1 Основы гладкой оптимизации

1.1 Обзорная лекция по методам оптимизации.

- Предмет оптимизации.
- Типы задачи оптимизации.
- Решение задачи оптимизации.
- Обзор градиентных методов.
- Обзор методов математической оптимизации для задач с ограничениями.
- Эвристические методы.
- Метод Нелдера-Мида (Nelder-Mead).
- Литература.

Задачи.

1. Доказать, что при рассмотрении задачи оптимизации для квадратичной функции без ограничения общности матрицу A можно считать симметричной.
2. Доказать, что минимум квадратичной функции $f(x) = \frac{1}{2}x^T Ax - b^T x + c$ существует тогда и только тогда, когда матрица A положительно определена.
3. Найти оценку скорости сходимости для квадратичной функции $f(x) = \frac{1}{2}x^T Ax - b^T x + c$.
4. Вычислить $\partial f(x)$ для $f(x) = |x|$. Показать, что для $f(x) = |x|$ при постоянном шаге субградиентный метод не сходится.
5. Вычислить производную по направлению $f'(0; y)$ для $f(x) = \|x\|$.
6. Найти $\nabla f(x)$ для $f(x) = \|x\|$ при $x \neq 0$.
7. Найти $\nabla f(x)$ для $f(x) = \|x_+\|^2$.
8. Найти $\nabla \|(Ax - b)_+\|$.

1.2 Общая постановка задачи оптимизации. Аналитическая и арифметическая сложность. Оценка сложности перебора. [N, §1.1]

- Общая постановка задачи оптимизации.
- Локальное и глобальное решение.
- Классификация задач оптимизации (линейная, гладкая, уловная, выпуклая...).
- Строго допустимая задача - уловие Слейтера.
- Общие примеры задач оптимизации.
- Понятие Оракула. Типы оракула: 0-го порядка, 1-го порядка, 2-го порядка.
- Аналитическая сложность и аримметическая сложность.
- Постановка задачи оптимизации в $B_n \subset \mathbb{R}^n$ для l_∞-липшицевой функции $f(x)$.
- Метод перебора (по p-сетке). Верхние оценки точности и сложности.

$$f(\overline{x}) - f^* \leq \frac{L}{2p}$$

- Понятие сопротивляющегося Оракула. Ничние оценки сложности.
- Асимптоическая оптимальность метода перебора на классе рассматриваемых задач.
- Неразрешимость общей задачи нелинейной оптимизации: оценка времени работы на примере.

Задачи.

1. $^\circ$ Пусть $f(x) : B_n \to \mathbb{R}$ удовлетворяет уловию Липшица (l_∞) с константой L. Пусть f^* — оптимальное значение функции $f(x)$ в посталвенной задаче. Пусть $\overline{x} = \operatorname{argmin} f(x)$ на p-сетке. Доказать, что $f(\overline{x}) - f^* \leq \frac{L}{2p}$.
2. $^\circ$ Доказать, что аналитическая сложность задачи 1. для метда перебора по p-сетке не превосходит $\left(\left[\frac{L}{2\varepsilon}\right] + 2\right)^n$.
3. $^\circ$ Пусть $\varepsilon < L/2$. Тогда аналитическая сложность задачи 1. не менее $\left(\left[\frac{L}{2\varepsilon}\right]\right)^n$.
4. $^\circ$ Оценить время работы метода перебора (количество лет) при следующих значениях параметров: $L = 2, n = 10, \varepsilon = 0,01$, производительность компьютера 10^9 арифм. опер. в сек.

1.3 Необходимые условия экстремума 1-го порядка. Необходимые условия экстремума 2-го порядка. Достаточные условия экстремума 2-го порядка. [N, §1.2.1]

- Идея релаксации.
- Постановка задачи безусловной гладкой минимизации.
- Линейные аппроксимации функции.
- Множество уровней $\mathcal{L}_f(a)$ функции $f(x)$.
- Локальное поведение функции $f(x)$ вдоль направления $s \in \mathbb{R}^n$, $||s|| = 1$:

$$\Delta(s) = \lim_{t \to 0} \frac{f(\overline{x} + ts) - f(\overline{x})}{t} = (f'(\overline{x}), s).$$

- Необходимые условия экстремума 1-го порядка.
- Понятие стационарной точки.
- Квадратичные аппроксимации функции. Гессиан.
- Положительно определенная матрица $A > 0$.
- Необходимые условия экстремума 2-го порядка.

$$\nabla f(x^*) = 0, \quad \nabla^2 f(x^*) \succeq 0.$$

- Достаточные условия экстремума 2-го порядка.

$$\nabla f(x^*) = 0, \quad \nabla^2 f(x^*) \succ 0.$$

Задачи.

1. Доказать, что функция $f : Q \to \mathbb{R}$ имеет глобальное решение

$$Q = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 > 0\},$$
$$f(x_1, x_2) = x_1^2 + x_2 + \frac{1}{x_1 + x_2}.$$

2. С помощью условий первого и второго порядка найти экстремальные точки функции $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1x_2.$$

3. Найти глобальные решения

$$Q = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 > 0,\ x_2 > 0,\ x_3 > 0\},$$
$$f(x_1, x_2, x_3) = x_1 + \frac{x_2^2}{4x_1} + \frac{x_3^2}{x_2} + \frac{2}{x_3} \ \to\ \min.$$

4. Построить пример функции $f \in C^1(\mathbb{R}^n)$, имеющей ровно одну стационарную точку, являющуюся ее локальным, но не глобальным экстремумом. Существует ли такой пример при $n = 1$?

5. Привести пример функции $f \in C[a, b]$, не удовлетворяющей условию Липшица (0-го порядка) ни для какого L.

6. Доказать, что при любом значении параметра $\sigma > 1$ система уравнений относительно $x \in \mathbb{R}^2$ имеет ненулевое решение
$$\begin{cases} \sigma \cos x_1 \sin x_2 + x_1 e^{x_1^2+x_2^2} = 0, \\ \sigma \sin x_1 \cos x_2 + x_2 e^{x_1^2+x_2^2} = 0. \end{cases}$$

1.4 Классы дифференцируемых функций. [N, §1.2.2]

- Класс функций $C_L^{k,p}(\mathbb{R}^n)$.
- Неравенства для функций из $C_L^{2,1}(\mathbb{R}^n)$ и $C_L^{1,1}(\mathbb{R}^n)$.

$$f(x+y) = f(x) + \big(\nabla f(x), y\big) + \int_0^1 \big(\nabla f(x+\tau y) - \nabla f(x), y\big)\ \mathrm{d}\tau,$$

$$f(x+y) = f(x) + \big(\nabla f(x), y\big) + \int_0^1 \int_0^t \big(\nabla^2 f(x+\tau y)y, y\big)\ \mathrm{d}\tau\ \mathrm{d}t,$$

$$f(x) \in C_L^{k,p}(\mathbb{R}^n) \quad <=> \quad \|\nabla^p f(x) - \nabla^p f(y)\| \le L\|x-y\|,$$

$$f(x) \in C_L^{2,1}(\mathbb{R}^n) \quad <=> \quad \|\nabla^2 f(x)\| \le L,$$

$$f(x) \in C_L^{1,1}(\mathbb{R}^n) \quad => \quad |f(y) - f(x) - \big(\nabla f(x), y-x\big)| \le \frac{L}{2}\|y-x\|^2,$$

$$f(x) \in C_L^{2,2}(\mathbb{R}^n) \quad => \quad \|\nabla f(y) - \nabla f(x) - \nabla^2 f(x)(y-x)\| \le \frac{M}{2}\|y-x\|^2,$$

$$f(x) \in C_L^{2,2}(\mathbb{R}^n) \quad => \quad |f(y) - f(x) - \big(\nabla f(x), y-x\big) - \big(\nabla^2 f(x)(y-x), y-x\big)\| \le \frac{M}{6}\|y-x\|^3,$$

$$\nabla^2 f(x) - MrI_n \preceq \nabla^2 f(y) \preceq \nabla^2 f(x) + MrI_n.$$

Задачи.

1. Пусть $\phi : \mathbb{R}^n \to \mathbb{R}^m$, $f : \mathbb{R}^m \to \mathbb{R}$ дифференцируемые функции и $h(x) = f(\phi(x))$. Выписать $\mathrm{d}h(x)$ и $\mathrm{d}^2 h(x)$ через $\mathrm{d}\phi_j$ и $\mathrm{d}x$.

2. Пусть $f \in C^2(\mathbb{R}^n)$. Доказать, что

   а)
   $$f(y) = f(x) + \int_0^1 \big(f'(x+t(y-x)),\ (y-x)\big)\mathrm{d}t;$$

   б)
   $$f'(y) = f'(x) + \int_0^1 f''(x+t(y-x))(y-x)\mathrm{d}t;$$

3. Найти $f'(x)$ и $f''(x)$ для $f(x) = \alpha + (a,x) + \frac{1}{2}(Ax,x)$.

4. Пусть $f \in C^1(\mathbb{R}^m)$, $\phi(x) = f(Ax+b)$, $b \in \mathbb{R}^m$, матрица $A : \mathbb{R}^n \to \mathbb{R}^m$. Доказать, что для $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$:

   а) $(y, Ax) = (A^T y, x)$;

   б) $\phi'(x) = A^T f'(Ax+b)$.

5. Пусть $f \in C_M^{2,2}(\mathbb{R}^n)$. Тогда для любых $x, y \in \mathbb{R}^n$ имеет место неравенства

$$|f(y) - f(x) - \big(f'(x), y-x\big) - \frac{1}{2}\big(f''(x)(y-x), y-x\big)| \le \frac{M}{6}||y-x||^3.$$

2 Методы гладкой оптимизации

2.1 Градиентные методы. [N, §1.2.3-1.2.4]

- Градиентный метод $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$.
- Варианты выбора длины шага h_k.

$$\alpha_k = h, \quad \alpha_k = \frac{h}{\sqrt{k+1}},$$
$$\alpha_k = \arg\min_{h\geq 0} f\big(x_k - h\nabla f(x_k)\big),$$
$$\begin{cases} a\big(\nabla f(x_k), x_k - x_{k+1}\big) \leq f(x_k) - f(x_{k+1}), \\ b\big(\nabla f(x_k), x_k - x_{k+1}\big) \geq f(x_k) - f(x_{k+1}). \end{cases}$$

- Оценки убывания целевой функции.

$$f(x) \in C_L^{1,1}(\mathbb{R}^n) \quad => \quad f(y) \leq f(x) - \frac{1}{2L}\|\nabla f(x)\|^2.$$

- Верхние оценка скорости сходимости для $f \in C_L^{1,1}(\mathbb{R}^n)$.

$$\min_{1\leq k\leq N} \|\nabla f(x_k)\| \leq \frac{1}{N+1}\left(\frac{L}{\omega}\big(f(x_0) - f^*\big)\right)^{1/2} \leq \varepsilon, \quad => \quad N+1 \geq \frac{L}{\omega\varepsilon^2}\big(f(x_0) - f^*\big).$$

- Пример, когда градиентный метод не сходится к точкам локального минимума.
- Верхняя оценка скорости сходимости для $f \in C_M^{2,2}(\mathbb{R}^n)$: линейная скорость.

$$lI_n \preceq \nabla^2 f(x) \preceq LI_n \quad => \quad \|x_k - x^*\| \leq \frac{rr_0}{r - r_0}\left(1 - \frac{2l}{L+3l}\right).$$

- Метод Ньютона $x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k)$.
- Вывод через квадратичную аппроксимацию.
- Пример сходимости и расходимости метода Ньютона.
- Демпфированный метод Ньютона с длиной шага h_k.
- Квадратичная сходимость метода Ньютона.

$$f \in C_M^{2,2}(\mathbb{R}^n),\ lI_n \preceq \nabla^2 f(x) \quad => \quad r_{k+1} \leq \frac{Mr_k^2}{2(l - Mr_k)}, \quad r_k = x_k - x^*.$$

- Классификация скорости сходимости метода.

Задачи.

1. Из начального приближения $x^0 = (-2, 1)$ сделать два шага метода скорейшего спуска для задачи безусловной оптимизации
$$f(x_1, x_2) = x_1^2 + x_1 x_2 + 2x_2^2 + x_1 \ \to \ \min .$$

2. Из начального приближения $x^0 = (1, -1)$ сделать шаг градиентного метода по правилу Голдштейна-Армийо (подобрать параметры α, β) для задачи безусловной оптимизации
$$f(x_1, x_2) = 2x_1^2 + x_1 x_2 + 3x_2^2 \ \to \ \min .$$

3. Из начального приближения $x^0 = 1$ сделать два шага метода Ньютона для уравнения $x^p = 0$, где $p \geq 2$ — целочисленный параметр. Проанализировать скорость сходимости, объяснить наблюдаемый эффект. Модифицировать метод Ньютона, введя параметр длины шага, равный p. Проанализировать скорость сходимости.

4. Из начального приближения $x^0 = (1, 1)$ сделать два шага метода Ньютона для задачи безусловной оптимизации
$$f(x_1, x_2) = x_1^2 + e^{x_2^2} \ \to \ \min .$$
Проанализировать сходимость. Показать, что в данном случае метод Ньютона сходится глобально.

5. * Установить верхние границы аналитической сложности (количество итераций для ε-решения) для сублинейной, линейной и квадратичной скорости сходимости.

2.2 Квазиньютоновские методы. [N, §1.3.1]

- Сравнение градиентного метода и метода Ньютона.
- Квазиньютоновские методы.

Задачи.

1. * Проверить, что следующие квазиньютоновские правила удовлетворяют соотношению

$$H_{k+1}\big(f'(x_{k+1}) - f'(x_k)\big) = x_{k+1} - x_k,$$

$$1)\ \Delta H_k = \frac{(\delta_k - H_k\gamma_k)(\delta_k - H_k\gamma_k)^T}{(\delta_k - H_k\gamma_k,\ \gamma_k)},$$

$$2)\ \Delta H_k = \frac{\delta_k\delta_k^T}{(\gamma_k,\ \delta_k)} - \frac{H_k\gamma_k\gamma_k^T H_k}{(H_k\gamma_k,\ \gamma_k)},$$

$$3)\ \Delta H_k = \frac{H_k\gamma_k\delta_k^T + \delta_k\gamma_k^T H_k}{(H_k\gamma_k,\ \gamma_k)} - \beta_k\frac{H_k\gamma_k\gamma_k^T H_k}{(H_k\gamma_k,\ \gamma_k)}, \quad \beta_k = 1 + \frac{(\gamma_k,\ \delta_k)}{(H_k\gamma_k,\ \gamma_k)},$$

где $\quad \Delta H_k = H_{k+1} - H_k, \quad \gamma_k = f'(x_{k+1}) - f'(x_k), \quad \delta_k = x_{k+1} - x_k.$

2.3 Метод сопряженных градиентов. [N, §1.3.2]

- Линейные подпространства Крылова.
- Метод сопряженных градиентов.

Задачи.

1. * Оценить количество шагов для решения задачи минимизации квадратичной функции $f(x) = \alpha + (a, x) + \frac{1}{2}(Ax, x)$ методом сопряженных градиентов:

$$x_{k+1} = x_k + h_k p_k, \quad p_{k+1} = f'(x_{k+1}) - \beta_k p_k, \quad \text{где}$$

$$1)\ \beta_k = \frac{||f'(x_{k+1})||^2}{(f'(x_{k+1}) - f'(x_k), p_k)},$$

$$2)\ \beta_k = -\frac{||f'(x_{k+1})||^2}{||f'(x_k)||^2},$$

$$3)\ \beta_k = -\frac{(f'(x_{k+1}), f'(x_{k+1}) - f'(x_k))}{||f'(x_k)||^2}.$$

2.4 Условная гладкая минимизация. [N, §1.3.3]

- Сравнение условной и безусловной минимизации.
- Метод штрафных функций.
- Метод барьерных функций.

3 Основы гладкой выпуклой оптимизации

3.1 Теория выпуклых функций. [N, §2.1.1]

- Класс гладких выпуклых функций $\mathbb{F}_L^{k,l}(Q)$.
- Существование глобального минимума.
- Эквивалентные определения выпуклой функции из $\mathbb{F}^1(\mathbb{R}^n)$.
- Эквивалентные условия для функций из $\mathbb{F}^2(\mathbb{R}^n)$.

Задачи.

1. Пусть $f \in C[0,1]$. Доказать равенство $\lim_{\tau\to 0}\int_0^\tau f(x)\,\mathrm{d}x = f(0)$.
2. Доказать, что следующие функции лежат в классе $\mathcal{F}^1(\mathbb{R}^n)$:
   а) $f(x) = \alpha + (a,x) + \frac{1}{2}(Ax,x)$;
   б) $f(x) = e^x$, $n = 1$;
   в) $f(x) = |x|^p$, $p > 1$, $n = 1$;
   г) $f(x) = \frac{x^2}{1-|x|}$, $n = 1$;
   д) $f(x) = |x| - \ln(1+|x|)$, $n = 1$;
   е) $f(x) = \sum_{i=1}^m e^{\alpha_i + (a_i,x)}$;
   ж) $f(x) = \sum_{i=1}^m |(a_i,x) - b_i|^p$, $p \geq 1$;
3. * Доказать, что следующие условия на функцию $f \in C^1(\mathbb{R}^n)$, выполняющиеся для всех $x, y \in \mathbb{R}^n$ и $\alpha \in [0,1]$, эквивалентны включению $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$:
   а)
   $$0 \leq f(y) - f(x) - \big(f'(x), y - x\big) \leq \frac{L}{2}\|x-y\|^2;$$
   б)
   $$f(x) + \big(f'(x), y - x\big) + \frac{L}{2}\|f'(x) - f'(y)\|^2 \leq f(y);$$
   в)
   $$\frac{1}{L}\|f'(x) - f'(y)\|^2 \leq \big(f'(x) - f'(y), x - y\big);$$
   г)
   $$0 \leq \big(f'(x) - f'(y), x - y\big) \leq L\|x-y\|^2;$$
   д)
   $$\alpha f(x) + (1-\alpha)f(y) \leq f(\alpha x + (1-\alpha)y) + \frac{\alpha(1-\alpha)}{2L}\|f'(x) - f'(y)\|^2;$$

e)

$$\alpha f(x) + (1-\alpha) f(y) - f(\alpha x + (1-\alpha) y) \le \frac{\alpha(1-\alpha)L}{2}\|x-y\|^2.$$

4. * Доказать, что дважды непрерывно дифференцируемая функция $f(x)$ лежит в классе $\mathcal{F}_L^{2,1}(\mathbb{R}^n)$ тогда и только тогда, когда для любого $x \in \mathbb{R}^n$ выполнено условие $0 \preceq f''(x) \preceq LI_n$.

3.2 Нижние границы аналитической сложности для класса $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$. [N, §2.1.2]

- Постановка задачи об оценке аналитической сложности
- Построение и исследование “наихудшей функции” в классе $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$.
- Получение нижних границ аналитической сложности.

Задачи.

1. ° Пусть $L > 0$. Рассмотрим семейство квадратичных функций

$$f_k(x) = \frac{L}{4}\left(\frac{1}{2}\left((x^{(1)})^2 + \sum_{i=1}^{k-1}(x^{(i)} - x^{(i+1)})^2 + (x^{(k)})^2\right) - x^{(1)}\right), \quad k = 1, 2, \ldots, n.$$

   1.1 Доказать, что $0 \preceq f_k''(x) \preceq LI_n$.

   1.2 Найти явное представление матрицы $f_k''(x)$.

   1.3 Найти точку $\overline{x}_k$ минимума функции $f_k(x)$ и найти $f_k^* = f_k(\overline{x}_k)$.

2. ° Пусть $x_0 = 0$ и функция f_p определена как в предыдущей задаче. Доказать, что для любой последовательности $\{x_k\}_{k=0}^p$, удовлетворяющей условию $x_k \in \mathcal{L}_k = \mathrm{Lin}\{f_p'(x_0), \ldots, f_p'(x_k)\}$, имеет место включение $\mathcal{L}_k \subset \mathbb{R}^{k,n}$.

3. ° Доказать, что для любого k, $1 \le k \le (n-1)/2$, и любого $x_0 \in \mathbb{R}^n$ найдется такая функция $f \in \mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$, что для любого метода первого порядка, удовлетворяющего условию $[x_k \in x_0 + \mathrm{Lin}\{f'(x_0), \ldots, f'(x_{k-1})\}$, $k \ge 1]$, выполняются неравенства

$$f(x_k) - f^* \ge \frac{3L||x_0 - x^*||^2}{32(k+1)^2},$$
$$||x_k - x^*||^2 \ge \frac{1}{8}||x_0 - x^*||^2,$$

   где x^* — точка минимума функции $f(x)$, а $f^* = f(x^*)$. (*Указание*: рассмотреть $f = f_{2k+1}$, $x_0 = 0$.)

3.3 Сильно выпуклые функции. [N, §2.1.3]

- Класс сильно выпуклых функций $\mathcal{J}_\mu^1(\mathbb{R}^n)$.
- Эквивалентные определения сильно выпуклой функции из $\mathcal{J}_\mu^1(\mathbb{R}^n)$.
- Эквивалентные условия для функций из $\mathbb{F}^2(\mathbb{R}^n)$.
-

Задачи.

1. $^\circ$ Доказать, что следующие условия на функцию $f \in C^1(\mathbb{R}^n)$, выполняющиеся для всех $x, y \in \mathbb{R}^n$ и $\alpha \in [0,1]$, эквивалентны включению $f \in \mathcal{J}_\mu^1(\mathbb{R}^n)$:

   а)
$$\big(f'(x) - f'(y), x - y\big) \geq \mu\|x - y\|^2;$$

   б)
$$\alpha f(x) + (1-\alpha)f(y) \geq f(\alpha x + (1-\alpha)y) + \alpha(1-\alpha)\frac{\mu}{2}\|x - y\|^2.$$

2. $^\circ$ Пусть $f \in C^2(\mathbb{R}^n)$. Доказать, что $f \in \mathcal{J}_\mu^2(\mathbb{R}^n)$ тогда и только тогда, когда $f''(x) \geq \mu I_n$.

3. Какие дополнительные условия необходимы, чтобы функция $f(x) = \alpha + (a, x) + \frac{1}{2}(Ax, x)$ принадлежала классу $\mathcal{J}_{\mu,L}^{\infty,1}(\mathbb{R}^n)$.

4. * Доказать, что если $f \in \mathcal{J}_{\mu,L}^{1,1}(\mathbb{R}^n)$, то для любых $x, y \in \mathbb{R}^n$ выполняется неравенство
$$\big(f'(x) - f'(y), x - y\big) \geq \frac{\mu L}{\mu + L}||x - y||^2 + \frac{1}{\mu + L}||f'(x) - f'(y)||^2.$$

3.4 Нижние границы аналитической сложности для класса $\mathcal{J}^{\infty,1}_{\mu,L}(\mathbb{R}^n)$. [N, §2.1.4]

- Класс сильно выпуклых функций $\mathcal{J}^1_\mu(\mathbb{R}^n)$.
- Эквивалентные определения сильно выпуклой функции из $\mathcal{J}^1_\mu(\mathbb{R}^n)$.
- Нижние границы аналитической сложности.

Задачи.

1. $^\circ$ Рассмотрим $\mathbb{R}^\infty = \ell_2$ — пространство всех последовательностей $x = \{x^{(i)}\}_{i=1}^\infty$ с конечной нормой $||x||^2 = \sum_{i=1}^\infty (x^{(i)})^2 < \infty$. Выберем параметры $\mu > 0$ и $Q_f > 1$ и определим

$$f(x) = \frac{\mu(Q_f - 1)}{8}\left((x^{(1)})^2 + \sum_{i=1}^{\infty}(x^{(i)} - x^{(i+1)})^2 - 2x^{(1)}\right) + \frac{\mu}{2}||x||^2.$$

   1.1 Найти явное представление матрицы $f''(x)$.

   1.2 Доказать, что $\mu I \preceq f''(x) \preceq \mu Q_f I$.

   1.3 Найти точку $x^* \in \mathbb{R}^\infty$ минимума функции $f(x)$ и найти $f^* = f(x^*)$.

2. * Доказать, что для любого $x_0 \in \mathbb{R}^\infty$ и произвольных констант $\mu > 0$ и $Q_f > 1$ найдется такая функция $f \in \mathcal{J}^{\infty,1}_{\mu,\mu Q_f}(\mathbb{R}^\infty)$, что для любого метода первого порядка $\mathcal{M}$, удовлетворяющего условию $[x_k \in x_0 + \mathrm{Lin}\{f'(x_0), \ldots, f'(x_{k-1})\},\ k \geq 1]$, выполняются неравенства

$$||x_k - x^*||^2 \geq \left(\frac{\sqrt{Q_f} - 1}{\sqrt{Q_f} + 1}\right)^{2k} ||x_0 - x^*||^2,$$

$$f(x_k) - f^* \geq \frac{\mu}{2}\left(\frac{\sqrt{Q_f} - 1}{\sqrt{Q_f} + 1}\right)^{2k} ||x_0 - x^*||^2,$$

   где x^* — точка минимума функции $f(x)$, а $f^* = f(x^*)$. (*Указание*: рассмотреть f из задачи 1 и доказать, что $x_k \in \mathbb{R}^{k,\infty}$.)

4 Методы гладкой выпуклой оптимизации

4.1 Градиентный метод для выпуклых функций. [N, §2.1.5]

- Верхняя оценка $f(x_k) - f^{\infty}$ для градиентного метода и функций из $\mathcal{F}_L^{1,1}(\mathbb{R}^n)$.
- Верхняя оценка $||x_k - x^*||$ и $f(x_k) - f^{\infty}$ для градиентного метода и функций из $\mathcal{J}_L^{1,1}(\mathbb{R}^n)$.
- Нижние границы аналитической сложности.

Задачи.

1. Провести сравнительный анализ верхних оценок для градиентного метода для функций из классов $C_L^{\infty,1}(\mathbb{R}^n)$, $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$, $\mathcal{J}_L^{\infty,1}(\mathbb{R}^n)$ между собой и с нижними оценками аналитической сложности для этих классов.

4.2 Оптимальные методы для гладких выпуклых функций. [N, §2.2.1]

- Оценивающие последовательности.
- Общая схема оптимального метода и верхняя оценка на $f(x_k) - f^*$.

Задачи.

1. * Пусть $\phi_0(x) = \phi_0^* + \frac{\gamma_0}{2}||x - v_0||^2$. Пусть $f \in \mathcal{J}^{1,1}_{\mu,L}(\mathbb{R}^n)$, $\alpha_k \in (0,1)$, $\sum_{k=0}^{\infty} \alpha_k = \infty$, $\lambda_0 = 1$, и $y_k \in \mathbb{R}^n$ — произвольные. Определим

$$\lambda_{k+1} = (1 - \alpha_k)\lambda_k,$$
$$\phi_{k+1}(x) = (1 - \alpha_k)\phi_k(x) + \alpha_k\big(f(y_k) + (f'(y_k), x - y_k) + \frac{\mu}{2}||x - y_k||^2\big).$$

   Доказать, что $\phi_k(x)$ имеет вид $\phi_k(x) = \phi_k^* + \frac{\gamma_0}{2}||x - v_k||^2$ и явно найти рекуррентные выражения для $\gamma_{k+1}, v_{k+1}, \phi_{k+1}^*$.

2. * Доказать, что общая схема с оценивающей последовательностью дает оптимальный метод для функций из класса $\mathcal{J}^{1,1}_{\mu,L}(\mathbb{R}^n)$. (*Указание:* сравнить верхнюю оценку метода с нижней оценкой для класса.)