Студент: Рыжиков Иван
Группа: 2381
Дата: 7 октября 2024 г.

---

# Методы оптимизации

Прим. во всех задачах используется обозначение $\boldsymbol{\nabla} f(\mathbf{x})$ — градиент функции $f(\mathbf{x})$, и $\langle \mathbf{x}, \mathbf{y} \rangle$ — скалярное произведение $\mathbf{x}$ и $\mathbf{y}$.

**Задание 7.** *Зная формулу Ньютона-Лейбница:*

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

*где $F(x)$ - первообразная функции $f(x)$, непрерывной на отрезке $[a, b]$.*
*Доказать, что:*

$$g(\mathbf{x} + \mathbf{y}) = g(\mathbf{x}) + \int_0^1 \langle \nabla g(\mathbf{x} + t\mathbf{y}), \mathbf{y} \rangle \; dt$$

*Доказательство.* Рассмотрим функцию $g(\mathbf{x} + t\mathbf{y})$ зависящую от параметра $t$.

- Продифференцируем её по $t$ по правилу дифференцирования сложной функции (chain rule):
$$\frac{d}{dt}(g(\mathbf{x} + t\mathbf{y})) = \langle \boldsymbol{\nabla} g(\mathbf{x} + t\mathbf{y}), \mathbf{y} \rangle$$
- Проинтегрируем обе части уравнения по $t$ от 0 до 1:
$$\int_0^1 \frac{d}{dt}(g(\mathbf{x} + t\mathbf{y}))\,dt = \int_0^1 \langle \boldsymbol{\nabla} g(\mathbf{x} + t\mathbf{y}), \mathbf{y} \rangle \; dt$$
- По формуле Ньютона-Лейбница левая часть равна:
$$\int_0^1 \frac{d}{dt}(g(\mathbf{x} + t\mathbf{y}))\,dt = g(\mathbf{x} + \mathbf{y}) - g(\mathbf{x})$$
- Получаем:
$$g(\mathbf{x} + \mathbf{y}) - g(\mathbf{x}) = \int_0^1 \langle \boldsymbol{\nabla} g(\mathbf{x} + t\mathbf{y}), \mathbf{y} \rangle \; dt$$
- Переносим $g(\mathbf{x})$ в правую часть:
$$g(\mathbf{x} + \mathbf{y}) = g(\mathbf{x}) + \int_0^1 \langle \boldsymbol{\nabla} g(\mathbf{x} + t\mathbf{y}), \mathbf{y} \rangle \; dt$$

■

**Задание 8.** *Для метода наискорейшего спуска доказать, что:*

$$\frac{d}{dt}(f(\mathbf{x} - t\boldsymbol{\nabla} f(\mathbf{x})))|_{t=t_k} = \langle \boldsymbol{\nabla} f(\mathbf{x}_{k+1}), \boldsymbol{\nabla} f(\mathbf{x}_k) \rangle$$

*Доказательство.* Метод наискорейшего спуска заключается в минимизации функции $f(\mathbf{x})$ по направлению антиградиента. Точка минимума функции $f(\mathbf{x})$ находится по формуле:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \boldsymbol{\nabla} f(\mathbf{x}_k)$$

где $\alpha_k$ — шаг метода вдоль антиградиента и $\boldsymbol{\nabla} f(\mathbf{x}_k)$ — градиент функции $f$ в точке $\mathbf{x}_k$.

- Введем функцию
$$\mathbf{x}(t) = \mathbf{x}_k - t\boldsymbol{\nabla} f(\mathbf{x}_k)$$
как некоторую кривую, проходящую через точку $\mathbf{x}_k$ в направлении антиградиента на шаг $t$.

- Далее найдем производную функции $f\left(\mathbf{x}\left(t\right)\right)$ по $t$, используя правило дифференцирования сложной функции (chain rule):

$$\frac{d}{dt}f\left(\mathbf{x}\left(t\right)\right)=\left\langle \boldsymbol{\nabla} f\left(\mathbf{x}\left(t\right)\right),\frac{d}{dt}\mathbf{x}\left(t\right)\right\rangle$$

- Так как $\frac{d}{dt}\left(\mathbf{x}\left(t\right)\right)=-\boldsymbol{\nabla} f\left(\mathbf{x}_k\right)$, то производная функции $f\left(\mathbf{x}\left(t\right)\right)$ по $t$ равна:

$$\frac{d}{dt}f\left(\mathbf{x}\left(t\right)\right)=-\left\langle \boldsymbol{\nabla} f\left(\mathbf{x}\left(t\right)\right),\boldsymbol{\nabla} f\left(\mathbf{x}_k\right)\right\rangle$$

- Подставим $t=t_k$:

$$\left.\frac{d}{dt}f\left(\mathbf{x}_k-t\boldsymbol{\nabla} f\left(\mathbf{x}_k\right)\right)\right|_{t=t_k}=-\left\langle \boldsymbol{\nabla} f\left(\mathbf{x}_{k+1}\right),\boldsymbol{\nabla} f\left(\mathbf{x}_k\right)\right\rangle$$

Таким образом, доказано, что производная функции $f\left(\mathbf{x}\left(t\right)\right)$ по $t$ в точке $t_k$ равна скалярному произведению градиентов функции $f$ в точках $\mathbf{x}_{k+1}$ и $\mathbf{x}_k$. ■

**Задание 9.** *Пусть* $f(\mathbf{x})=0.5\mathbf{x}^{\mathsf{T}}A\mathbf{x}+\langle\mathbf{b},\mathbf{x}\rangle$ *— квадратичная функция.*

*Доказать, что для метода наискорейшего спуска справедливо соотношение*

$$t_k=\frac{\|\boldsymbol{\nabla} f(x_k)\|}{\langle\boldsymbol{\nabla} f(x_k),A\boldsymbol{\nabla} f(x_k)\rangle}$$

*Доказательство.*

- Градиент функции $f(\mathbf{x})$ равен:

$$\boldsymbol{\nabla} f(\mathbf{x})=A\mathbf{x}+\mathbf{b}$$

- На каждом шаге метода наискорейшего спуска выбирается такой шаг $t_k$, чтобы минимизировать функцию $f(\mathbf{x}_{k+1})$:

$$\mathbf{x}_{k+1}=\mathbf{x}_k-t_k\boldsymbol{\nabla} f(\mathbf{x}_k)$$

  Найдем такое $t_k$, при котором функция $f(\mathbf{x}_{k+1})$ минимальна:

$$\arg\min_{t_k} f(\mathbf{x}_{k+1})=\arg\min_{t_k} f(\mathbf{x}_k-t_k\boldsymbol{\nabla} f(\mathbf{x}_k))$$

- Подставим $\mathbf{x}_{k+1}$ в функцию $f(\mathbf{x})$:

$$\begin{aligned}f(\mathbf{x}_{k+1})&=0.5(\mathbf{x}_k-t_k\boldsymbol{\nabla} f(\mathbf{x}_k))^{\mathsf{T}}A(\mathbf{x}_k-t_k\boldsymbol{\nabla} f(\mathbf{x}_k))+\mathbf{b}^{\mathsf{T}}(\mathbf{x}_k-t_k\boldsymbol{\nabla} f(\mathbf{x}_k))\\&=0.5\mathbf{x}_k^{\mathsf{T}}A\mathbf{x}_k-t_k\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\mathbf{x}_k+0.5t_k^2\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\boldsymbol{\nabla} f(\mathbf{x}_k)+\mathbf{b}^{\mathsf{T}}\mathbf{x}_k-t_k\mathbf{b}^{\mathsf{T}}\boldsymbol{\nabla} f(\mathbf{x}_k)\end{aligned}$$

- Найдем производную функции $f(x_{k+1})$ по $t_k$:

$$\frac{d}{dt_k}f(x_{k+1})=-\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\mathbf{x}_k+t_k\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\boldsymbol{\nabla} f(\mathbf{x}_k)-\mathbf{b}^{\mathsf{T}}\boldsymbol{\nabla} f(\mathbf{x}_k)$$

- Сгруппируем слагаемые:

$$\frac{d}{dt_k}f(x_{k+1})=t_k\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\boldsymbol{\nabla} f(\mathbf{x}_k)-\left(\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\mathbf{x}_k+\mathbf{b}^{\mathsf{T}}\boldsymbol{\nabla} f(\mathbf{x}_k)\right)$$

- Воспользуемся тем, что $\mathbf{b}^{\mathsf{T}}\boldsymbol{\nabla} f(\mathbf{x}_k)=\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}\mathbf{b}$:

$$\begin{aligned}\frac{d}{dt_k}f(x_{k+1})&=t_k\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\boldsymbol{\nabla} f(\mathbf{x}_k)-\left(\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\mathbf{x}_k+\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}\mathbf{b}\right)\\&=t_k\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\boldsymbol{\nabla} f(\mathbf{x}_k)-\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}\left(A\mathbf{x}_k+\mathbf{b}\right)\end{aligned}$$

- Подставим вместо $A\mathbf{x}_k+\mathbf{b}$ градиент функции $f(\mathbf{x}_k)$:

$$\frac{d}{dt_k}f(x_{k+1})=t_k\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}A\boldsymbol{\nabla} f(\mathbf{x}_k)-\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}}\boldsymbol{\nabla} f(\mathbf{x}_k)$$

- Воспользуемся условием оптимальности шага $t_k$:

$$\frac{d}{dt_k} f(x_{k+1}) = 0$$

Получаем:

$$t_k \boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}} A \boldsymbol{\nabla} f(\mathbf{x}_k) = \boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}} \boldsymbol{\nabla} f(\mathbf{x}_k)$$

- Откуда получаем формулу для шага $t_k$:

$$t_k = \frac{\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}} \boldsymbol{\nabla} f(\mathbf{x}_k)}{\boldsymbol{\nabla} f(\mathbf{x}_k)^{\mathsf{T}} A \boldsymbol{\nabla} f(\mathbf{x}_k)} = \frac{\|\boldsymbol{\nabla} f(\mathbf{x}_k)\|^2}{\langle \boldsymbol{\nabla} f(\mathbf{x}_k), A \boldsymbol{\nabla} f(\mathbf{x}_k) \rangle}$$

Итого, доказано, что для метода наискорейшего спуска справедливо соотношение

$$t_k = \frac{\|\boldsymbol{\nabla} f(x_k)\|}{\langle \boldsymbol{\nabla} f(x_k), A \boldsymbol{\nabla} f(x_k) \rangle}$$

■

**Задание 10.** *Найти минимум функции $f(x_1, x_2) = x_1 x_2$ при ограничениях $x_1^2 + x_2^2 \leqslant 1$, используя теорему Каруша-Джона.*

*Решение.*

- Запишем функцию Лагранжа:

$$L(x_1, x_2, \lambda) = x_1 x_2 + \lambda(x_1^2 + x_2^2 - 1)$$

- Найдем градиент функции Лагранжа:

$$\boldsymbol{\nabla} L = \begin{pmatrix} x_2 + 2\lambda x_1 \\ x_1 + 2\lambda x_2 \\ x_1^2 + x_2^2 - 1 \end{pmatrix}$$

- По условию оптимальности из теоремы Каруша-Джона градиент функции Лагранжа равен нулю:

$$\begin{cases} x_2 + 2\lambda x_1 = 0 \\ x_1 + 2\lambda x_2 = 0 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases}$$

- Решим систему уравнений:

$$\begin{cases} x_2 + 2\lambda x_1 = 0 \\ x_1 + 2\lambda x_2 = 0 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases}$$

Получаем:

$$\begin{cases} x_2 + 2\lambda x_1 = 0 \\ x_1 + 2\lambda x_2 = 0 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases} \Rightarrow \left[ \begin{array}{l} \begin{cases} \lambda = -\frac{1}{2} \\ x_1 = x_2 = \pm\frac{1}{\sqrt{2}} \end{cases} \\ \\ \begin{cases} \lambda = \frac{1}{2} \\ x_1 = -x_2 = \pm\frac{1}{\sqrt{2}} \end{cases} \end{array} \right.$$

- Подставим найденные значения $x_1$ и $x_2$ в функцию $f(x_1, x_2) = x_1 x_2$:

$$f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \frac{1}{2}$$

$$f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = -\frac{1}{2}$$

$$f\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = -\frac{1}{2}$$

$$f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = \frac{1}{2}$$

Таким образом, минимум функции $f(x_1, x_2) = x_1 x_2$ при ограничениях $x_1^2 + x_2^2 \leqslant 1$ равен $-\frac{1}{2}$ и достигается в точках $x_1 = -x_2 = \pm\frac{1}{\sqrt{2}}$. ■

**Задание 11.** *Пусть $f(y, y')$ не зависит от $x$.*
*Чему равно значение $\frac{d}{dx}(f - f_{y'}y')$?*

*Решение.* Введем обозначение $f = f(y, y')$, $f_y = \frac{d}{dy}f$, $f_{y'} = \frac{d}{dy'}f$.

- Продифференцируем $f - f_{y'}y'$ по $x$:

$$\frac{d}{dx}(f - f_{y'}y') = \frac{d}{dx}f - \frac{d}{dx}(f_{y'}y')$$

- Так как $f$ не зависит от $x$, то $\frac{d}{dx}f = 0$:

$$\frac{d}{dx}(f - f_{y'}y') = -\frac{d}{dx}(f_{y'}y')$$

- Применим правило дифференцирования произведения:

$$\frac{d}{dx}(f_{y'}y') = -f_{y'}\frac{d}{dx}(y') - \frac{d}{dx}(f_{y'})y'$$

- Так как $f$ не зависит от $x$, то $\frac{d}{dx}(f_{y'}) = 0$:

$$\frac{d}{dx}(f_{y'}y') = -f_{y'}\frac{d}{dx}(y') = -f_{y'}y''$$

Итого, значение $\frac{d}{dx}(f - f_{y'}y')$ равно $-f_{y'}y''$. ■

**Задание 12.** *Составить уравнение Эйлера-Лагранжа для $F = x^2y + y'^2x$.*

*Решение.* Формула Эйлера-Лагранжа имеет вид:

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0$$

- Найдем частную производную $F$ по $y'$:

$$\frac{\partial F}{\partial y'} = 2xy'$$

- Найдем производную $\frac{\partial F}{\partial y'}$ по $x$:

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 2y' + 2xy''$$

- Найдем частную производную $F$ по $y$:

$$\frac{\partial F}{\partial y} = x^2$$

- Подставим найденные значения в уравнение Эйлера-Лагранжа:

$$2y' + 2xy'' - x^2 = 0$$

Таким образом, уравнение Эйлера-Лагранжа для $F = x^2y + y'^2x$ имеет вид $2y' + 2xy'' - x^2 = 0$. ■