x} \in \mathrm{int}\,(\mathrm{dom}\,(f))$ если найдётся вектор $\mathbf{g}$ такой что

$$\lim_{\mathbf{d}\to 0} \frac{f(\mathbf{x}+\mathbf{d}) - f(\mathbf{x}) - \langle \mathbf{g}, \mathbf{d} \rangle}{\|\mathbf{d}\|} = 0.$$

Этот вектор единственный (проверьте!) и называется градиентом.

Утверждение

Если f дифференцируема в точке $\mathbf{x}$, то для любого направления $\mathbf{d}$ выполнено $f'(\mathbf{x}, \mathbf{d}) = \langle f'(\mathbf{x}), \mathbf{d} \rangle$.

L-гладкая функция

Определение

Пусть $L > 0$. Функция f называется L-гладкой, если она дифференцируема и выполнено следующее неравенство

$$\|f'(\mathbf{x}) - f'(\mathbf{y})\|_* \leq L\|\mathbf{x} - \mathbf{y}\|,$$

для любых $\mathbf{x}, \mathbf{y}$. Норма $\|\cdot\|_*$ является сопряжённой нормой для $\|\cdot\|$

L-гладкая функция

Определение

Пусть $L > 0$. Функция f называется L-гладкой, если она дифференцируема и выполнено следующее неравенство

$$\|f'(\mathbf{x}) - f'(\mathbf{y})\|_* \leq L\|\mathbf{x} - \mathbf{y}\|,$$

для любых $\mathbf{x}, \mathbf{y}$. Норма $\|\cdot\|_*$ является сопряжённой нормой для $\|\cdot\|$

- L является константой Липшица для градиента

L-гладкая функция

Определение

Пусть $L > 0$. Функция f называется L-гладкой, если она дифференцируема и выполнено следующее неравенство

$$\|f'(\mathbf{x}) - f'(\mathbf{y})\|_* \leq L\|\mathbf{x} - \mathbf{y}\|,$$

для любых $\mathbf{x}, \mathbf{y}$. Норма $\|\cdot\|_*$ является сопряжённой нормой для $\|\cdot\|$

- L является константой Липшица для градиента
- Интерес представляет минимально возможное L, при котором выполняется неравенство в определении

L-гладкая функция

Определение

Пусть $L > 0$. Функция f называется L-гладкой, если она дифференцируема и выполнено следующее неравенство

$$\|f'(\mathbf{x}) - f'(\mathbf{y})\|_* \leq L\|\mathbf{x} - \mathbf{y}\|,$$

для любых $\mathbf{x}, \mathbf{y}$. Норма $\|\cdot\|_*$ является сопряжённой нормой для $\|\cdot\|$

- L является константой Липшица для градиента
- Интерес представляет минимально возможное L, при котором выполняется неравенство в определении
- На свойстве L-гладкости основан анализ многих численных методов

Пример

- Рассмотрим $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$, где $\mathbf{A} \in \mathbf{S}^n$

Пример

- Рассмотрим $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$, где $\mathbf{A} \in \mathbf{S}^n$
- Её градиент равен $f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$

Пример

- Рассмотрим $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$, где $\mathbf{A} \in \mathbf{S}^n$
- Её градиент равен $f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$
- Тогда по определению

$$\|\mathbf{A}\mathbf{x} - \mathbf{b} - \mathbf{A}\mathbf{y} + \mathbf{b}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{y}\|_2$$

так как для $\|\cdot\|_2$ выполнено $\|\cdot\|_* = \|\cdot\|_2$.

Пример

- Рассмотрим $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$, где $\mathbf{A} \in \mathbf{S}^n$
- Её градиент равен $f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$
- Тогда по определению

$$\|\mathbf{A}\mathbf{x} - \mathbf{b} - \mathbf{A}\mathbf{y} + \mathbf{b}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{y}\|_2$$

  так как для $\| \cdot \|_2$ выполнено $\| \cdot \|_* = \| \cdot \|_2$.
- Функция f является $\|\mathbf{A}\|_2$-гладкой

Пример

- Рассмотрим $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$, где $\mathbf{A} \in \mathbf{S}^n$
- Её градиент равен $f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$
- Тогда по определению

$$\|\mathbf{A}\mathbf{x} - \mathbf{b} - \mathbf{A}\mathbf{y} + \mathbf{b}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{y}\|_2$$

  так как для $\|\cdot\|_2$ выполнено $\|\cdot\|_* = \|\cdot\|_2$.
- Функция f является $\|\mathbf{A}\|_2$-гладкой
- $\|\mathbf{A}\|_2 = \lambda_{\max}(\mathbf{A})$

Пример

- Рассмотрим $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$, где $\mathbf{A} \in \mathbf{S}^n$
- Её градиент равен $f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$
- Тогда по определению

$$\|\mathbf{A}\mathbf{x} - \mathbf{b} - \mathbf{A}\mathbf{y} + \mathbf{b}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{y}\|_2$$

  так как для $\|\cdot\|_2$ выполнено $\|\cdot\|_* = \|\cdot\|_2$.
- Функция f является $\|\mathbf{A}\|_2$-гладкой
- $\|\mathbf{A}\|_2 = \lambda_{\max}(\mathbf{A})$
- Покажем, что $\|\mathbf{A}\|_2$ действительно минимально возможное значение L

Пример

- Рассмотрим $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$, где $\mathbf{A} \in \mathbf{S}^n$
- Её градиент равен $f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$
- Тогда по определению

$$\|\mathbf{A}\mathbf{x} - \mathbf{b} - \mathbf{A}\mathbf{y} + \mathbf{b}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{y}\|_2$$

  так как для $\|\cdot\|_2$ выполнено $\|\cdot\|_* = \|\cdot\|_2$.
- Функция f является $\|\mathbf{A}\|_2$-гладкой
- $\|\mathbf{A}\|_2 = \lambda_{\max}(\mathbf{A})$
- Покажем, что $\|\mathbf{A}\|_2$ действительно минимально возможное значение L
- Пусть f L-гладкая, тогда рассмотрим вектор $\mathbf{z}$ такой что $\|\mathbf{z}\|_2 = 1$ и $\|\mathbf{A}\mathbf{z}\|_2 = \|\mathbf{A}\|_2$

Пример

- Рассмотрим $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$, где $\mathbf{A} \in \mathbf{S}^n$
- Её градиент равен $f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$
- Тогда по определению

$$\|\mathbf{A}\mathbf{x} - \mathbf{b} - \mathbf{A}\mathbf{y} + \mathbf{b}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{y}\|_2$$

  так как для $\|\cdot\|_2$ выполнено $\|\cdot\|_* = \|\cdot\|_2$.
- Функция f является $\|\mathbf{A}\|_2$-гладкой
- $\|\mathbf{A}\|_2 = \lambda_{\max}(\mathbf{A})$
- Покажем, что $\|\mathbf{A}\|_2$ действительно минимально возможное значение L
- Пусть f L-гладкая, тогда рассмотрим вектор $\mathbf{z}$ такой что $\|\mathbf{z}\|_2 = 1$ и $\|\mathbf{A}\mathbf{z}\|_2 = \|\mathbf{A}\|_2$
- Тогда $\|\mathbf{A}\|_2 = \|\mathbf{A}\mathbf{z}\|_2 = \|f'(\mathbf{z}) - f'(0)\|_2 \leq L\|\mathbf{z} - 0\|_2 = L$

Пример

- Рассмотрим $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}$, где $\mathbf{A} \in \mathbf{S}^n$
- Её градиент равен $f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$
- Тогда по определению

$$\|\mathbf{A}\mathbf{x} - \mathbf{b} - \mathbf{A}\mathbf{y} + \mathbf{b}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x} - \mathbf{y}\|_2$$

  так как для $\| \cdot \|_2$ выполнено $\| \cdot \|_* = \| \cdot \|_2$.
- Функция f является $\|\mathbf{A}\|_2$-гладкой
- $\|\mathbf{A}\|_2 = \lambda_{\max}(\mathbf{A})$
- Покажем, что $\|\mathbf{A}\|_2$ действительно минимально возможное значение L
- Пусть f L-гладкая, тогда рассмотрим вектор $\mathbf{z}$ такой что $\|\mathbf{z}\|_2 = 1$ и $\|\mathbf{A}\mathbf{z}\|_2 = \|\mathbf{A}\|_2$
- Тогда $\|\mathbf{A}\|_2 = \|\mathbf{A}\mathbf{z}\|_2 = \|f'(\mathbf{z}) - f'(0)\|_2 \leq L\|\mathbf{z} - 0\|_2 = L$
- Значит $\|\mathbf{A}\|_2$ действительно наименьшее значение L

Лемма (descent lemma)

Descent lemma

Пусть f L-гладкая функция. Тогда для любых пар $\mathbf{x}, \mathbf{y}$

$$f(\mathbf{y}) \leq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$$

Доказательство

Лемма (descent lemma)

Descent lemma

Пусть f L-гладкая функция. Тогда для любых пар $\mathbf{x}, \mathbf{y}$

$$f(\mathbf{y}) \leq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$$

Доказательство

- По формуле Ньютона-Лейбница

$$f(\mathbf{y}) - f(\mathbf{x}) = \int_0^1 \langle f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})), \mathbf{y} - \mathbf{x} \rangle dt$$

Лемма (descent lemma)

Descent lemma

Пусть f L-гладкая функция. Тогда для любых пар $\mathbf{x}, \mathbf{y}$

$$f(\mathbf{y}) \leq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$$

Доказательство

- По формуле Ньютона-Лейбница

$$f(\mathbf{y}) - f(\mathbf{x}) = \int_0^1 \langle f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})), \mathbf{y} - \mathbf{x} \rangle dt$$

- $f(\mathbf{y}) - f(\mathbf{x}) =$
  $\langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \int_0^1 \langle f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle dt$

Лемма (descent lemma)

Descent lemma

Пусть f L-гладкая функция. Тогда для любых пар $\mathbf{x}, \mathbf{y}$

$$f(\mathbf{y}) \leq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$$

Доказательство

- По формуле Ньютона-Лейбница

$$f(\mathbf{y}) - f(\mathbf{x}) = \int_0^1 \langle f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})), \mathbf{y} - \mathbf{x} \rangle dt$$

- $f(\mathbf{y}) - f(\mathbf{x}) =$
  $\langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \int_0^1 \langle f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle dt$
- $|f(\mathbf{y}) - f(\mathbf{x}) - \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle| \leq \int_0^1 |\langle f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle| dt \leq \int_0^1 \|f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - f'(\mathbf{x})\| \|\mathbf{y} - \mathbf{x}\| dt$

Лемма (descent lemma)

Descent lemma

Пусть f L-гладкая функция. Тогда для любых пар $\mathbf{x}, \mathbf{y}$

$$f(\mathbf{y}) \leq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$$

Доказательство

- По формуле Ньютона-Лейбница

$$f(\mathbf{y}) - f(\mathbf{x}) = \int_0^1 \langle f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})), \mathbf{y} - \mathbf{x} \rangle dt$$

- $f(\mathbf{y}) - f(\mathbf{x}) =$
  $\langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \int_0^1 \langle f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle dt$
- $|f(\mathbf{y}) - f(\mathbf{x}) - \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle| \leq \int_0^1 |\langle f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) -$
  $f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle| dt \leq \int_0^1 \|f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - f'(\mathbf{x})\| \|\mathbf{y} - \mathbf{x}\| dt$
- $\int_0^1 \|f'(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - f'(\mathbf{x})\| \|\mathbf{y} - \mathbf{x}\| dt \leq$
  $\int_0^1 tL \|\mathbf{y} - \mathbf{x}\|_2^2 dt = \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$

L-гладкость и свойства гессиана

Критерий второго порядка

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема, тогда для заданного $L > 0$ следующие условия эквивалентны

- f является L-гладкой
- $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

L-гладкость и свойства гессиана

Критерий второго порядка

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема, тогда для заданного $L > 0$ следующие условия эквивалентны

- f является L-гладкой
- $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

Доказательство

L-гладкость и свойства гессиана

Критерий второго порядка

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема, тогда для заданного $L > 0$ следующие условия эквивалентны

- f является L-гладкой
- $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

Доказательство

1. Пусть f L-гладкая

L-гладкость и свойства гессиана

Критерий второго порядка

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема, тогда для заданного $L > 0$ следующие условия эквивалентны

- f является L-гладкой
- $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

Доказательство

1. Пусть f L-гладкая
    - Для любого направления $\mathbf{d}$ и $\alpha > 0$:
      $\|f'(\mathbf{x} + \alpha\mathbf{d}) - f'(\mathbf{x})\|_2 \leq \alpha L \|\mathbf{d}\|_2$

L-гладкость и свойства гессиана

Критерий второго порядка

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема, тогда для заданного $L > 0$ следующие условия эквивалентны

- f является L-гладкой
- $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

Доказательство

1. Пусть f L-гладкая
    - Для любого направления $\mathbf{d}$ и $\alpha > 0$: $\|f'(\mathbf{x} + \alpha\mathbf{d}) - f'(\mathbf{x})\|_2 \leq \alpha L \|\mathbf{d}\|_2$
    - $\lim\limits_{\alpha \to +0} \frac{\|f'(\mathbf{x}+\alpha\mathbf{d}) - f'(\mathbf{x})\|_2}{\alpha} = \|f''(\mathbf{x})\mathbf{d}\|_2 \leq L\|\mathbf{d}\|_2$

L-гладкость и свойства гессиана

Критерий второго порядка

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема, тогда для заданного $L > 0$ следующие условия эквивалентны

- f является L-гладкой
- $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

Доказательство

1. Пусть f L-гладкая
    - Для любого направления $\mathbf{d}$ и $\alpha > 0$: $\|f'(\mathbf{x} + \alpha\mathbf{d}) - f'(\mathbf{x})\|_2 \leq \alpha L \|\mathbf{d}\|_2$
    - $\lim\limits_{\alpha \to +0} \frac{\|f'(\mathbf{x}+\alpha\mathbf{d}) - f'(\mathbf{x})\|_2}{\alpha} = \|f''(\mathbf{x})\mathbf{d}\|_2 \leq L\|\mathbf{d}\|_2$
    - Так как это выполнено для любого $\mathbf{d}$, то $\|f''(\mathbf{x})\|_2 \leq L$

L-гладкость и свойства гессиана

Критерий второго порядка

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема, тогда для заданного $L > 0$ следующие условия эквивалентны

- f является L-гладкой
- $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

Доказательство

1. Пусть f L-гладкая
   - Для любого направления $\mathbf{d}$ и $\alpha > 0$: $\|f'(\mathbf{x} + \alpha\mathbf{d}) - f'(\mathbf{x})\|_2 \leq \alpha L\|\mathbf{d}\|_2$
   - $\lim\limits_{\alpha \to +0} \frac{\|f'(\mathbf{x}+\alpha\mathbf{d}) - f'(\mathbf{x})\|_2}{\alpha} = \|f''(\mathbf{x})\mathbf{d}\|_2 \leq L\|\mathbf{d}\|_2$
   - Так как это выполнено для любого $\mathbf{d}$, то $\|f''(\mathbf{x})\|_2 \leq L$
2. Пусть $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

L-гладкость и свойства гессиана

Критерий второго порядка

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема, тогда для заданного $L > 0$ следующие условия эквивалентны

- f является L-гладкой
- $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

Доказательство

1. Пусть f L-гладкая
   - Для любого направления $\mathbf{d}$ и $\alpha > 0$: $\|f'(\mathbf{x} + \alpha\mathbf{d}) - f'(\mathbf{x})\|_2 \leq \alpha L\|\mathbf{d}\|_2$
   - $\lim\limits_{\alpha \to +0} \frac{\|f'(\mathbf{x}+\alpha\mathbf{d}) - f'(\mathbf{x})\|_2}{\alpha} = \|f''(\mathbf{x})\mathbf{d}\|_2 \leq L\|\mathbf{d}\|_2$
   - Так как это выполнено для любого $\mathbf{d}$, то $\|f''(\mathbf{x})\|_2 \leq L$
2. Пусть $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$
   - По формуле Ньютона-Лейбница $f'(\mathbf{y}) - f'(\mathbf{x}) = \int_0^1 f''(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})dt$

L-гладкость и свойства гессиана

Критерий второго порядка

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ дважды непрерывно дифференцируема, тогда для заданного $L > 0$ следующие условия эквивалентны

- f является L-гладкой
- $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$

Доказательство

1. Пусть f L-гладкая
   - Для любого направления $\mathbf{d}$ и $\alpha > 0$: $\|f'(\mathbf{x} + \alpha\mathbf{d}) - f'(\mathbf{x})\|_2 \leq \alpha L \|\mathbf{d}\|_2$
   - $\lim\limits_{\alpha \to +0} \frac{\|f'(\mathbf{x}+\alpha\mathbf{d}) - f'(\mathbf{x})\|_2}{\alpha} = \|f''(\mathbf{x})\mathbf{d}\|_2 \leq L\|\mathbf{d}\|_2$
   - Так как это выполнено для любого $\mathbf{d}$, то $\|f''(\mathbf{x})\|_2 \leq L$
2. Пусть $\|f''(\mathbf{x})\|_2 \leq L$ для любого $\mathbf{x}$
   - По формуле Ньютона-Лейбница $f'(\mathbf{y}) - f'(\mathbf{x}) = \int_0^1 f''(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})dt$
   - $\|f'(\mathbf{y}) - f'(\mathbf{x})\|_2 \leq \left(\int_0^1 \|f''(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))\|_2 dt\right) \|\mathbf{y} - \mathbf{x}\|_2 \leq L\|\mathbf{y} - \mathbf{x}\|_2$

L-гладкая выпуклая функция

Утверждение

Пусть f дифференцируемая выпуклая функция и $L > 0$. Тогда следующие условия эквивалентны:

L-гладкая выпуклая функция

Утверждение

Пусть f дифференцируемая выпуклая функция и $L > 0$. Тогда следующие условия эквивалентны:

- f является L-гладкой

L-гладкая выпуклая функция

Утверждение

Пусть f дифференцируемая выпуклая функция и $L > 0$. Тогда следующие условия эквивалентны:

- f является L-гладкой
- $f(\mathbf{y}) \leq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$ для любых пар $\mathbf{x}, \mathbf{y}$

L-гладкая выпуклая функция

Утверждение

Пусть f дифференцируемая выпуклая функция и $L > 0$. Тогда следующие условия эквивалентны:

- f является L-гладкой
- $f(\mathbf{y}) \leq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$ для любых пар $\mathbf{x}, \mathbf{y}$
- $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2L} \|f'(\mathbf{x}) - f'(\mathbf{y})\|_2^2$ для любых пар $\mathbf{x}, \mathbf{y}$

L-гладкая выпуклая функция

Утверждение

Пусть f дифференцируемая выпуклая функция и $L > 0$. Тогда следующие условия эквивалентны:

- f является L-гладкой
- $f(\mathbf{y}) \leq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2}\|\mathbf{y} - \mathbf{x}\|_2^2$ для любых пар $\mathbf{x}, \mathbf{y}$
- $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2L}\|f'(\mathbf{x}) - f'(\mathbf{y})\|_2^2$ для любых пар $\mathbf{x}, \mathbf{y}$
- $\langle f'(\mathbf{x}) - f'(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \geq \frac{1}{L}\|f'(\mathbf{x}) - f'(\mathbf{y})\|_2^2$

L-гладкая выпуклая функция

Утверждение

Пусть f дифференцируемая выпуклая функция и $L > 0$. Тогда следующие условия эквивалентны:

- f является L-гладкой
- $f(\mathbf{y}) \le f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{L}{2}\|\mathbf{y} - \mathbf{x}\|_2^2$ для любых пар $\mathbf{x}, \mathbf{y}$
- $f(\mathbf{y}) \ge f(\mathbf{x}) + \langle f'(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2L}\|f'(\mathbf{x}) - f'(\mathbf{y})\|_2^2$ для любых пар $\mathbf{x}, \mathbf{y}$
- $\langle f'(\mathbf{x}) - f'(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \ge \frac{1}{L}\|f'(\mathbf{x}) - f'(\mathbf{y})\|_2^2$
- $f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \ge \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) - \frac{L}{2}\lambda(1-\lambda)\|\mathbf{x} - \mathbf{y}\|_2^2$ для любых пар $\mathbf{x}, \mathbf{y}$

Главное

- Непрерывность выпуклых функций

Главное

- Непрерывность выпуклых функций
- Производная по направлению

Главное

- Непрерывность выпуклых функций
- Производная по направлению
- L-гладкость выпуклых функций