

## Определения и формулировки

### 1. Положительно определенная матрица.

💡 Матрица  $A \in \mathbb{S}^n$  называется положительно (отрицательно) определённой, если для всех  $x \in \mathbb{R}^n, x \neq 0$  выполнено  $x^\top A x > 0$  (соотв.  $< 0$ ). Обозначение:  $A \succ 0$  ( $A \prec 0$ ). Аналогично определяется полуопределённость, только там неравенства нестрогие.

### 2. Евклидова норма вектора.



$$\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$$

Данная норма соответствует расстоянию в реальном мире. Иначе называется 2-норма (см.  $p$ -норма вектора)

### 3. Неравенство треугольника для нормы.

💡 Норма должна удовлетворять следующим свойствам:

1.  $\|\alpha x\| = |\alpha| \|x\|, \alpha \in \mathbb{R}$
2.  $\|x\| = 0 \iff x = 0$
3.  $\|x + y\| \leq \|x\| + \|y\|$  – неравенство треугольника

### 4. $p$ -норма вектора.



$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

Важные частные случаи:

- Норма Чебышёва:  $\|x\|_\infty = \max_i |x_i|$
- Манхэттенское расстояние или  $\ell_1$ -норма:  $\|x\|_1 = \sum_{i=1}^n |x_i|$

### 5. Как выглядит единичный шар в $p$ -норме на плоскости для $p = 1, 2, \infty$ ?

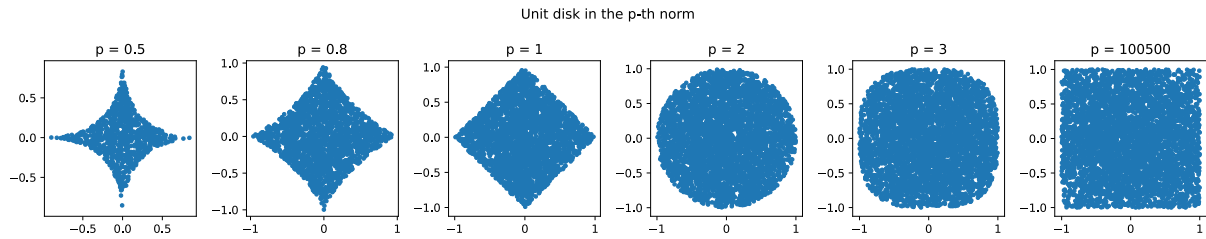


Рисунок 1: Шары в разных нормах

6. Норма Фробениуса для матрицы.



$$\|A\|_F = \left( \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}}$$

7. Спектральная норма матрицы.



$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_1(A) = \sqrt{\lambda_{\max}(A^T A)}$$

Где  $\sigma_1(A)$  – старшее сингулярное значение  $A$ ,  $\lambda_{\max}(A^T A)$  – наибольшее собственное значение  $A^T A$ .

8. Скалярное произведение двух векторов.



Пусть  $x, y \in \mathbb{R}^n$ , тогда их скалярное произведение это

$$\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i = y^T x = \langle y, x \rangle$$

9. Скалярное произведение двух матриц, согласованное с нормой Фробениуса.



Пусть  $X, Y \in \mathbb{R}^{m \times n}$ , тогда их скалярное произведение это

$$\langle X, Y \rangle = \text{tr}(X^T Y) = \sum_{i=1}^m \sum_{j=1}^n X_{ij} Y_{ij} = \text{tr}(Y^T X) = \langle Y, X \rangle$$

Связь с нормой Фробениуса:  $\langle X, X \rangle = \|X\|_F^2$

10. Собственные значения матрицы. Спектр матрицы.



Скаляр  $\lambda$  является собственным значением для матрицы  $A$ , если существует вектор  $q$ , такой что  $Aq = \lambda q$ . В таком случае ненулевой вектор  $q$  называют собственным вектором. Спектр матрицы – совокупность её собственных значений.

## 11. Связь спектра матрицы и ее определенности.

💡 Симметричная матрица положительно (неотрицательно) определена  $\iff$  её спектр (все её собственные значения) положителен (неотрицателен).

## 12. Спектральное разложение матрицы.

💡 Спектральное разложение матрицы, или разложение матрицы на основе собственных векторов, — это представление квадратной матрицы  $A$  в виде произведения трёх матриц  $A = S\Lambda S^{-1}$ , где  $S$  — матрица, столбцы которой являются собственными векторами матрицы  $A$ ,  $\Lambda$  — диагональная матрица с соответствующими собственными значениями на главной диагонали. В таком виде могут быть представлены только матрицы, обладающие полным набором собственных векторов.

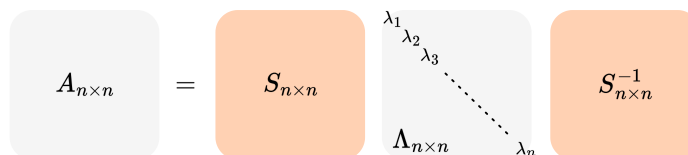


Рисунок 2: Спектральное разложение матрицы

## 13. Сингулярное разложение матрицы.

💡  $A \in \mathbb{R}^{m \times n}$ ,  $\text{rank}(A) = r$ .

$$A = U\Sigma V^T$$

$U \in \mathbb{R}^{m \times r}$ ,  $U^T U = I$ ,  $V \in \mathbb{R}^{n \times r}$ ,  $V^T V = I$ ,  $\Sigma$  диагональная матрица

$$\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r)$$

такая что

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$$

Столбцы  $U$ ,  $V$  — левые и правые сингулярные векторы  $A$ ,  $\sigma_i$  — сингулярные значения.

$$A = \sum_{i=1}^r \sigma_i u_i v_i^T$$

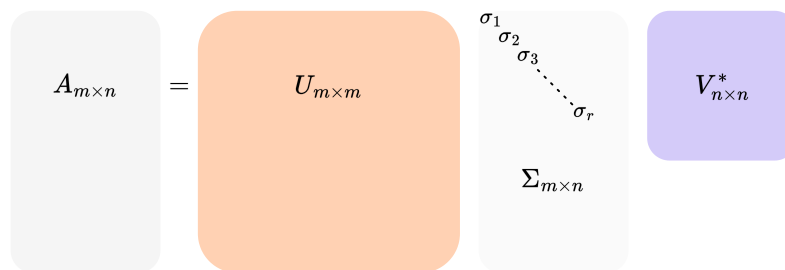


Рисунок 3: Сингулярное разложение матрицы

## 14. Связь определителя и собственных чисел для квадратной матрицы.

- 💡 Если у матрицы  $A$  собственные значения  $\lambda_1, \lambda_2, \dots, \lambda_n$ , то её определитель равен:

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n$$

15. Связь следа и собственных чисел для квадратной матрицы.

- 💡 Если у матрицы  $A$  собственные значения  $\lambda_1, \lambda_2, \dots, \lambda_n$ , то её след равен:

$$\text{tr}(A) = \lambda_1 + \lambda_2 + \dots + \lambda_n$$

16. Градиент функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ .

- 💡  $\nabla f(x)$ , вектор частных производных функции  $f$ .

17. Гессиан функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ .



$$f''(x) = \nabla^2 f(x) = \left[ \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right]_{i,j=1}^n = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

18. Якобиан функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .



$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

19. Формула для аппроксимации Тейлора первого порядка  $f_{x_0}^I(x)$  функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$  в точке  $x_0$ .

- 💡 Для дифференцируемой  $f$ :

$$f_{x_0}^I(x) = f(x_0) + \nabla f(x_0)^T (x - x_0)$$

20. Формула для аппроксимации Тейлора второго порядка  $f_{x_0}^{II}(x)$  функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$  в точке  $x_0$ .

- 💡 Для дважды дифференцируемой  $f$ :

$$f_{x_0}^{II}(x) = f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{1}{2} (x - x_0)^T \nabla^2 f(x_0) (x - x_0)$$

21. Определение дифференцируемости функции в точке через производную как линейный оператор.

- 💡 Пусть  $x \in S$  - внутренняя точка множества  $S$ , и пусть  $D : U \rightarrow V$  - линейный оператор. Функция  $f$  называется дифференцируемой в точке  $x$  с производной  $D$  если для всех достаточно малых  $h \in U$  верно следующее:

$$f(x+h) = f(x) + D[h] + o(\|h\|)$$

Если для любого линейного оператора  $D : U \rightarrow V$  функция  $f$  не является дифференцируемой в точке  $x$  с производной  $D$ , тогда мы говорим, что  $f$  не дифференцируема в точке  $x$ .

22. Связь дифференциала функции  $df$  и градиента  $\nabla f$  для функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ .



$$df(x) = \langle \nabla f(x), dx \rangle$$

23. Связь второго дифференциала функции  $d^2 f$  и гессиана  $\nabla^2 f$  для функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ .



$$d(df) = d^2 f(x) = \langle \nabla^2 f(x) dx_1, dx \rangle = \langle H_f(x) dx_1, dx \rangle$$

24. Формула для приближенного вычисления производной функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$  по  $k$ -ой координате с помощью метода конечных разностей.



$$\frac{\partial f}{\partial x_k}(x) \approx \frac{f(x + \varepsilon e_k) - f(x)}{\varepsilon}, \quad e_k = (0, \dots, \frac{1}{k}, \dots, 0)$$

Время работы:  $(d+1)T$ , где вызов  $f(x)$  занимает  $T$ ,  $x \in \mathbb{R}^d$

25. Пусть  $f = f(x_1(t), \dots, x_n(t))$ . Формула для вычисления  $\frac{\partial f}{\partial t}$  через  $\frac{\partial x_i}{\partial t}$  (Forward chain rule).



$$\frac{\partial f}{\partial t} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial t}$$

26. Пусть  $L$  - функция, возвращающая скаляр, а  $v_k$  - функция, возвращающая вектор  $x \in \mathbb{R}^t$ . Формула для вычисления  $\frac{\partial L}{\partial v_k}$  через  $\frac{\partial L}{\partial x_i}$  (Backward chain rule).



$$\frac{\partial L}{\partial v_k} = \sum_{i=1}^t \frac{\partial L}{\partial x_i} \frac{\partial x_i}{\partial v_k}$$

27. Аффинное множество. Аффинная комбинация. Аффинная оболочка.

- 💡 Множество  $A$  называется аффинным если для любых  $x_1, x_2$  из  $A$  прямая, проходящая через  $x_1, x_2$ , тоже лежит в  $A$ . То есть:

$$\forall \theta \in \mathbb{R}, \forall x_1, x_2 \in A : \theta x_1 + (1 - \theta) x_2 \in A$$

Пример аффинного множества:  $\mathbb{R}^n$

Пусть  $x_1, x_2, \dots, x_k \in S$ . Тогда точка  $\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k$  называется аффинной комбинацией, если

$$\forall i \in \{1, \dots, k\} : \theta_i \in \mathbb{R}, \quad \sum_{i=1}^k \theta_i = 1$$

Аффинная оболочка – множество всех возможных аффинных комбинаций элементов множества.

$$\text{aff}(S) = \left\{ \sum_{i=1}^k \theta_i x_i \mid k > 0, x_i \in S, \theta_i \in \mathbb{R}, \sum_{i=1}^k \theta_i = 1 \right\}$$

28. Выпуклое множество. Выпуклая комбинация. Выпуклая оболочка.

- 💡 Множество  $S$  называется выпуклым если для любых  $x_1, x_2$  из  $S$  отрезок между  $x_1, x_2$  тоже лежит в  $S$ . То есть:

$$\forall \theta \in [0, 1], \forall x_1, x_2 \in S : \theta x_1 + (1 - \theta) x_2 \in S$$

Пусть  $x_1, x_2, \dots, x_k \in S$ . Тогда точка  $\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k$  называется выпуклой комбинацией, если

$$\forall i \in \{1, \dots, k\} : \theta_i \geq 0, \quad \sum_{i=1}^k \theta_i = 1$$

Выпуклая оболочка – множество всех возможных выпуклых комбинаций элементов множества.

$$\text{conv}(S) = \left\{ \sum_{i=1}^k \theta_i x_i \mid k > 0, x_i \in S, \theta_i \geq 0, \sum_{i=1}^k \theta_i = 1 \right\}$$

29. Конус. Выпуклый конус. Коническая комбинация. Коническая оболочка.

- 💡 Множество  $S$  называется конусом если для любого  $x$  из  $S$  луч, проходящий из 0 через  $x$ , тоже лежит в  $S$ . То есть:

$$\forall \theta \geq 0, \forall x \in S : \theta x \in S$$

Множество  $S$  называется выпуклым конусом если для любых  $x_1, x_2 \in S$  их коническая комбинация тоже лежит в  $S$ . То есть:

$$\forall x_1, x_2 \in S, \theta_1, \theta_2 \geq 0 : \theta_1 x_1 + \theta_2 x_2 \in S$$

Пусть  $x_1, x_2, \dots, x_k \in S$ . Тогда точка  $\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k$  называется конической комбинацией, если

$$\forall i \in \{1, \dots, k\} : \theta_i \geq 0$$

Коническая оболочка - множество всех возможных конических комбинаций элементов множества.

$$\text{coni}(S) = \left\{ \sum_{i=1}^k \theta_i x_i \mid k > 0, x_i \in S, \theta_i \geq 0 \right\}$$

### 30. Внутренность множества.

- 💡 Внутренность множества - совокупность всех точек множества, содержащих вместе с собой в множестве некоторую окрестность вокруг себя.

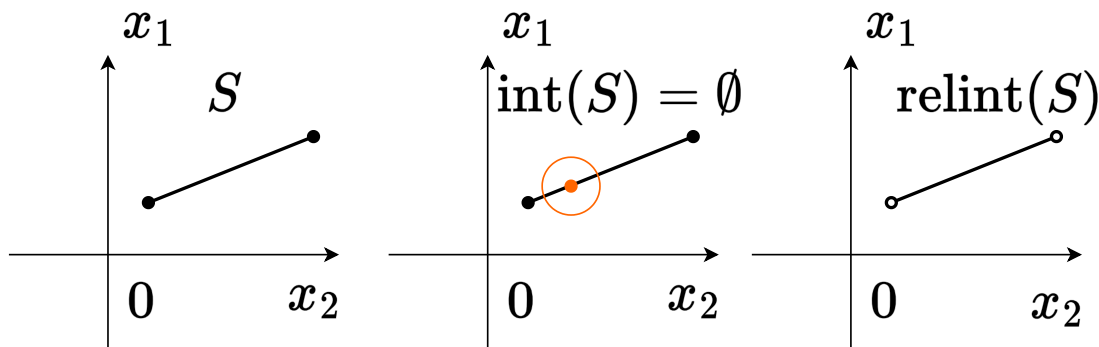
### 31. Относительная внутренность множества.

- 💡 Относительная внутренность множества - внутренность множества в его аффинной оболочке. Может быть полезной при работе с множествами меньшей размерности чем пространство, в котором они находятся.

$$\text{relint}(S) = \{x \in S \mid \exists \varepsilon > 0, N_\varepsilon(x) \cap \text{aff}(S) \subseteq S\}$$

$N_\varepsilon(x)$  – шар радиуса  $\varepsilon$  с центром в  $x$ ,  $\text{aff}(S)$  – аффинная оболочка  $S$

Пример: отрезок на плоскости имеет пустую внутренность, но его относительная внутренность – тот же отрезок без концов.



### 32. Сумма Минковского.

💡 Сумма Минковского двух множеств  $S_1, S_2$  — это множество

$$S_1 + S_2 = \{s_1 + s_2 \mid s_1 \in S_1, s_2 \in S_2\}.$$

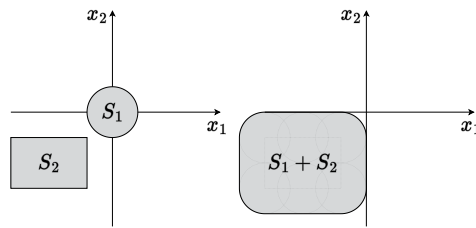


Рисунок 4: Сумма Минковского двух множеств

33. Любые 2 операции с множествами, сохраняющие выпуклость.



1. Линейная комбинация:

$$S = \{s \mid s = c_1x + c_2y, x \in S_x, y \in S_y, c_1, c_2 \in \mathbb{R}\}$$

2. Пересечение любого числа выпуклых множеств

3. Образ множества в аффинном преобразовании:

$$S \subseteq \mathbb{R}^n \text{ выпукло} \rightarrow f(S) = \{f(x) \mid x \in S\} \text{ выпукло} \quad (f(x) = Ax + b)$$

34. Выпуклая функция.



Функция  $f(x)$ , определённая на выпуклом множестве  $S \subseteq \mathbb{R}^n$  называется выпуклой на  $S$  если:

$$\forall x_1, x_2 \in S, \quad \forall \lambda \in [0, 1]$$

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

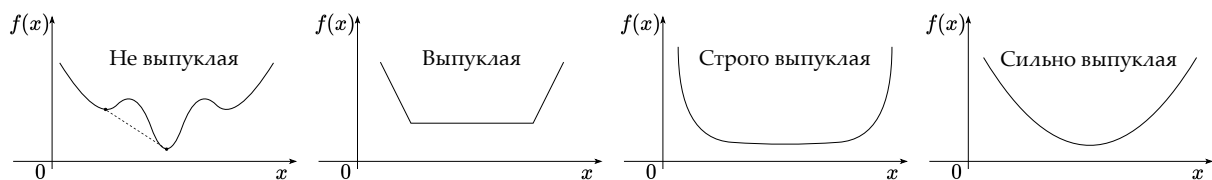


Рисунок 5: Иллюстрация выпуклой функции

35. Строго выпуклая функция.



💡 Функция  $f(x)$ , определённая на выпуклом множестве  $S \subseteq \mathbb{R}^n$  называется строго выпуклой на  $S$  если:

$$\forall x_1, x_2 \in S : x_1 \neq x_2, \quad \forall \lambda \in (0, 1) \\ f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2)$$

36. Надграфик функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ .

💡 Для функции, определённой на  $S \subseteq \mathbb{R}^n$ , множество:

$$\text{epi } f = \{[x, \mu] \in S \times \mathbb{R} : f(x) \leq \mu\}$$

называется надграфиком функции  $f(x)$ .

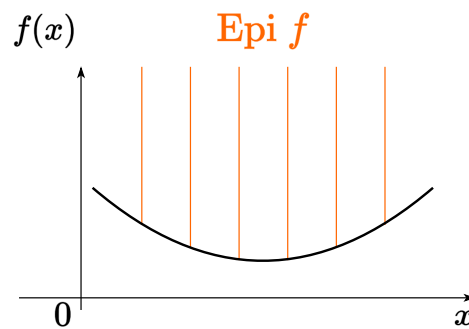


Рисунок 6: Иллюстрация надграфика функции

37. Множество подуровней функции  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ .

💡 Для функции, определённой на  $S \subseteq \mathbb{R}^n$ , множество:

$$\mathcal{L}_\beta = \{x \in S : f(x) \leq \beta\}$$

называется множеством подуровней или множеством Лебега функции  $f(x)$

Если функция выпукла, то множество её подуровней выпукло. Обратное - не верно ( $f(x) = \sqrt{|x|}$ ).

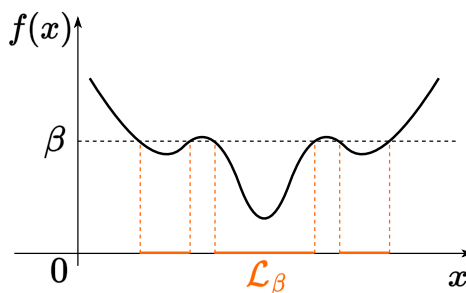


Рисунок 7: Множество подуровней функции для выбранного уровня  $\beta$

38. Дифференциальный критерий выпуклости первого порядка.

💡 Дифференцируемая функция определенная на выпуклом множестве  $S \subseteq \mathbb{R}^n$  выпукла тогда и только тогда когда  $\forall x, y \in S$ :

$$f(y) \geq f(x) + \nabla f(x)^T(y - x)$$

39. Дифференциальный критерий выпуклости второго порядка.

💡 Дважды дифференцируемая функция определенная на выпуклом множестве  $S \subseteq \mathbb{R}^n$  выпукла тогда и только тогда когда для любой внутренней точки  $x \forall x \in \text{int}(S) \neq \emptyset$ :

$$\nabla^2 f(x) \succeq 0$$

40. Связь выпуклости функции и ее надграфика.

💡 Функция выпукла тогда и только тогда, когда её надграфик - выпуклое множество.

41.  $\mu$ -сильно выпуклая функция.

💡 Функция определенная на выпуклом множестве  $S \subseteq \mathbb{R}^n$  называется сильно выпуклой если  $\forall x_1, x_2 \in S, 0 \leq \lambda \leq 1$  и  $\mu > 0$ :

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) - \frac{\mu}{2} \lambda(1 - \lambda) \|x_1 - x_2\|^2$$

42. Дифференциальный критерий сильной выпуклости первого порядка.

- 💡 Дифференцируемая функция определенная на выпуклом множестве  $S \subseteq \mathbb{R}^n$  является сильно выпуклой тогда и только тогда, когда  $\forall x, y \in S$ :

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) + \frac{\mu}{2}\|y - x\|^2$$

43. Дифференциальный критерий сильной выпуклости второго порядка.

- 💡 Дважды дифференцируемая функция определенная на выпуклом множестве  $S \subseteq \mathbb{R}^n$  является сильно выпуклой тогда и только тогда, когда существует  $\mu > 0$

$$\nabla^2 f(x) \succeq \mu I$$

44. Любые 2 операции с функциями, сохраняющие выпуклость.



1. Сумма выпуклых функций с не отрицательными коэффициентами является выпуклой функцией.
2. Композиция выпуклой функции с аффинной выпукла:  $g(x) = f(Ax + b)$
3. Поточечный максимум любого числа выпуклых функций есть выпуклая функция.

45. Является ли задача линейных наименьших квадратов для переопределенной линейной системы выпуклой/сильно выпуклой?



Рассмотрим задачу минимизации функции:

$$\|Ax - b\|^2 \rightarrow \min_{x \in \mathbb{R}^d},$$

где матрица  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $m > n$  (стоячая). Легко заметить, что гессиан минимизируемой функции  $A^T A$  - невырожденная матрица размера  $n \times n$ . Она является положительно определенной. То есть задача в классической постановке является сильно выпуклой. Т.е. содержит единственный локальный минимум (единственное решение). Отметим, что мы предполагаем, что матрица имеет полный ранг.

46. Является ли задача линейных наименьших квадратов для недоопределенной линейной системы выпуклой/сильно выпуклой?

💡 Рассмотрим задачу минимизации функции:

$$\|Ax - b\|^2 \rightarrow \min_{x \in \mathbb{R}^d},$$

где матрица  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $m < n$  (лежачая). Легко заметить, что гессиан минимизируемой функции  $2A^T A$  - вырожденная матрица размера  $n \times n$ . Однако, она является положительно полуопределенной. То есть задача в классической постановке является выпуклой, но не сильно выпуклой. Т.е. содержит бесконечное количество локальных минимумов, каждый из которых - глобальный. Стоит отметить, что добавление  $\ell_2$  регуляризации к минимизируемой функции изменит задачу, однако, будет гарантировать сильную выпуклость. Отметим, что мы предполагаем, что матрица имеет полный ранг.

47. Сопряженное множество.

💡 Пусть  $S \subseteq \mathbb{R}^n$  - произвольное непустое множество. Тогда его сопряженное множество определяется как:

$$S^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in S\}$$

48. Любые 2 нетривиальных свойства сопряженного множества.



1. Сопряженное множество всегда замкнуто, выпукло и содержит ноль.
2. Для произвольного множества  $S \subseteq \mathbb{R}^n$ :  $S^{**} = \text{conv}(S \cup \{0\})$
3. Если  $S_1 \subseteq S_2$ , то  $S_2^* \subseteq S_1^*$ .
4.  $\left(\bigcup_{i=1}^m S_i\right)^* = \bigcap_{i=1}^m S_i^*$ .
5. Если  $S$  замкнуто, выпукло и включает 0, то  $S^{**} = S$ .
6.  $S^* = (\overline{S})^*$ .

49. Сопряженный конус.



Сопряженным конусом к конусу  $K$  называется множество  $K^*$  такое, что:

$$K^* = \{y \mid \langle x, y \rangle \geq 0 \quad \forall x \in K\}$$

50. Сопряженная функция.

- 💡 Для функции  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , сопряженной к ней называется функция  $f^*$ , причем область её определения можно считать теми  $y$ , для которых  $\max$  конечен.

$$f^*(y) = \max_x [y^T x - f(x)]$$

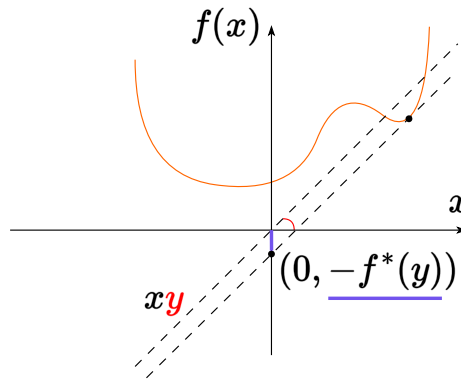


Рисунок 8: Сопряженная функция

51. Связь сильной выпуклости функции и гладкости сопряженной функции.

- 💡 Пусть  $f$  - замкнутая и выпуклая. Тогда  $f$  - сильно выпуклая с константой выпуклости  $\mu \Leftrightarrow \nabla f^*$  - липшицев с параметром  $\frac{1}{\mu}$ .

52. Сопряженная норма. Сопряженная норма к векторной  $p$ -норме.

- 💡 Сопряжённой нормой  $\|\cdot\|_*$  к норме  $\|\cdot\|$  называется норма, определённая как:

$$\|y\|_* = \sup_{\|x\| \leq 1} \langle y, x \rangle,$$

где  $x, y \in \mathbb{R}^n$ .

Для  $p$ -нормы сопряжённой является  $q$ -норма, где  $p$  и  $q$  связаны соотношением:

$$\frac{1}{p} + \frac{1}{q} = 1, \quad p, q \geq 1.$$

Например:

- Для  $p = 1$  сопряжённой является  $q = \infty$ .
- Для  $p = 2$  сопряжённая норма также является 2-нормой.
- Для  $p = \infty$  сопряжённой является  $q = 1$ .

53. Субградиент. Субдифференциал.

💡 Субградиент функции  $f$  в точке  $x$  — это вектор  $g$ , удовлетворяющий условию:

$$f(y) \geq f(x) + g^T(y - x), \quad \forall y.$$

Множество всех субградиентов в точке  $x$  называется субдифференциалом и обозначается как  $\partial f(x)$ .

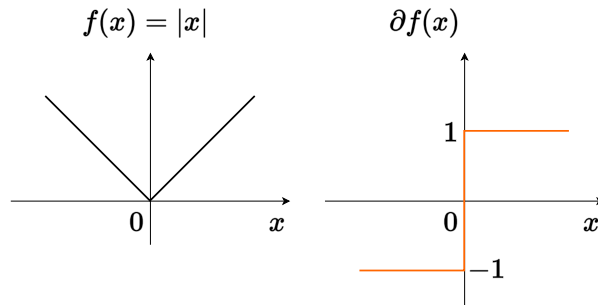


Рисунок 9: Субдифференциал функции ReLU.

#### 54. Нормальный конус.

💡 Для  $x \in S$ ,  $\partial I_S(x) = \mathcal{N}_S(x)$ , **нормальный конус** к  $S$  в точке  $x$ , напомним.

$$\mathcal{N}_S(x) = \{g \in \mathbb{R}^n : g^T x \geq g^T y \text{ для любого } y \in S\}$$

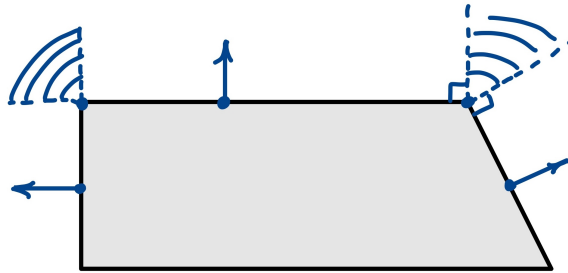


Рисунок 10: Нормальные конусы к множеству  $S$  в точках

#### 55. Теорема Моро - Рокафеллара.

💡 Пусть  $f_i(x)$  — выпуклые функции, определённые на выпуклых множествах  $S_i$ ,  $i = \overline{1, n}$ . Если выполнено условие  $\bigcap_{i=1}^n S_i \neq \emptyset$ , то функция  $f(x) = \sum_{i=1}^n a_i f_i(x)$ ,  $a_i > 0$  имеет субдифференциал  $\partial_S f(x)$  на множестве  $S = \bigcap_{i=1}^n S_i$ , и его можно выразить следующим образом:

$$\partial_S f(x) = \sum_{i=1}^n a_i \partial_{S_i} f_i(x).$$

Это означает, что субдифференциал линейной комбинации выпуклых функций равен взвешенной сумме их субдифференциалов, взятых на пересечении соответствующих множеств.

56. Теорема Дубовицкого - Милютина.

💡 Пусть  $f_i(x)$  — выпуклые функции, определённые на открытом выпуклом множестве  $S \subseteq \mathbb{R}^n$ , и  $x_0 \in S$ . Пусть  $f(x)$  определяется как покомпонентный максимум этих функций:

$$f(x) = \max_i f_i(x).$$

Тогда субдифференциал  $f(x_0)$  выражается следующим образом:

$$\partial_S f(x_0) = \text{conv} \left\{ \bigcup_{i \in I(x_0)} \partial_S f_i(x_0) \right\},$$

где множество  $I(x)$  определяется как индексы функций, достигающих максимума:

$$I(x) = \{i \in [1 : m] : f_i(x) = f(x)\}.$$

Это утверждение говорит, что субдифференциал максимума выпуклых функций представляет собой выпуклую оболочку объединения субдифференциалов тех функций, которые достигают максимума в данной точке.

57. Теорема Вейерштрасса.

💡 Пусть  $S \subset \mathbb{R}^n$  - компакт, а  $f(x)$  - непрерывная функция на  $S$ . Значит, точка глобального минимума функции  $f(x)$  на  $S$  существует.

58. Теорема Тейлора.

- 💡  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  - непрерывная, дифференцируемая функция и  $p \in \mathbb{R}^n$ , тогда теорема Тейлора гласит:

$$f(x+p) = f(x) + \nabla f(x)^T p$$

Для некоторого  $t \in (0, 1)$

Более того, если  $f$  - дважды дифференцируема, то:

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x+tp) p$$

Для некоторого  $t \in (0, 1)$

59. Необходимые условия локального экстремума.

- 💡 Если  $x^*$  - локальный экстремум и  $f$  непрерывно дифференцируема в открытой окрестности  $x^*$ , то:

$$\nabla f(x^*) = 0$$

60. Достаточные условия локального экстремума.

- 💡 Если  $\nabla^2 f$  непрерывна в открытой окрестности  $x^*$  и

$$\nabla f(x^*) = 0$$

$$\nabla^2 f(x^*) \succ 0$$

То  $x^*$  - локальный минимум  $f(x)$ . Для локального максимума аналогично, только

$$0 \succ \nabla^2 f(x^*)$$

61. Принцип Ферма для минимума функции.

- 💡 Пусть  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ , тогда  $x^*$  является глобальным минимумом  $f$  тогда и только тогда, когда

$$0 \in \partial f(x^*)$$

62. Общая задача математического программирования. Функция Лагранжа.



$$\begin{cases} f_0(x) \rightarrow \min_{x \in \mathbb{R}^d} \\ f_i(x) \leq 0, & i = 1, \dots, m, \\ h_i(x) = 0, & i = 1, \dots, p. \end{cases}$$

Функция Лагранжа:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x),$$

63. Теорема Каруша - Куна - Таккера в форме необходимых условий решения задачи математического программирования.



💡 Пусть  $x_*$  - решение задачи с нулевым зазором двойственности

$$\begin{cases} f_0(x) \rightarrow \min_{x \in \mathbb{R}^d} \\ f_i(x) \leq 0, \quad i = 1, \dots, m, \\ h_i(x) = 0, \quad i = 1, \dots, p. \end{cases}$$

Функция Лагранжа:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x),$$

Тогда найдутся такие векторы  $\lambda^*$  и  $\nu^*$ , что выполнены условия

$$\begin{cases} \nabla f_0(x_*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x_*) + \sum_{i=1}^p \nu_i^* \nabla h_i(x_*) = 0 \\ f_i(x_*) \leq 0, \quad i = 1, \dots, m \\ h_i(x_*) = 0, \quad i = 1, \dots, p \\ \lambda_i^* \geq 0, \quad i = 1, \dots, m \\ \lambda_i^* f_i(x_*) = 0, \quad i = 1, \dots, m \end{cases}$$

64. Условие Слейтера.



1. Если задача выпуклая (т.е., говоря о задаче минимизации, оптимизируемая функция  $f_0$  и ограничения вида неравенство  $f_i$  – выпуклые, ограничения вида равенства  $h_i$  – аффинные)
2. И существует точка  $x$  такая, что  $h_i(x) = 0$  и  $f_i(x) < 0$  (ограничения вида равенства активные, а ограничения вида неравенства выполняются строго)

То тогда задача имеет нулевой зазор двойственности и условия ККТ становятся необходимыми и достаточными.

65. Задача выпуклого программирования.



Задача выпуклого программирования — это задача оптимизации, в которой целевая функция является выпуклой функцией и область допустимых решений выпукла. В форме ниже функции  $f_0, \dots, f_m$  - выпуклые, а функции  $h_i$  - аффинные.

$$\begin{cases} f_0(x) \rightarrow \min_{x \in \mathbb{R}^d} \\ f_i(x) \leq 0, \quad i = 1, \dots, m, \\ h_i(x) = 0, \quad i = 1, \dots, p. \end{cases}$$

66. Двойственная функция в задаче математического программирования.

💡 Предположим, что  $D = \bigcap_{i=0}^m \text{dom } f_i \cap \bigcap_{i=0}^p \text{dom } h_i$  непустое. Определим двойственную функцию  $g : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$  как минимум лагранжиана по  $x$  : для  $\lambda \in \mathbb{R}^m, \nu \in \mathbb{R}^p$

$$g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu) = \inf_{x \in D} f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$

Так как двойственная функция это поточечный инфимум семейства аффинных функций от  $(\lambda, \nu)$ , она вогнутая, даже если изначальная задача не выпуклая.

67. Двойственная задача для задачи математического программирования.

💡 Пусть  $p^*$  - оптимальное значение исходной задачи. Пусть  $\hat{x}$  достижимая точка для исходной задачи, т.е.  $f_i(\hat{x}) \leq 0$  и  $h_i(\hat{x}) = 0, \lambda \geq 0$ . Тогда имеем:

$$L(\hat{x}, \lambda, \nu) = f_0(\hat{x}) + \underbrace{\lambda^T f(\hat{x})}_{\leq 0} + \underbrace{\nu^T h(\hat{x})}_{=0} \leq f_0(\hat{x})$$

Тогда

$$g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu) \leq L(\hat{x}, \lambda, \nu) \leq f_0(\hat{x})$$

$$g(\lambda, \nu) \leq p^*$$

Двойственной задачей называется

$$g(\lambda, \nu) \rightarrow \max_{\lambda \in \mathbb{R}^m, \nu \in \mathbb{R}^p}$$

$$s.t. \lambda \geq 0$$

68. Сильная двойственность. Зазор двойственности.

💡 Пусть  $p^*$  - оптимальное значение прямой задачи,  $d^*$  - оптимальное значение двойственной задачи. Зазором двойственности называется

$$p^* - d^* \geq 0$$

Сильная двойственность возникает, если зазор равен нулю

$$p^* = d^*$$

69. Локальный анализ чувствительности с помощью множителей Лагранжа.

💡 Перейдем к возмущенной версии задачи:

$$f_0(x) \rightarrow \min_x$$

$$f_i(x) \leq u_i, \quad i = 1, \dots, m$$

$$h_i(x) = v_i, \quad i = 1, \dots, p,$$

Обозначим  $p^*(u, v)$  - оптимальное значение этой задачи. Если имеет место сильная двойственность, то выполнено:

$$p^*(u, v) \geq p^*(0, 0) - (\lambda^*)^T u - (\nu^*)^T v$$

Если множители Лагранжа  $\lambda_i^*, \nu_i^*$  большие, то небольшое изменение ограничений приведет к существенному изменению оптимального решения. То есть соответствующие ограничения очень сильно влияют на задачу.

Если множители Лагранжа маленькие, то соответствующие ограничения мало влияют на задачу.

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i}$$

#### 70. Задача линейного программирования.

💡 **Задача линейного программирования (ЛП)** — это задача оптимизации с **линейной** целевой функцией и **линейными** ограничениями.

Общий вид (один из вариантов записи):

$$\min_{x \in \mathbb{R}^n} c^T x \quad \text{при} \quad Ax \leq b, \quad A_{eq}x = b_{eq}.$$

Здесь  $c \in \mathbb{R}^n$ , матрицы  $A, A_{eq}$  и векторы  $b, b_{eq}$  — заданные.

#### 71. Стандартная форма задачи линейного программирования.

💡 **Стандартная форма ЛП** (классическая для симплекс-метода):

$$\min_{x \in \mathbb{R}^n} c^T x \quad \text{при} \quad Ax = b, \quad x \geq 0.$$

(Иногда используют эквивалентную форму максимизации; различие лишь в знаках.)

Любая ЛП с неравенствами и/или свободными переменными сводится к стандартной форме введением добавочных переменных и разбиением свободных переменных на разность двух неотрицательных.

#### 72. Возможные случаи двойственности в задаче линейного программирования.

💡 Двойственная задача:

$$\begin{aligned} \max_{\nu \in \mathbb{R}^m} \quad & -b^T \nu \\ \text{s.t.} \quad & -A^T \nu \leq c \end{aligned}$$

1. Если либо у прямой, либо у двойственной задачи есть конечное решение, то и у другой тоже, и целевые переменные равны.
2. Если либо прямая, либо двойственная задача неограничена, то вторая из них невыполнима.

73. Идея симплекс метода.

💡 Симплекс метод решает следующую задачу:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \end{aligned}$$

Шаги выполнения симплекс метода:

1. Поиск начальной базисной допустимой точки
2. Проверка оптимальности. Если решение оптимально, то алгоритм завершается, иначе переходим к следующему шагу
3. Замена базиса
4. Повторяем предыдущие два шага до достижения оптимального решения или установления, что задача не имеет допустимого решения

74. Нахождение первоначальной угловой точки (идея).

💡 Если очевидного допустимого базиса нет, применить **двухфазный метод**:

- *Фаза 1*: добавить искусственные переменные и решить вспомогательную задачу, минимизируя сумму искусственных переменных, чтобы найти допустимую базисную точку (если минимум  $> 0$ , исходная задача недопустима). Для этой задачи можно легко указать начальную базисную допустимую точку.
- *Фаза 2*: стартовать симплекс-метод для исходной целевой функции из найденного допустимого базиса.

Это и даёт первоначальную угловую точку для запуска симплекс-метода.

75. Сходимость симплекс метода.

💡 В худшем случае симплекс метод сходится экспоненциально от размерности задачи, но на практике в среднем алгоритм работает сильно лучше. Задача, на которой симплекс метод работает экспоненциальное время, называется примером Klee Minty.

76. Теорема о связи задач max-flow и min-cut (надо суметь описать обе задачи).

💡 **Задача Max-Flow (максимальный поток):** Дано ориентированное взвешенное графовое представление сети, где узлы — это вершины графа, а ребра имеют пропускную способность (capacity). Требуется найти максимальный поток из источника (source) в сток (sink), при условии:

1. Поток на каждом ребре не превышает его пропускную способность.
2. Сохраняется закон сохранения потока в промежуточных узлах (входящий поток равен исходящему, за исключением источника и стока).

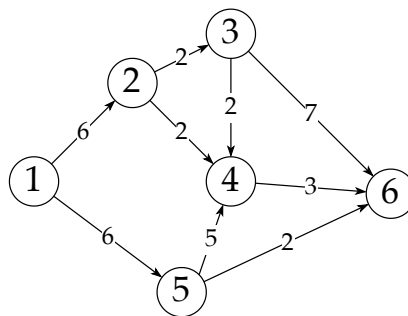
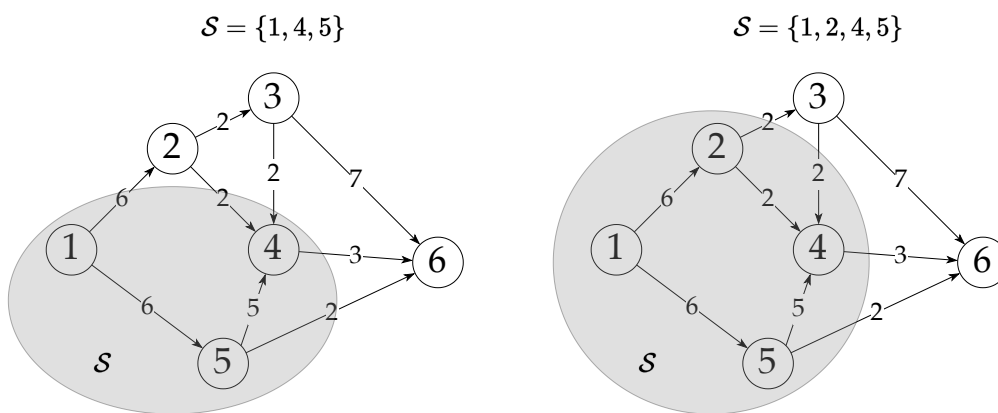


Рисунок 11: Транспортный граф

**Задача Min-Cut (минимальный разрез):** Для той же сети требуется найти разрез — разделение вершин графа на два множества (одно включает источник, другое — сток), при котором суммарная пропускная способность ребер, пересекающих разрез, минимальна.



**Теорема Max-Flow Min-Cut:** Максимальный поток из источника в сток равен минимальной пропускной способности разреза между источником и стоком.  
Формально:

$$\text{MAXFLOW} = \text{MINCUT}.$$

Эта теорема утверждает, что задачи поиска максимального потока и минимального разреза являются двойственными: решение одной задачи предоставляет решение другой.

- 💡 Последовательность  $\{r_k\}$  называется линейно сходящейся, если существуют константы  $C > 0$  и  $q \in (0, 1)$ , такие что

$$r_k \leq Cq^k, \quad k \geq m.$$

Наименьшее возможное  $q$  называется константой (скоростью) линейной сходимости.

78. Сублинейная сходимость последовательности.

- 💡 Если последовательность  $r_k$  сходится к нулю, но не обладает линейной сходимостью, то говорят, что она сходится сублинейно. Иногда мы можем рассматривать следующий класс сублинейной сходимости:

$$r_k \leq Ck^q,$$

где  $q < 0$  и  $0 < C < \infty$ .

79. Сверхлинейная сходимость последовательности.

- 💡 Мы определяем сверхлинейную сходимость как сходимость последовательности, которая быстрее любой линейной сходимости. Иногда рассматривают более специальный класс. Тогда говорят, что последовательность  $r_k > 0$  имеет сверхлинейную сходимость порядка  $p$ , если существуют  $C > 0$  и  $p > 1$  такие, что выполнено для всех достаточно больших  $k$ :

$$r_{k+1} \leq Cr_k^p$$

80. Квадратичная сходимость последовательности.

- 💡 Говорят, что последовательность  $r_k > 0$  сходится квадратично, если существует  $C > 0$  такое, что выполнено для всех достаточно больших  $k$ :

$$r_{k+1} \leq Cr_k^2$$

81. Тест корней для определения скорости сходимости последовательности.

- 💡 Пусть  $(r_k)_{k=m}^\infty$  - последовательность неотрицательных чисел, сходящаяся к нулю, и пусть  $\alpha := \limsup_{k \rightarrow \infty} r_k^{1/k}$ . (Заметим, что  $\alpha \geq 0$ .)

1. Если  $0 \leq \alpha < 1$ , то  $(r_k)_{k=m}^\infty$  сходится линейно с константой  $\alpha$ .
2. В частности, если  $\alpha = 0$ , то  $(r_k)_{k=m}^\infty$  сходится сверхлинейно.
3. Если  $\alpha = 1$ , то  $(r_k)_{k=m}^\infty$  сходится сублинейно.
4. Случай  $\alpha > 1$  невозможен.

82. Тест отношений для определения скорости сходимости последовательности.

💡 Пусть  $r_{k=m}^{\infty}$  - последовательность строго положительных чисел, сходящаяся к нулю. Пусть

$$q = \lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k}$$

1. Если существует  $q$  и  $0 \leq q < 1$ , то  $r_{k=m}^{\infty}$  имеет линейную сходимость с константой  $q$ .
2. В частности, если  $q = 0$ , то  $r_{k=m}^{\infty}$  имеет сверхлинейную сходимость.
3. Если  $q$  не существует, но  $q = \lim_{k \rightarrow \infty} \sup_k \frac{r_{k+1}}{r_k} < 1$ , то  $r_{k=m}^{\infty}$  имеет линейную сходимость с константой, не превышающей  $q$ .
4. Если  $\lim_{k \rightarrow \infty} \inf_k \frac{r_{k+1}}{r_k} = 1$ , то  $r_{k=m}^{\infty}$  имеет сублинейную сходимость.
5. Случай  $\lim_{k \rightarrow \infty} \inf_k \frac{r_{k+1}}{r_k} > 1$  невозможен.

### 83. Унимодальная функция.

💡 Функция  $f(x)$  называется унимодальной на  $[a, b]$ , если существует  $x^* \in [a, b]$ , такое, что

1.  $f(x_1) > f(x_2)$  для всех  $a \leq x_1 < x_2 < x^*$
2.  $f(x_1) < f(x_2)$  для всех  $x^* < x_1 < x_2 \leq b$

### 84. Метод дихотомии.

💡 Решаемая задача:  $\min_{x \in [a, b]} f(x)$ . В начале берётся середина отрезка и одна из четвертей (например, левая). Тогда по результатам измерения значений функции в этих двух точках решение может оказаться либо в левой половине отрезка, либо в правых трёх четвертях. Во втором случае нам необходимо провести дополнительное измерение значения функции в правой четверти исходного отрезка, тогда мы гарантируем уменьшение области поиска вдвое. Таким образом, в методе дихотомии мы гарантируем уменьшение области поиска вдвое на каждой итерации, однако, на каждой итерации (кроме самой первой) может потребоваться не более двух измерений функции.

### 85. Метод золотого сечения.

💡 Решаемая задача:  $\min_{x \in [a, b]} f(x)$ . На отрезке выбираются правая и левая точки золотого сечения. По результатам измерения значений функции в этих двух точках решение может оказаться либо от начала до правой точки золотого сечения, либо от левой точки золотого сечения до конца отрезка. В любом из этих двух случаев на следующей итерации метода, одна из предыдущих измеренных точек золотого сечения станет точкой (левая на предыдущей итерации станет правой на новой или правая на предыдущей итерации станет левой на новой) золотого сечения отрезка на следующей итерации. Таким образом, нам останется доизмерить значение функции не более, чем один раз за итерацию.

### 86. Метод параболической интерполяции (без точных формул).

💡 Идея метода: берем 3 точки, по этим 3 точкам однозначно строим параболу, находим ее минимум, и из этих 4 точек оставляем 3 так, чтобы между первой и третьей находился минимум.

Сходится сверхлинейно, но метод довольно неустойчивый. Если  $f(x)$  не похожа на параболу, нам конец. Если она обратна параболе, то мы и вовсе уйдём искать максимум.

87. Условие достаточного убывания для неточного линейного поиска.

💡 Неточный линейный поиск:

$$x_{k+1} = x_k - \alpha \nabla f(x_k), \alpha = \arg \min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k)).$$

Хотим приближенно найти  $\alpha$ . Сведем задачу к поиску минимума следующей функции:

$$\phi(\alpha) = f(x_k - \alpha \nabla f(x_k)), \alpha \geq 0$$

Приблизим ее через первые 2 члена ряда Тейлора:

$$\phi(\alpha) \approx f(x_k) - \alpha \nabla f(x_k)^\top \nabla f(x_k)$$

Тогда условием достаточного убывания (Armijo condition) является:

$$f(x_k - \alpha \nabla f(x_k)) \leq f(x_k) - c_1 \cdot \alpha \nabla f(x_k)^\top \nabla f(x_k), c_1 \in (0, 1)$$

Иллюстрация для понимания:

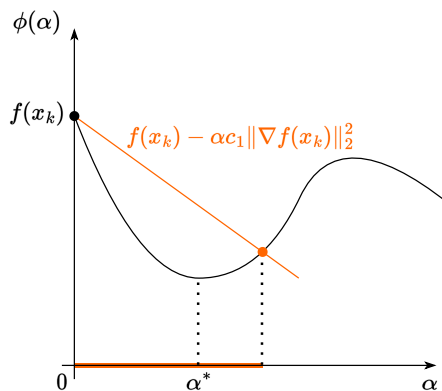


Рисунок 12: Иллюстрация условия достаточного убывания

88. Условия Гольдштейна для неточного линейного поиска.



💡 Определим  $\phi_1$  и  $\phi_2$  следующим образом ( $0 < c_1 < c_2 < 1$ )

$$\phi_1(\alpha) = f(x_k) - c_1 \alpha \|\nabla f(x_k)\|_2^2$$

$$\phi_2(\alpha) = f(x_k) - c_2 \alpha \|\nabla f(x_k)\|_2^2$$

Тогда условие Гольдштейна заключается в том, что  $\phi_2(\alpha) \leq \phi(\alpha) \leq \phi_1(\alpha)$ .  
Иллюстрация для понимания:

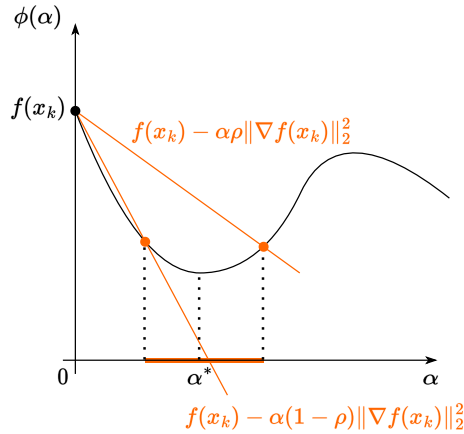


Рисунок 13: Иллюстрация условий Гольдштейна

89. Условие ограничения на кривизну для неточного линейного поиска.



$$-\nabla f(x_k - \alpha \nabla f(x_k))^\top \nabla f(x_k) \geq c_2 \nabla f(x_k)^\top (-\nabla f(x_k)),$$

где  $c_2 \in (c_1, 1)$ , и  $c_1$  взято из условия достаточного убывания.

Иллюстрация для понимания:

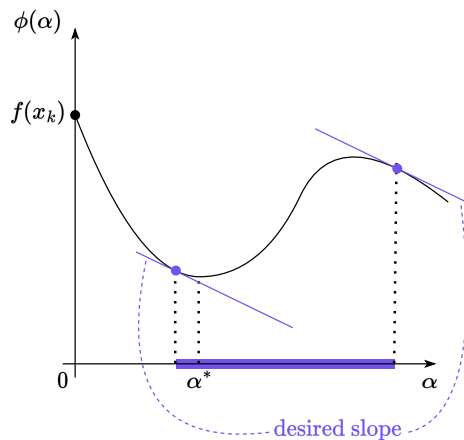


Рисунок 14: Иллюстрация условия ограничения на кривизну

90. Показать, что направление антиградиента - направление наискорейшего локального убывания

функции.

💡 Пусть  $f$  дифференцируема, зададим искомое направление локального убывания -  $h$  -  $\|h\| = 1$ . Тогда её аппроксимация:  $f(x + \alpha h) = f(x) + \alpha \langle \nabla f(x), h \rangle + o(\alpha)$

$$f(x + \alpha h) < f(x) \Rightarrow \alpha \langle \nabla f(x), h \rangle + o(\alpha) < 0.$$

При  $\alpha \rightarrow +0$  получаем:  $\alpha \langle \nabla f(x), h \rangle \leq 0$

$$\|\langle \nabla f(x), h \rangle\| \leq \|\nabla f(x)\| \|h\| \leq \|\nabla f(x)\|$$

$$\langle \nabla f(x), h \rangle \geq -\|\nabla f(x)\| \Rightarrow h = \frac{-\nabla f(x)}{\|\nabla f(x)\|}, \text{ ч.т.д.}$$

91. Дифференциальное уравнение градиентного потока.

💡 **Градиентный поток** (continuous-time аналог градиентного спуска) для гладкой функции  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  задаётся ОДУ

$$\dot{x}(t) = -\nabla f(x(t)), \quad x(0) = x_0.$$

В евклидовой метрике это движение в направлении наискорейшего убывания.

По правилу цепочки вдоль траектории:

$$\frac{d}{dt} f(x(t)) = \nabla f(x(t))^\top \dot{x}(t) = -\|\nabla f(x(t))\|_2^2 \leq 0,$$

т.е. значение  $f$  не возрастает.

92. Метод градиентного спуска.

💡 Решаем задачу минимизации

$$f(x) \rightarrow \min_{x \in \mathbb{R}^d}$$

Если  $f$  дифференцируема, то тогда для решения этой задачи можно использовать метод градиентного спуска:

$$x_{k+1} = x_k - \alpha \nabla f(x_k)$$

93. Наискорейший спуск.

💡 Решаем задачу минимизации

$$f(x) \rightarrow \min_{x \in \mathbb{R}^d}$$

Если  $f$  дифференцируема, то тогда для решения этой задачи можно использовать метод наискорейшего спуска:

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

$$\alpha_k = \arg \min_{\alpha \in \mathbb{R}^+} f(x_k - \alpha \nabla f(x_k)),$$

т.е. выбираем наилучший шаг спуска на каждой итерации метода.

94. Как направлены две соседние итерации метода наискорейшего спуска по отношению друг к другу?

💡 Шаги между итерациями ортогональны друг другу.

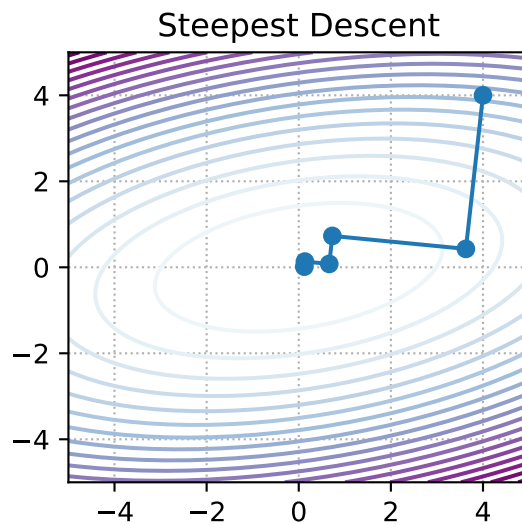


Рисунок 15: Несколько итераций наискорейшего спуска

95. Липшицева парабола для гладкой функции.

💡 Если  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  - непрерывно дифференцируема и градиент Липшицев с константой  $L$ , то  $\forall x, y \in \mathbb{R}^n$ :

$$|f(y) - f(x) - \langle \nabla f(x), y - x \rangle| \leq \frac{L}{2} \|y - x\|^2$$

Если зафиксируем  $x_0 \in \mathbb{R}^n$ , то:

$$\varphi_1(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle - \frac{L}{2} \|x - x_0\|^2$$

$$\varphi_2(x) = f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{L}{2} \|x - x_0\|^2$$

Это две параболы, и для них верно, что  $\varphi_1(x) \leq f(x) \leq \varphi_2(x) \forall x$

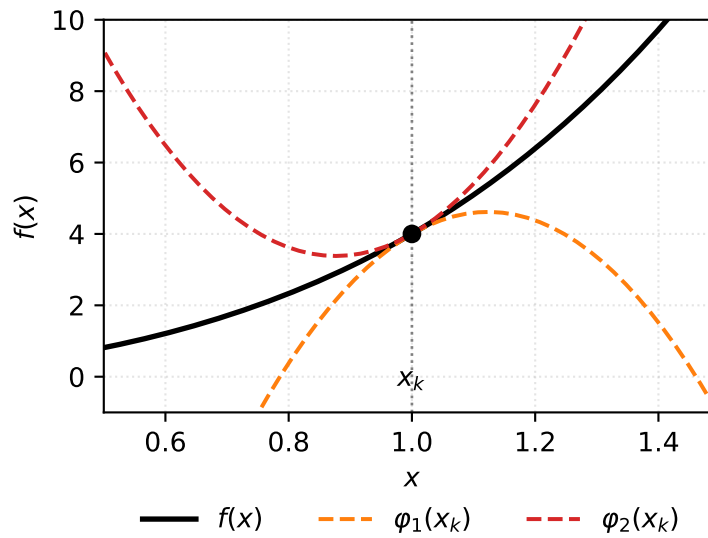


Рисунок 16: Иллюстрация Липшицевых парабол, между которыми зажата гладкая функция. Чаще нас интересует мажорирующая из них.

96. Размер шага наискорейшего спуска для квадратичной функции.

💡 Решаем задачу минимизации методом наискорейшего спуска

$$f(x) = \frac{1}{2} x^T A x - b^T x + c \rightarrow \min_{x \in \mathbb{R}^d}$$

$$\nabla f = \frac{1}{2} (A + A^T) x - b$$

Из условия  $\nabla f(x_{k+1})^T \nabla f(x_k) = 0$  получаем:

$$\alpha_k = \frac{2 \nabla f(x_k)^T \nabla f(x_k)}{\nabla f(x_k)^T (A + A^T) \nabla f(x_k)} = \frac{\nabla f(x_k)^T \nabla f(x_k)}{\nabla f(x_k)^T \nabla^2 f(x_k) \nabla f(x_k)}.$$

97. Характер сходимости градиентного спуска к локальному экстремуму для гладких невыпуклых функций в терминах  $\mathcal{O}$  от числа итераций метода.

- i** Пусть  $f$   $L$ -гладкая (градиент  $L$ -липшицев), и  $f$  ограничена снизу:  $f(x) \geq f_{\inf}$ . Рассмотрим градиентный спуск

$$x_{k+1} = x_k - \alpha \nabla f(x_k), \quad \alpha \in (0, 1/L].$$

Тогда выполняется **лемма о спуске**:

$$f(x_{k+1}) \leq f(x_k) - \frac{\alpha}{2} \|\nabla f(x_k)\|_2^2.$$

Суммируя по  $k = 0, \dots, N-1$ , получаем

$$\min_{0 \leq k \leq N-1} \|\nabla f(x_k)\|_2^2 \leq \frac{2(f(x_0) - f_{\inf})}{\alpha N} = \mathcal{O}\left(\frac{1}{N}\right).$$

Следовательно, чтобы найти  $\varepsilon$ -стационарную точку ( $\|\nabla f(x_k)\|_2 \leq \varepsilon$ ), достаточно

$$N = \mathcal{O}\left(\frac{1}{\varepsilon^2}\right).$$

В невыпуклом случае в общем виде гарантируется сходимость к **стационарной точке** (а не обязательно к локальному минимуму), при стандартных предположениях о шаге/линейном поиске и отсутствии вырождения.

98. Характер сходимости градиентного спуска для гладких выпуклых функций в терминах  $\mathcal{O}$  от числа итераций метода.

$f(x_k) - f^* \sim \mathcal{O}\left(\frac{1}{k}\right).$

99. Характер сходимости градиентного спуска для гладких и сильно выпуклых функций в терминах  $\mathcal{O}$  от числа итераций метода.

$\|x_k - x^*\|^2 \sim \mathcal{O}\left(\left(1 - \frac{\mu}{L}\right)^k\right).$

100. Связь спектра гессиана с константами сильной выпуклости и гладкости функции.

$\mu = \min_{x \in \text{dom} f} \lambda_{\min}(\nabla^2 f(x)), \quad L = \max_{x \in \text{dom} f} \lambda_{\max}(\nabla^2 f(x)).$

101. Условие Поляка-Лоясевича (градиентного доминирования) для функций.

$\exists \mu > 0 : \quad \|\nabla f(x)\|^2 \geq 2\mu(f(x) - f^*) \quad \forall x, \text{ где } f^* - \text{минимум функции } f(x).$

102. Сходимость градиентного спуска для сильно выпуклых квадратичных функций. Оптимальные гиперпараметры.

- 💡 Решаем задачу минимизации методом градиентного спуска. Пусть  $A \in \mathbb{S}_{++}^n \Rightarrow \nabla f = Ax - b$ .

$$f(x) = \frac{1}{2}x^T Ax - b^T x + c \rightarrow \min_{x \in \mathbb{R}^d}$$

$$x_{k+1} = x_k - \alpha(Ax_k - b)$$

$$\alpha_{opt} = \frac{2}{\mu + L}, \text{ где } \mu = \lambda_{\min}(A), L = \lambda_{\max}(A)$$

$$\kappa = \frac{L}{\mu} \geq 1$$

$$\rho = \frac{\kappa - 1}{\kappa + 1}$$

$$\|x_k - x^*\| \leq \rho^k \|x_0 - x^*\|$$

103. Связь PL-функций и сильно выпуклых функций.

- 💡 Пусть  $f$   $\mu$ -сильно выпуклая и дифференцируемая  $\Rightarrow f \in \text{PL}$ .  
Обратное неверно -  $f(x) = x^2 + 3\sin^2 x \in \text{PL}$ , но не сильно выпуклая (она вообще не выпуклая).

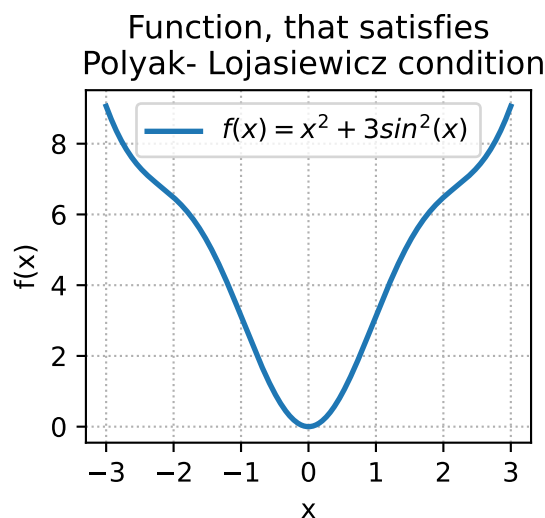


Рисунок 17: Пример невыпуклой PL функции

104. Привести пример выпуклой, но не сильно выпуклой задачи линейных наименьших квадратов (возможно, с регуляризацией).

- 💡 Рассмотрим задачу минимизации функции:

$$\|Ax - b\|^2 \rightarrow \min_{x \in \mathbb{R}^d},$$

где матрица  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $m < n$  (лежащая).

105. Привести пример сильно выпуклой задачи линейных наименьших квадратов (возможно, с

регуляризацией).

💡 Рассмотрим задачу минимизации функции:

$$f(x) = \|Ax - b\|_2^2,$$

где  $A \in \mathbb{R}^{n \times n}$  (ранг  $A = n$ ). Эта функция сильно выпукла, так как гессиан положительно определен.

106. Нижние оценки для гладкой выпуклой оптимизации с помощью методов первого порядка в терминах  $\mathcal{O}$  от числа итераций метода.

💡  $f(x_k) - f^* \sim \Omega\left(\frac{1}{k^2}\right)$

107. Нижние оценки для гладкой сильно выпуклой оптимизации с помощью методов первого порядка в терминах  $\mathcal{O}$  от числа итераций метода.

💡 
$$f(x_k) - f^* = \Omega\left(\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k}\right)$$

108. Отличие ускоренной и неускоренной линейной сходимости для методов первого порядка.

💡

Функция	Неускоренная	Ускоренная
Гладкая и сильно-выпуклая (или PL)	$\mathcal{O}\left(\kappa \log \frac{1}{\varepsilon}\right)$	$\mathcal{O}\left(\sqrt{\kappa} \log \frac{1}{\varepsilon}\right)$
Гладкая и выпуклая (в вопросе не требуется)	$\mathcal{O}\left(\frac{1}{\varepsilon}\right)$	$\mathcal{O}\left(\frac{1}{\sqrt{\varepsilon}}\right)$

109. Метод тяжелого шарика (Поляка).

💡 Задача:  $f(x) \rightarrow \min_{x \in \mathbb{R}^d}$ ,  $f(x)$  - непрерывно дифференцируемая функция

$$x_{k+1} = x_k - \alpha \nabla f(x_k) + \beta(x_k - x_{k-1}), \quad 0 < \beta < 1.$$

110. Понятие локальной и глобальной сходимости численного метода оптимизации.

💡 Локальная сходимость означает, что последовательность итераций сходится к решению  $x^*$  внутри некоторого множества. За пределами этого множества сходимости нет. Глобальная сходимость означает, что последовательность итераций сходится к решению  $x^*$  вне зависимости от начального приближения.

111. Ускоренный градиентный метод Нестерова для выпуклых гладких функций.

💡 Предположим, что  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  является выпуклой и  $L$ -гладкой. Ускоренный градиентный метод Нестерова (NAG) предназначен для решения задачи минимизации, начиная с начальной точки  $x_0 = y_0 \in \mathbb{R}^n$  и  $\lambda_0 = 0$ . Алгоритм выполняет следующие шаги:

**Обновление градиента:**  $x_{k+1} = y_k - \frac{1}{L} \nabla f(y_k)$

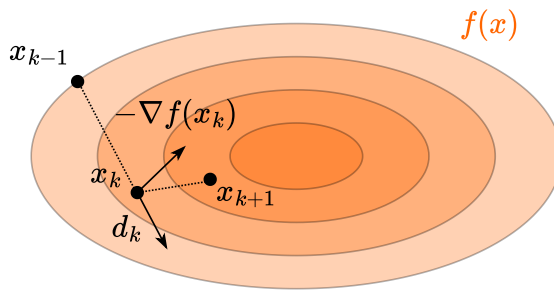
**Вес экстраполяции:**  $\lambda_{k+1} = \frac{1 + \sqrt{1 + 4\lambda_k^2}}{2}$   
 $\gamma_k = \frac{\lambda_k - 1}{\lambda_{k+1}}$

**Экстраполяция:**  $y_{k+1} = x_{k+1} + \gamma_k (x_{k+1} - x_k)$

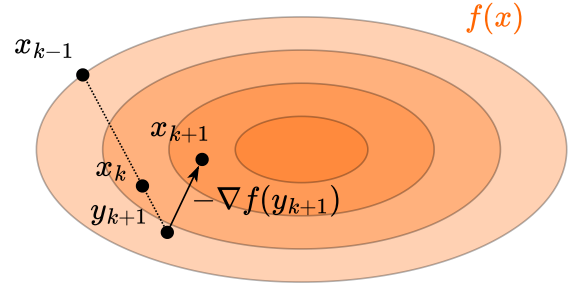
Последовательность  $\{f(x_k)\}_{k \in \mathbb{N}}$ , генерируемая алгоритмом, сходится к оптимальному значению  $f^*$  со скоростью  $\mathcal{O}\left(\frac{1}{k^2}\right)$ , в частности:

$$f(x_k) - f^* \leq \frac{2L\|x_0 - x^*\|^2}{k^2}$$

Polyak momentum



Nesterov momentum



112. Ускоренный градиентный метод Нестерова для сильно выпуклых гладких функций.

💡 Предположим, что  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  является  $\mu$ -сильно выпуклой и  $L$ -гладкой. Ускоренный градиентный метод Нестерова (NAG) предназначен для решения задачи минимизации, начиная с начальной точки  $x_0 = y_0 \in \mathbb{R}^n$  и  $\lambda_0 = 0$ . Алгоритм выполняет следующие шаги:

**Обновление градиента:**  $x_{k+1} = y_k - \frac{1}{L} \nabla f(y_k)$

**Экстраполяция:**  $y_{k+1} = x_{k+1} + \gamma (x_{k+1} - x_k)$

**Вес экстраполяции:**  $\gamma = \frac{\sqrt{L} - \sqrt{\mu}}{\sqrt{L} + \sqrt{\mu}}$

Последовательность  $\{f(x_k)\}_{k \in \mathbb{N}}$ , генерируемая алгоритмом, сходится к оптимальному значению  $f^*$  линейно:

$$f(x_k) - f^* \leq \frac{\mu + L}{2} \|x_0 - x^*\|_2^2 \exp\left(-\frac{k}{\sqrt{\kappa}}\right)$$

113.  $A$ -сопряженность двух векторов.  $A$ -ортогональность. Скалярное произведение  $\langle \cdot, \cdot \rangle_A$ .



💡  $A$ -ортогональность (сопряженность):

$$x \perp_A y \iff x^T A y = 0.$$

114. Процедура ортогонализации Грама-Шмидта.

💡 **Классический Грам–Шмидт.** Даны линейно независимые векторы  $a_1, \dots, a_m \in \mathbb{R}^n$ . Строим ортонормированный набор  $q_1, \dots, q_m$ :

$$u_1 = a_1, \quad q_1 = \frac{u_1}{\|u_1\|_2},$$
$$u_k = a_k - \sum_{j=1}^{k-1} \langle q_j, a_k \rangle q_j, \quad q_k = \frac{u_k}{\|u_k\|_2}, \quad k = 2, \dots, m.$$

В результате  $q_i^T q_j = \delta_{ij}$  и  $\text{span}\{a_1, \dots, a_m\} = \text{span}\{q_1, \dots, q_m\}$ .

Для матрицы  $A = [a_1 \dots a_m]$  получаем **QR-разложение**:  $A = QR$ , где  $Q = [q_1 \dots q_m]$ , а  $R$  – верхнетреугольная,  $R_{jk} = \langle q_j, a_k \rangle$ .

(На практике часто используют модифицированный Грам–Шмидт из-за лучшей численной устойчивости.)

115. Метод сопряженных направлений.

💡 **Метод сопряжённых направлений** применяется для минимизации квадратичной функции

$$f(x) = \frac{1}{2} x^T A x - b^T x, \quad A \succ 0.$$

Векторы  $p_0, \dots, p_{n-1}$  называются  **$A$ -сопряжёнными** (conjugate), если

$$p_i^T A p_j = 0, \quad i \neq j.$$

Итерации метода:

$$x_{k+1} = x_k + \alpha_k p_k, \quad \alpha_k = -\frac{\nabla f(x_k)^T p_k}{p_k^T A p_k} = \frac{r_k^T p_k}{p_k^T A p_k},$$

где  $r_k = b - A x_k = -\nabla f(x_k)$  (невязка).

При точном одномерном поиске и выборе  $n$  линейно независимых  $A$ -сопряжённых направлений метод находит точный минимум за **не более чем  $n$  шагов** (в точной арифметике), т.к. минимизация последовательно выполняется по взаимно  $A$ -ортогональным направлениям.

116. Метод сопряженных градиентов.

💡 **Метод сопряжённых градиентов (CG)** — частный случай метода сопряжённых направлений, где направления строятся рекуррентно по градиентам.

Эквивалентные постановки:

- решить СЛАУ  $Ax = b$  при  $A \succ 0$ ;
- минимизировать  $f(x) = \frac{1}{2}x^\top Ax - b^\top x$ .

Обозначим невязку  $r_k = b - Ax_k = -\nabla f(x_k)$  и направление  $p_k$ .

**Алгоритм (линейный CG):**

$$r_0 = b - Ax_0, \quad p_0 = r_0;$$

для  $k = 0, 1, 2, \dots$ :

$$\alpha_k = \frac{r_k^\top r_k}{p_k^\top A p_k}, \quad x_{k+1} = x_k + \alpha_k p_k,$$

$$r_{k+1} = r_k - \alpha_k A p_k,$$

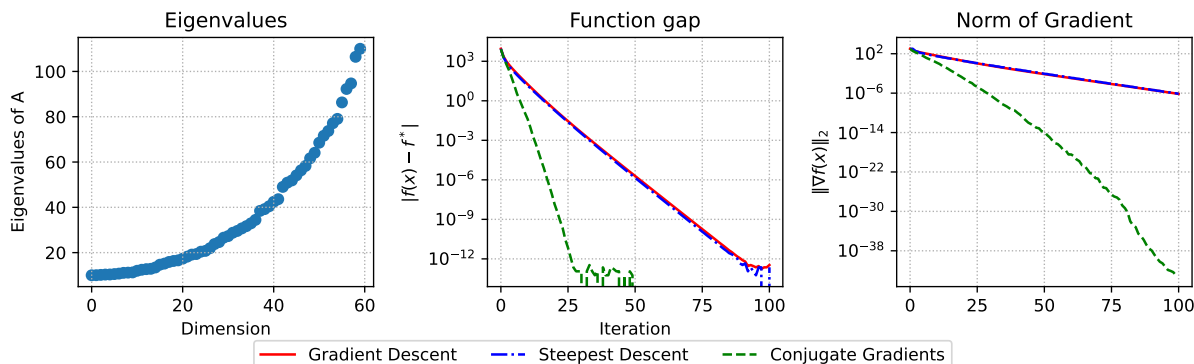
$$\beta_k = \frac{r_{k+1}^\top r_{k+1}}{r_k^\top r_k}, \quad p_{k+1} = r_{k+1} + \beta_k p_k.$$

Свойства (в точной арифметике):  $p_k$  —  $A$ -сопряжённые,  $r_k$  попарно ортогональны, а  $x_k$  минимизирует  $f$  на аффинном подпространстве  $x_0 + \mathcal{K}_k(A, r_0)$  (крыловском подпространстве).

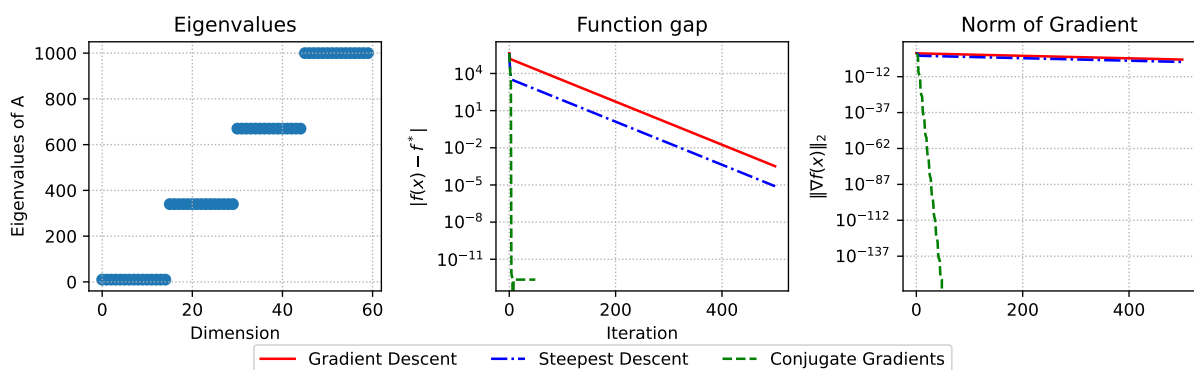
117. Зависимость сходимости метода сопряженных градиентов от спектра матрицы.

- 💡 Если матрица  $A$  имеет только  $r$  различных собственных чисел, тогда метод сопряжённых градиентов сходится за  $r$  итераций.

Strongly convex quadratics.  $n=60$ , random matrix.



Strongly convex quadratics.  $n=60$ , clustered matrix.



118. Характер сходимости метода сопряжённых градиентов в терминах  $\mathcal{O}$  от числа итераций метода.

❗ Пусть  $A \succ 0$  и  $\kappa = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$  — число обусловленности.

Для линейного CG выполняется оценка ошибки в  $A$ -норме:

$$\|x_k - x_*\|_A \leq 2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k \|x_0 - x_*\|_A, \quad \|v\|_A := \sqrt{v^\top A v}.$$

То есть сходимость **линейная (геометрическая)** с фактором  $q = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} < 1$ .

Для достижения относительной точности  $\varepsilon$  достаточно

$$k = \mathcal{O}(\sqrt{\kappa} \log(1/\varepsilon)).$$

Кроме того, в точной арифметике метод завершится за  $\leq n$  итераций (не позднее размера задачи), т.к. строит решение в возрастающих крыловских подпространствах.

119. Метод Флетчера-Ривза.

💡 **Метод Флетчера–Ривза (FR)** — вариант *нелинейного* метода сопряжённых градиентов для минимизации гладкой функции  $f$ .

Обозначим  $g_k = \nabla f(x_k)$ . Инициализация:  $d_0 = -g_0$ .

Итерация:

$$x_{k+1} = x_k + \alpha_k d_k,$$

где  $\alpha_k$  выбирается одномерным поиском (обычно условия Вольфа/сильные Вольфа).

Коэффициент:

$$\beta_k^{\text{FR}} = \frac{\|g_{k+1}\|_2^2}{\|g_k\|_2^2},$$

направление:

$$d_{k+1} = -g_{k+1} + \beta_k^{\text{FR}} d_k.$$

При квадратичной  $f$  и точном линейном поиске FR совпадает с линейным CG.

## 120. Метод Полака-Рибьера.

💡 **Метод Полака–Рибьера (PR)** — ещё один вариант *нелинейного* сопряжённого градиента.

Обозначим  $g_k = \nabla f(x_k)$ ,  $d_0 = -g_0$ , и делаем шаг

$$x_{k+1} = x_k + \alpha_k d_k$$

с одномерным поиском (условия Вольфа/сильные Вольфа).

Коэффициент Полака–Рибьера:

$$\beta_k^{\text{PR}} = \frac{g_{k+1}^\top (g_{k+1} - g_k)}{\|g_k\|_2^2},$$

направление:

$$d_{k+1} = -g_{k+1} + \beta_k^{\text{PR}} d_k.$$

Часто используют модификацию **PR+**:  $\beta_k = \max\{\beta_k^{\text{PR}}, 0\}$ , чтобы гарантировать спуск при стандартных условиях линейного поиска.

## Теоремы с доказательствами

1. Критерий положительной определенности матрицы через знаки собственных значений симметричной матрицы.

**i**  $A \succeq (>)0 \iff$  все собственные значения симметричной матрицы  $A \geq (>)0$

→ Пусть некоторые собственные значения  $\lambda$  отрицательны, и  $x$  - соответствующий ему собственный вектор. Тогда:

$$Ax = \lambda x, x^T Ax \geq 0 \rightarrow x^T Ax = \lambda x^T x, x^T x \geq 0 \rightarrow \lambda \geq 0 - \text{противоречие}$$

← Помним, что положительная определённость задаётся для симметричных матриц. Для симметричной матрицы можем выбрать собственные векторы  $v_i$ , образующие ортогональный базис ( $i \neq j : v_i^T v_j = 0$  - выкидываем часть слагаемых из суммы в доказательстве). Тогда для  $x \in \mathbb{R}^n$

$$x^T Ax = (\alpha_1 v_1 + \dots + \alpha_n v_n)^T A(\alpha_1 v_1 + \dots + \alpha_n v_n) = \sum \alpha_i^2 v_i^T A v_i = \sum \alpha_i^2 v_i^T \lambda_i v_i$$

Так как  $\lambda_i \geq 0$ , то и вся сумма неотрицательна.

2. Базовые операции, сохраняющие выпуклость множеств: пересечение бесконечного числа множеств, линейная комбинация множеств, образ аффинного отображения.

- i**
- Пересечение любого (!) количества выпуклых множеств — выпуклое множество.
  - Линейная комбинация выпуклых множеств выпукла.
  - Образ выпуклого множества после применения аффинного отображения — выпуклое множество.

### Пересечение бесконечного числа множеств

Пересечение любого (!) количества выпуклых множеств — выпуклое множество.

Если итоговое пересечение пустое или содержит одну точку, то свойство выпуклости выполняется по определению. Иначе возьмем 2 точки и отрезок между ними. Эти точки должны лежать во всех пересекаемых множествах. Так как все пересекаемые множества выпуклы, отрезок между этими двумя точками лежит во всех множествах. А значит, отрезок лежит и в их пересечении.

### Линейная комбинация множеств

Линейная комбинация выпуклых множеств выпукла.

Пусть есть 2 выпуклых множества  $S_x, S_y$ , рассмотрим их линейную комбинацию

$$S = \{s \mid s = c_1 x + c_2 y, x \in S_x, y \in S_y, c_1, c_2 \in \mathbb{R}\}$$

Возьмем две точки из  $S$ :  $s_1 = c_1 x_1 + c_2 y_1, s_2 = c_1 x_2 + c_2 y_2$  и докажем, что отрезок между ними  $\theta s_1 + (1 - \theta) s_2, \theta \in [0, 1]$  также принадлежит  $S$

$$\theta s_1 + (1 - \theta) s_2$$

$$\begin{aligned} & \theta(c_1 x_1 + c_2 y_1) + (1 - \theta)(c_1 x_2 + c_2 y_2) \\ & c_1(\theta x_1 + (1 - \theta)x_2) + c_2(\theta y_1 + (1 - \theta)y_2) \end{aligned}$$

$$c_1x + c_2y \in S$$

### Образ аффинного отображения

Образ выпуклого множества после применения аффинного отображения — выпуклое множество.

$$S \subseteq \mathbb{R}^n \text{ выпукло} \rightarrow f(S) = \{f(x) \mid x \in S\} \text{ выпукло} \quad (f(x) = Ax + b)$$

### Доказательство

При  $\theta \in [0, 1]$ ;  $x, y \in S$ ,  $S$  — выпуклое. Тогда и  $\theta x + (1 - \theta)y \in S$ . В то же время  $f(\theta x + (1 - \theta)y) = \theta Ax + \theta b + (1 - \theta)Ay + (1 - \theta)b = \theta Ax + (1 - \theta)Ay + b = \theta f(x) + (1 - \theta)f(y)$ . В итоге мы доказали, что образ  $f(S)$  — тоже выпуклый, так как  $\forall \theta \in [0, 1], x, y \in S$  выполняется  $\theta f(x) + (1 - \theta)f(y) \in f(S)$ .

Примеры аффинных функций: растяжение, сжатие, проекция, транспонирование, множество решений линейного матричного неравенства  $\{x \mid x_1 A_1 + \dots + x_m A_m \preceq B\}$ . Здесь  $A_i, B \in \mathbb{S}^p$  — симметричные матрицы  $p \times p$ .

Заметим также, что прообраз выпуклого множества при аффинном отображении также является выпуклым.

$$S \subseteq \mathbb{R}^m \text{ выпукло} \rightarrow f^{-1}(S) = \{x \in \mathbb{R}^n \mid f(x) \in S\} \text{ выпукло} \quad (f(x) = Ax + b)$$

### 3. Неравенство Йенсена для выпуклой функции и выпуклой комбинации точек.

**i** Пусть  $f(x)$  — выпуклая функция, определённая на выпуклом множестве  $S \subseteq \mathbb{R}^n$ . Тогда для точек  $x_1, \dots, x_m \in S$  выполнено неравенство:

$$f\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i f(x_i)$$

$$\lambda = [\lambda_1, \dots, \lambda_m] \in \Delta_m.$$

1. Заметим, что  $\sum_{i=1}^m \lambda_i x_i$  является выпуклой комбинацией элементов  $S$  и лежит в  $S$ .
2. Доказательство по индукции. Для  $m = 1$  очевидно, для  $m = 2$  следует из определения выпуклой функции.
3. Пусть неравенство верно для  $m = 1, \dots, k$ , докажем для  $m = k + 1$ . Пусть  $\lambda \in \Delta_{k+1}$ ,  $x = \sum_{i=1}^{k+1} \lambda_i x_i = \lambda_{k+1} x_{k+1} + \sum_{i=1}^k \lambda_i x_i$ . При  $\lambda_i = 0$  либо 1 выражение сводится к уже рассмотренным случаям, далее полагаем  $0 < \lambda_i < 1$ :

$$x = \lambda_{k+1} x_{k+1} + (1 - \lambda_{k+1}) \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x_i = \lambda_{k+1} x_{k+1} + (1 - \lambda_{k+1}) \hat{x}$$

$$\text{где } \hat{x} = \sum_{i=1}^k \gamma_i x_i \text{ и } \gamma_i = \frac{\lambda_i}{1 - \lambda_{k+1}} \geq 0, \quad 1 \leq i \leq k.$$

4. Так как  $\lambda \in \Delta_{k+1}$ , то  $\gamma = [\gamma_1, \dots, \gamma_k] \in \Delta_k$ . Значит,  $\hat{x} \in S$ , из выпуклости  $f(x)$  и предположения индукции следует:

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) = f(\lambda_{k+1} x_{k+1} + (1 - \lambda_{k+1})\hat{x}) \leq \lambda_{k+1} f(x_{k+1}) + (1 - \lambda_{k+1}) f(\hat{x}) \leq \sum_{i=1}^{k+1} \lambda_i f(x_i)$$

4. Выпуклость надграфика как критерий выпуклости функции.

**i** Чтобы функция  $f(x)$ , определенная на выпуклом множестве  $X$ , была выпуклой на  $X$ , необходимо и достаточно чтобы надграфик  $f$  был выпуклым множеством.

Для функции  $f(x)$ , определенной на  $X \subseteq \mathbb{R}^n$ , множество:

$$\text{epi } f = \{[x, \mu] \in X \times \mathbb{R} : f(x) \leq \mu\}$$

называется **надграфиком** функции  $f(x)$  (здесь  $\mu \in \mathbb{R}, x \in S$ ).

### Необходимость

Предположим, что  $f(x)$  выпукла на  $X$ . Возьмем две произвольные точки  $[x_1, \mu_1] \in \text{epi } f$  и  $[x_2, \mu_2] \in \text{epi } f$ . Также возьмем  $0 \leq \lambda \leq 1$  и обозначим  $x_\lambda = \lambda x_1 + (1 - \lambda)x_2$ ,  $\mu_\lambda = \lambda \mu_1 + (1 - \lambda)\mu_2$ . Тогда,

$$\lambda \begin{bmatrix} x_1 \\ \mu_1 \end{bmatrix} + (1 - \lambda) \begin{bmatrix} x_2 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} x_\lambda \\ \mu_\lambda \end{bmatrix}.$$

Из выпуклости  $X$  следует, что  $x_\lambda \in X$ . Более того, так как  $f(x)$  – выпуклая функция, то

$$f(x_\lambda) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda \mu_1 + (1 - \lambda)\mu_2 = \mu_\lambda$$

Из неравенства выше по определению надграфика следует, что  $\begin{bmatrix} x_\lambda \\ \mu_\lambda \end{bmatrix} \in \text{epi } f$ . Следовательно, надграфик  $f$  – выпуклое множество.

### Достаточность

Предположим, что надграфик  $f$ ,  $\text{epi } f$ , выпуклое множество. Тогда, исходя из того что  $[x_1, \mu_1] \in \text{epi } f$  и  $[x_2, \mu_2] \in \text{epi } f$ , получаем

$$\begin{bmatrix} x_\lambda \\ \mu_\lambda \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ \mu_1 \end{bmatrix} + (1 - \lambda) \begin{bmatrix} x_2 \\ \mu_2 \end{bmatrix} \in \text{epi } f$$

для любого  $0 \leq \lambda \leq 1$ .

Следовательно, из определения надграфика, подставив значение  $\mu_\lambda$ , получаем, что  $f(x_\lambda) \leq \mu_\lambda = \lambda \mu_1 + (1 - \lambda)\mu_2$ .

$$f(x_\lambda) = f(\lambda x_1 + (1 - \lambda)x_2) \leq \mu_\lambda = \lambda \mu_1 + (1 - \lambda)\mu_2$$

Но это верно для всех  $\mu_1 \geq f(x_1)$  и  $\mu_2 \geq f(x_2)$ , в том числе и при  $\mu_1 = f(x_1)$  и  $\mu_2 = f(x_2)$ . Тогда мы получаем неравенство:

$$f(x_\lambda) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

Так как  $x_1 \in X$  и  $x_2 \in X$  выбирались произвольно,  $f(x)$  – выпуклая функция на  $X$ .

5. Дифференциальный критерий сильной выпуклости первого порядка.

**i** Пусть  $f(x)$  — дифференцируемая функция на выпуклом множестве  $X \subseteq \mathbb{R}^n$ . Тогда  $f(x)$  сильно выпукла на  $X$  с константой  $\mu > 0$  тогда и только тогда, когда

$$f(x) - f(x_0) \geq \langle \nabla f(x_0), x - x_0 \rangle + \frac{\mu}{2} \|x - x_0\|^2$$

для всех  $x, x_0 \in X$ .

### Необходимость

Пусть  $0 < \lambda \leq 1$ . Согласно определению сильно выпуклой функции,

$$f(\lambda x + (1 - \lambda)x_0) \leq \lambda f(x) + (1 - \lambda)f(x_0) - \frac{\mu}{2} \lambda(1 - \lambda) \|x - x_0\|^2$$

или эквивалентно,

$$\begin{aligned} f(x) - f(x_0) - \frac{\mu}{2} (1 - \lambda) \|x - x_0\|^2 &\geq \frac{1}{\lambda} [f(\lambda x + (1 - \lambda)x_0) - f(x_0)] = \\ &= \frac{1}{\lambda} [f(x_0 + \lambda(x - x_0)) - f(x_0)] = \frac{1}{\lambda} [\lambda \langle \nabla f(x_0), x - x_0 \rangle + o(\lambda)] = \\ &= \langle \nabla f(x_0), x - x_0 \rangle + \frac{o(\lambda)}{\lambda}. \end{aligned}$$

Таким образом, переходя к пределу при  $\lambda \rightarrow 0$ , мы приходим к первоначальному утверждению.

### Достаточность

Предположим, что неравенство в теореме выполнено для всех  $x, x_0 \in X$ . Возьмем  $x_0 = \lambda x_1 + (1 - \lambda)x_2$ , где  $x_1, x_2 \in X$ ,  $0 \leq \lambda \leq 1$ . Согласно неравенству из условия теоремы, выполняются следующие неравенства:

$$f(x_1) - f(x_0) \geq \langle \nabla f(x_0), x_1 - x_0 \rangle + \frac{\mu}{2} \|x_1 - x_0\|^2,$$

$$f(x_2) - f(x_0) \geq \langle \nabla f(x_0), x_2 - x_0 \rangle + \frac{\mu}{2} \|x_2 - x_0\|^2.$$

Умножая первое неравенство на  $\lambda$  и второе на  $1 - \lambda$  и складывая их, учитывая, что

$$x_1 - x_0 = (1 - \lambda)(x_1 - x_2), \quad x_2 - x_0 = \lambda(x_2 - x_1),$$

и что  $\lambda(1 - \lambda)^2 + \lambda^2(1 - \lambda) = \lambda(1 - \lambda)$ , получаем:

$$\begin{aligned} \lambda f(x_1) + (1 - \lambda)f(x_2) - f(x_0) - \frac{\mu}{2} \lambda(1 - \lambda) \|x_1 - x_2\|^2 &\geq \\ &\geq \langle \nabla f(x_0), \lambda x_1 + (1 - \lambda)x_2 - x_0 \rangle = 0. \end{aligned}$$



Таким образом, неравенство из определения сильно выпуклой функции выполнено. Важно отметить, что при  $\mu = 0$  получаем случай выпуклости и соответствующий дифференциальный критерий.

#### 6. Дифференциальный критерий сильной выпуклости второго порядка.

**i** Пусть  $X \subseteq \mathbb{R}^n$  — выпуклое множество с непустой внутренностью. Пусть также  $f(x)$  — дважды непрерывно дифференцируемая функция на  $X$ . Тогда  $f(x)$  сильно выпукла на  $X$  с константой  $\mu > 0$  тогда и только тогда, когда

$$\langle y, \nabla^2 f(x)y \rangle \geq \mu \|y\|^2$$

для всех  $x \in X$  и  $y \in \mathbb{R}^n$ .

Другая форма записи:

$$\nabla^2 f(x) \succcurlyeq \mu I$$

Целевое неравенство тривиально, когда  $y = 0_n$ , поэтому предположим, что  $y \neq 0_n$ .

#### Необходимость

Пусть  $x$  является внутренней точкой  $X$ . Тогда  $x + \alpha y \in X$  для всех  $y \in \mathbb{R}^n$  и достаточно малых  $\alpha$ . Поскольку  $f(x)$  дважды дифференцируема,

$$f(x + \alpha y) = f(x) + \alpha \langle \nabla f(x), y \rangle + \frac{\alpha^2}{2} \langle y, \nabla^2 f(x)y \rangle + o(\alpha^2).$$

Основываясь на критерии первого порядка сильной выпуклости, имеем

$$\frac{\alpha^2}{2} \langle y, \nabla^2 f(x)y \rangle + o(\alpha^2) = f(x + \alpha y) - f(x) - \alpha \langle \nabla f(x), y \rangle \geq \frac{\mu}{2} \alpha^2 \|y\|^2.$$

Это неравенство сводится к целевому неравенству после деления обеих частей на  $\alpha^2$  и перехода к пределу при  $\alpha \rightarrow 0$ .

Если  $x \in X$ , но  $x \notin \text{int} X$ , рассмотрим последовательность  $\{x_k\}$  такую, что  $x_k \in \text{int} X$  и  $x_k \rightarrow x$  при  $k \rightarrow \infty$ . Тогда мы приходим к целевому неравенству после перехода к пределу.

#### Достаточность

Формула Тейлора с остаточным членом Лагранжа второго порядка  $\forall x, y : x, x + y \in X$  найдется  $\alpha$  такая, что:

$$f(x + y) = f(x) + \langle \nabla f(x), y \rangle + \frac{1}{2} \langle y, \nabla^2 f(x + \alpha y)y \rangle$$

где  $0 < \alpha < 1$ .

Используя формулу Тейлора с остаточным членом Лагранжа и неравенство из условия, получаем для  $x + y \in X$ :

$$f(x + y) - f(x) - \langle \nabla f(x), y \rangle = \frac{1}{2} \langle y, \nabla^2 f(x + \alpha y)y \rangle \geq \frac{\mu}{2} \|y\|^2,$$

где  $0 \leq \alpha \leq 1$ . Следовательно,

$$f(x+y) - f(x) \geq \langle \nabla f(x), y \rangle + \frac{\mu}{2} \|y\|^2.$$

Таким образом, по критерию первого порядка сильной выпуклости, функция  $f(x)$  является сильно выпуклой с константой  $\mu$ . Важно отметить, что  $\mu = 0$  соответствует случаю выпуклости и соответствующему дифференциальному критерию.

7. Теорема о построении сопряженного множества к многогранному множеству.

**i** Пусть  $x_1, \dots, x_m \in \mathbb{R}^n$ . Сопряжённое к многогранному множеству:

$$S = \text{conv}(x_1, \dots, x_k) + \text{cone}(x_{k+1}, \dots, x_m)$$

будет многогранным множеством:

$$S^* = \{p \in \mathbb{R}^n \mid \langle p, x_i \rangle \geq -1, i = \overline{1, k}; \langle p, x_i \rangle \geq 0, i = \overline{k+1, m}\}$$

- Пусть  $S = X, S^* = Y$ . Возьмём произвольный  $p \in X^*$ , тогда  $\langle p, x_i \rangle \geq -1, i = \overline{1, k}$ . В то же время, для любого  $\theta > 0, i = \overline{k+1, m}$ :

$$\langle p, x_i \rangle \geq -1 \rightarrow \langle p, \theta x_i \rangle \geq -1$$

$$\langle p, x_i \rangle \geq -\frac{1}{\theta} \rightarrow \langle p, x_i \rangle \geq 0.$$

Таким образом,  $p \in Y \rightarrow X^* \subset Y$ .

- В обратную сторону: пусть  $p \in Y$ . Для любого  $x \in X$ :

$$x = \sum_{i=1}^m \theta_i x_i \quad \sum_{i=1}^k \theta_i = 1, \theta_i \geq 0$$

Тогда:

$$\langle p, x \rangle = \sum_{i=1}^m \theta_i \langle p, x_i \rangle = \sum_{i=1}^k \theta_i \langle p, x_i \rangle + \sum_{i=k+1}^m \theta_i \langle p, x_i \rangle \geq \sum_{i=1}^k \theta_i (-1) + \sum_{i=1}^k \theta_i \cdot 0 = -1.$$

Значит,  $p \in X^* \rightarrow Y \subset X^*$ .

8. Субдифференциальное условие оптимальности для условных выпуклых задач.

i

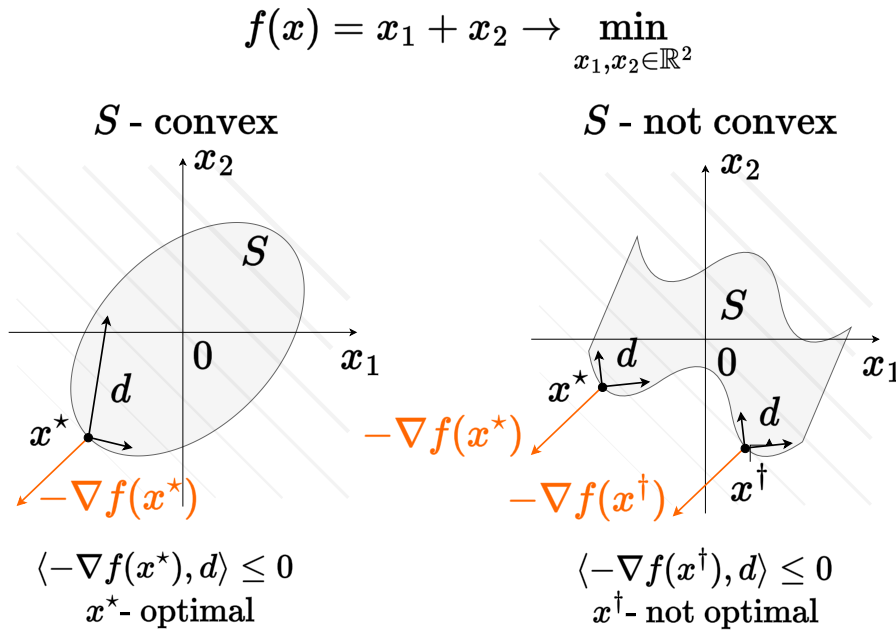


Рисунок 18: Условие оптимальности для условных выпуклых задач

Пусть  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  — выпуклая функция, а  $S \subset \mathbb{R}^n$  — некоторое (выпуклое) множество допустимых точек. Рассмотрим задачу оптимизации

$$\min_{x \in S} f(x).$$

Тогда точка  $x^*$  является решением этой задачи тогда и только тогда, когда

$$0 \in \partial f(x^*) + \mathcal{N}_S(x^*),$$

где  $\partial f(x^*)$  — субдифференциал функции  $f$  в точке  $x^*$ , а  $\mathcal{N}_S(x^*)$  — нормальный конус к множеству  $S$  в точке  $x^*$ .

Если же  $f$  дополнительно дифференцируема, то условие оптимальности принимает вид

$$\nabla f(x^*)^T (y - x^*) \geq 0 \quad \text{для всех } y \in S.$$

Рассмотрим задачу

$$\min_{x \in S} f(x).$$

### 1. Переход к неограниченной задаче:

Введём индикаторную функцию множества  $S$ , то есть

$$I_S(x) = \begin{cases} 0, & x \in S, \\ +\infty, & x \notin S. \end{cases}$$

Тогда исходная задача эквивалентна безусловной (без явных ограничений) задаче

$$\min_x (f(x) + I_S(x)).$$

**2. Условие оптимальности через субградиент:**

Из общего субдифференциального условия оптимальности следует, что точка  $x$  является решением  $\min_x \{f(x) + I_S(x)\}$  тогда и только тогда, когда

$$0 \in \partial(f(x) + I_S(x)).$$

**3. Свойство субградиента суммы:**

Поскольку  $f$  выпуклая и  $I_S$  — тоже выпуклая (это индикаторная функция выпуклого множества  $S$ ), имеем

$$\partial(f(x) + I_S(x)) = \partial f(x) + \partial I_S(x).$$

Но  $\partial I_S(x) = \mathcal{N}_S(x)$ , то есть *нормальный конус* к множеству  $S$  в точке  $x$ . Следовательно,

$$0 \in \partial f(x) + \partial I_S(x) \iff 0 \in \partial f(x) + \mathcal{N}_S(x).$$

**4. Интерпретация условия  $0 \in \partial f(x) + \mathcal{N}_S(x)$ :**

Это означает, что существует субградиент  $g \in \partial f(x)$  такой, что

$$-g \in \mathcal{N}_S(x).$$

**5. Частный случай дифференцируемой функции  $f$ :**

Если  $f$  дифференцируема, то  $\partial f(x) = \{\nabla f(x)\}$ . Условие

$$0 \in \{\nabla f(x)\} + \mathcal{N}_S(x) \iff -\nabla f(x) \in \mathcal{N}_S(x).$$

По определению нормального конуса,

$$-\nabla f(x) \in \mathcal{N}_S(x) \iff -\nabla f(x)^T x \geq -\nabla f(x)^T y \text{ для всех } y \in S,$$

что переписывается как

$$\nabla f(x)^T (y - x) \geq 0 \text{ для всех } y \in S.$$

Это и есть классическое *условие оптимальности первого порядка* для дифференцируемых выпуклых задач оптимизации с ограничениями.

**9. Необходимые условия безусловного экстремума.**

**i** Если в  $x^*$  достигается локальный минимум и  $f$  непрерывно дифференцируема в открытой окрестности  $x^*$ , то

$$\nabla f(x^*) = 0$$

Предположим обратное. Пусть  $\nabla f(x^*) \neq 0$ . Рассмотрим вектор  $p = -\nabla f(x^*)$  и заметим, что

$$p^T \nabla f(x^*) = -\|\nabla f(x^*)\|^2 < 0$$

Так как  $\nabla f$  непрерывна в окрестности  $x^*$ , то существует скаляр  $T > 0$  такой, что

$$p^T \nabla f(x^* + tp) < 0, \text{ для любого } t \in [0, T]$$

Для любого  $\bar{t} \in (0, T]$ , мы можем воспользоваться теоремой Тейлора:

$$f(x^* + \bar{t}p) = f(x^*) + \bar{t}p^T \nabla f(x^* + tp), \text{ для некоторого } t \in (0, \bar{t})$$

Следовательно,  $f(x^* + \bar{t}p) < f(x^*)$  для любого  $\bar{t} \in (0, T]$ . Мы нашли направление, идя вдоль которого из  $x^*$  функция  $f$  убывает. Тогда  $x^*$  — не точка локального минимума. Получили противоречие.

**10. Достаточные условия безусловного экстремума.**

**i** Пусть  $\nabla^2 f$  непрерывна в открытой окрестности  $x^*$  и

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

Тогда  $x^*$  – точка локального минимума  $f$ .

Так как гессиан непрерывен и положительно определен в  $x^*$ , то мы можем выбрать радиус  $r > 0$  такой, что  $\nabla^2 f(x)$  остается положительно определенной для всех  $x$  в открытом шаре  $B = \{z \mid \|z - x^*\| < r\}$ . Взяв любой ненулевой вектор  $p$ , для которого выполняется  $\|p\| < r$ , мы получаем  $x^* + p \in B$ , а также по формуле Тейлора:

$$\begin{aligned} f(x^* + p) &= f(x^*) + p^T \nabla f(x^*) + \frac{1}{2} p^T \nabla^2 f(z) p \\ &= f(x^*) + \frac{1}{2} p^T \nabla^2 f(z) p \end{aligned}$$

где  $z = x^* + tp$  для некоторого  $t \in (0, 1)$ . Так как  $z \in B$ , мы получаем  $p^T \nabla^2 f(z) p > 0$ , и следовательно  $f(x^* + p) > f(x^*)$ . Таким образом  $x^*$  – точка локального минимума.

11. Субдифференциальная форма теоремы Каруша Куна Таккера (доказательство). Необходимые условия ККТ для произвольной задачи математического программирования (только формулировка).

**i** Пусть  $X$  – линейное нормированное пространство, а  $f_j : X \rightarrow \mathbb{R}, j = 0, 1, \dots, m$ , – выпуклые собственные (никогда не принимающие значения  $-\infty$ , а также не тождественно равные  $\infty$ ) функции. Рассмотрим задачу

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in X} \\ \text{s.t. } f_j(x) &\leq 0, \quad j = 1, \dots, m \end{aligned}$$

Пусть  $x^* \in X$  – минимум в задаче выше, а функции  $f_j, j = 0, 1, \dots, m$ , непрерывны в точке  $x^*$ . Тогда существуют числа  $\lambda_j \geq 0, j = 0, 1, \dots, m$ , такие, что

$$\begin{aligned} \sum_{j=0}^m \lambda_j &= 1, \\ \lambda_j f_j(x^*) &= 0, \quad j = 1, \dots, m, \\ 0 &\in \sum_{j=0}^m \lambda_j \partial f_j(x^*). \end{aligned}$$

**Доказательство.**

1. Рассмотрим функцию

$$f(x) = \max\{f_0(x) - f_0(x^*), f_1(x), \dots, f_m(x)\}.$$

Точка  $x^*$  является глобальным минимумом этой функции. Действительно, если бы в некоторой точке  $x_e \in X$  выполнялось неравенство  $f(x_e) < 0$ , то из этого следовало бы, что  $f_0(x_e) < f_0(x^*)$  и  $f_j(x_e) < 0, j = 1, \dots, m$ , что противоречит минимальности  $x^*$  в задаче выше.

2. Тогда из теоремы Ферма в субдифференциальной форме следует, что

$$0 \in \partial f(x^*).$$

3. По теореме Дубовицкого-Милютина имеем

$$\partial f(x^*) = \text{conv} \left( \bigcup_{j \in I} \partial f_j(x^*) \right),$$

4. Поэтому существует  $g_j \in \partial f_j(x^*)$ ,  $j \in I$ , такой, что

$$\sum_{j \in I} \lambda_j g_j = 0, \quad \sum_{j \in I} \lambda_j = 1, \quad \lambda_j \geq 0, \quad j \in I.$$

Осталось задать  $\lambda_j = 0$  для  $j \notin I$ .

**i** Для задачи математического программирования в общем виде

$$\begin{aligned} f_0(x) &\rightarrow \min_{x \in \mathbb{R}^n} \\ \text{s.t. } f_i(x) &\leq 0, \quad i = 1, \dots, m \\ h_i(x) &= 0, \quad i = 1, \dots, p \end{aligned}$$

можно сформулировать лагранжиан:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$

Пусть  $x^*$ ,  $(\lambda^*, \nu^*)$  - решение задачи математического программирования с нулевым зазором двойственности (оптимальное значение для прямой задачи  $p^*$  равно оптимальному значению для двойственной задачи  $d^*$ ). Пусть также функции  $f_0, f_i, h_i$  дифференцируемы.

- $\nabla_x L(x^*, \lambda^*, \nu^*) = 0$
- $\nabla_{\nu} L(x^*, \lambda^*, \nu^*) = 0$
- $\lambda_i^* \geq 0, i = 1, \dots, m$
- $\lambda_i^* f_i(x^*) = 0, i = 1, \dots, m$
- $f_i(x^*) \leq 0, i = 1, \dots, m$

### Некоторые условия регулярности

Эти условия необходимы для того, чтобы сделать условия ККТ необходимыми. Некоторые из них даже превращают необходимые условия в достаточные (например, условия Слейтера). Более того, если у вас есть регулярность, вы можете записать необходимые условия второго порядка  $\langle y, \nabla_{xx}^2 L(x^*, \lambda^*, \nu^*) y \rangle \geq 0$  с *полуопределенным* гессианом лагранжиана.

- **Условие Слейтера.** Если для выпуклой задачи (т.е. в предположении минимизации,  $f_0, f_i$  выпуклы и  $h_i$  аффинны) существует точка  $x$  такая, что  $h_i(x) = 0$  и  $f_i(x) < 0$  (существование строго внутренней точки бюджетного множества), то мы имеем нулевой зазор двойственности и условия ККТ становятся необходимыми и достаточными.
- **Условие линейности ограничений.** Если  $f_i$  и  $h_i$  - аффинные функции, то других условий не требуется.
- **Условие линейной независимости ограничений.** Градиенты активных ограничений неравенств и градиенты ограничений равенств линейно независимы в точке  $x^*$ .
- Другие примеры см. в [wiki](#).

12. Формулировка симплекс метода для задачи линейного программирования в стандартной форме. Теорема о проверке оптимальности решения.

**i** Если все элементы  $\lambda_B$  неположительны и базис  $B$  допустимый, тогда базис  $B$  оптимален. Здесь  $\lambda_B$  это коэффициенты при разложении  $c$  по базису  $B$ :  $\lambda_B^T A_B = c^T \Rightarrow \lambda_B^T = c^T A_B^{-1}$ .

## Формулировка симплекс метода для задачи линейного программирования в стандартной форме. Теорема о проверке оптимальности решения

Задача линейного программирования:

Пусть  $c \in \mathbb{R}^n$ ,  $b \in \mathbb{R}^m$ ,  $A \in \mathbb{R}^{m \times n}$ , тогда задача формулируется так:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \end{aligned}$$

### Идейное описание симплекс метода:

1. Убедится, что точка, в которой мы находимся, является угловой
2. Проверить оптимальность точки
3. Если необходимо, сменить угол (то есть сменить базис)
4. Повторять до схождения

Шаги выполнения симплекс метода:

#### 1. Поиск начальной базисной допустимой точки:

- Выберем начальную базисную (она является решением системы  $A_B x = b_B$ , где  $B$  - базис размера  $n$  пространства, а матрица  $A$  обычно имеет больше  $n$  ограничений) допустимую ( $Ax_0 \leq b$ ) точку  $x_0$  (искать ее будем через двухфазный симплексметод). Если такая точка не найдена, задача не имеет допустимого решения.

#### 2. Проверка оптимальности:

- Разложение вектора  $c$  в данном базисе  $B$  с коэффициентами  $\lambda_B$ :

$$\lambda_B^T A_B = c^T \quad \text{или} \quad \lambda_B^T = c^T A_B^{-1}$$

- Если все компоненты  $\lambda_B$  неположительны, текущий базис является оптимальным. Иначе далее меняем вершину симплекса.

#### 3. Определение переменной для удаления из базиса:

- Если в разложении  $\lambda_B$  есть положительные координаты, продолжаем оптимизацию. Пусть  $\lambda_B^k > 0$ . Необходимо исключить  $k$  из базиса. Рассчитаем направляющий вектор  $d$ , идя вдоль которого изменим вершину следующим образом: во-первых, для векторов всех ограничений из базиса, которые мы оставляем, направление должно быть им ортогонально, и, во-вторых, вдоль него значение, связанное с нашим ограничением, должно убывать:

$$\begin{cases} A_{B \setminus \{k\}} d = 0 \\ a_k^T d < 0 \end{cases}$$

#### 4. Вычисление шага вдоль выбранного направления $d$ :

- Для всех  $j \notin B$  считаем шаг:

$$\mu_j = \frac{b_j - a_j^T x_B}{a_j^T d}$$

- Новая вершина, которую добавим в базис:

$$t = \arg \min_j \{ \mu_j \mid \mu_j > 0 \}$$

#### 5. Обновление базиса:



- Обновляем базис и текущее решение:

$$B' = B \setminus \{k\} \cup \{t\},$$

$$x_{B'} = x_B + \mu_t d = A_{B'}^{-1} b_{B'}$$

- Изменение базиса приводит к уменьшению значения целевой функции:

$$c^\top x_{B'} = c^\top (x_B + \mu_t d) = c^\top x_B + \mu_t c^\top d$$

## 6. Повторение:

- Далее повторяем шаги 2-5 до достижения оптимального решения или установления, что задача не имеет допустимого решения.

### Теорема о проверке оптимальности решения:

Если все элементы  $\lambda_B$  неположительны и базис  $B$  достижим, тогда базис  $B$  оптимален.

Здесь  $\lambda_B$  это коэффициенты при разложении  $c$  по базису  $B$ :  $\lambda_B^T A_B = c^T \Rightarrow \lambda_B^T = c^T A_B^{-1}$ .

### Доказательство:

Предположим противное (что этот базис не оптимален), пусть  $\exists x^* : Ax^* \leq b$  и при этом  $c^T x^* < c^T x_B$ . Так как для всей матрицы  $A$  и вектора  $b$  неравенство верно, то и для подматрицы оно верно:

$$A_B x^* \leq b_B$$

Так как все элементы  $\lambda_B$  неположительны, то домножим строки на соответствующие элементы и сложим:

$$\lambda_B^T A_B x^* \geq \lambda_B^T b_B$$

$$c^T x^* \geq \lambda_B^T b_B = \lambda_B^T A_B x_B = c^T x_B$$

Противоречие.

## 13. Доказательство работы теста корней.

**i** Пусть  $(r_k)_{k=m}^\infty$  последовательность неотрицательных чисел, сходящихся к нулю, и пусть  $\alpha := \limsup_{k \rightarrow \infty} r_k^{1/k}$ . (причем  $\alpha \geq 0$ .)

- (a) Если  $0 \leq \alpha < 1$ , то  $(r_k)_{k=m}^\infty$  сходится линейно с константой  $\alpha$ .
- (b) В частности, если  $\alpha = 0$ , то  $(r_k)_{k=m}^\infty$  сходится сверхлинейно.
- (c) Если  $\alpha = 1$ , то  $(r_k)_{k=m}^\infty$  сходится сублинейно.
- (d) Случай  $\alpha > 1$  невозможен.

### Доказательство.

1. Покажем, что если  $(r_k)_{k=m}^\infty$  сходится линейно с константой  $0 \leq \beta < 1$ , то обязательно  $\alpha \leq \beta$ .

Действительно, по определению константы линейной сходимости для любого  $\varepsilon > 0$ , удовлетворяющего условию  $\beta + \varepsilon < 1$ , существует  $C > 0$ , такое что  $r_k \leq C(\beta + \varepsilon)^k$  для всех  $k \geq m$ .

Отсюда  $r_k^{1/k} \leq C^{1/k}(\beta + \varepsilon)$  для всех  $k \geq m$ . Переходя к пределу при  $k \rightarrow \infty$  и используя  $C^{1/k} \rightarrow 1$ , получаем  $\alpha \leq \beta + \varepsilon$ . Учитывая произвольность  $\varepsilon$ , следует, что  $\alpha \leq \beta$ .

2. Таким образом, в случае  $\alpha = 1$  последовательность  $(r_k)_{k=m}^\infty$  не может иметь линейную сходимость согласно вышеуказанному результату (доказанному от противного). Поскольку, тем не менее,  $(r_k)_{k=m}^\infty$  сходится к нулю, она должна сходиться сублинейно.

3. Теперь рассмотрим случай  $0 \leq \alpha < 1$ . Пусть  $\varepsilon > 0$  — произвольное число, такое что  $\alpha + \varepsilon < 1$ .

Согласно свойствам  $\limsup$ , существует  $N \geq m$ , такое что  $r_k^{1/k} \leq \alpha + \varepsilon$  для всех  $k \geq N$ .

Следовательно,  $r_k \leq (\alpha + \varepsilon)^k$  для всех  $k \geq N$ . Поэтому последовательность  $(r_k)_{k=m}^\infty$  сходится линейно с параметром  $\alpha + \varepsilon$  (не имеет значения, что неравенство выполняется только начиная с числа  $N$ ).

Из произвольности  $\varepsilon$  следует, что константа линейной сходимости последовательности  $(r_k)_{k=m}^\infty$  не превышает  $\alpha$ .

Поскольку, как показано выше, константа линейной сходимости не может быть меньше  $\alpha$ , это означает, что константа линейной сходимости последовательности  $(r_k)_{k=m}^\infty$  равна именно  $\alpha$ .

4. Покажем, что случай  $\alpha > 1$  невозможен.

Действительно, предположим, что  $\alpha > 1$ . Тогда из определения  $\limsup$  следует, что для любого  $N \geq m$  существует  $k \geq N$ , такое что  $r_k^{1/k} \geq 1$ , и, в частности,  $r_k \geq 1$ .

Но это означает, что у  $r_k$  есть подпоследовательность, ограниченная от нуля. Следовательно, последовательность  $(r_k)_{k=m}^\infty$  не может сходиться к нулю, что противоречит условию.

14. Метод дихотомии и золотого сечения для унимодальных функций. Скорость сходимости.

**i** Методы локализации решения для скалярной минимизации. Сходятся линейно.

### Метод дихотомии

Решаем следующую задачу:

$$f(x) \rightarrow \min_{x \in [a,b]}$$

где  $f(x)$  — унимодальная функция.

Мы хотим на каждом шаге вдвое сокращать область, в которой ищем минимум. Для этого будем пользоваться основным свойством унимодальных функций:

$$\forall a \leq x_1 < x_2 \leq b :$$

$$f(x_1) \leq f(x_2) \Rightarrow x_* \in [a, x_2]$$

$$f(x_1) \geq f(x_2) \Rightarrow x_* \in [x_1, b]$$

где  $x_*$  — точка, в которой достигается минимум

Алгоритм:

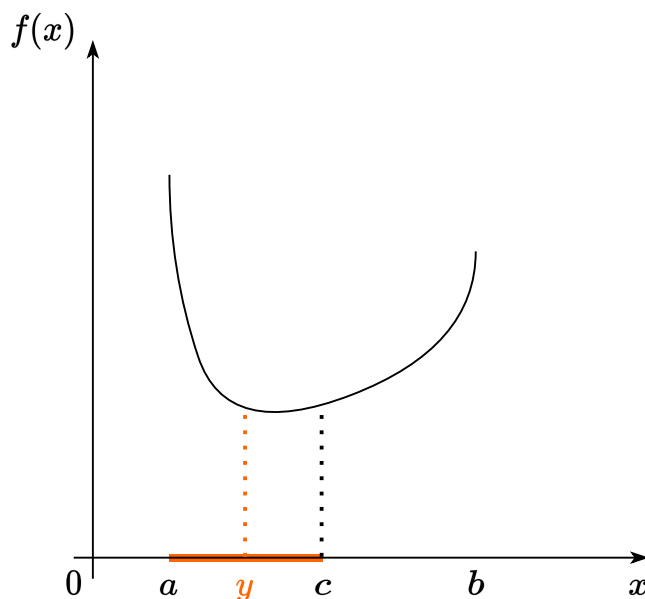


Рисунок 19: Алгоритм дихотомии

Можно заметить, что на каждой итерации требуется не более 2-х вычислений значения функции.

#### Сходимость метода дихотомии

Длина отрезка на  $k$ -ой итерации:

$$\Delta_k = b_k - a_k = \frac{1}{2^k}(b - a)$$

Если будем выбирать середину отрезка как выход  $k$ -ой итерации:

$$|x_k - x_*| \leq \frac{\Delta_k}{2}$$

Подставим полученное ранее выражение для длины отрезка:

$$|x_k - x_*| \leq \frac{1}{2^{k+1}}(b - a)$$

$$|x_k - x_*| \leq (0.5)^{k+1}(b - a)$$

Получили выражение для сходимости по итерациям. Отсюда также можно выразить необходимое количество итераций для достижения точности  $\varepsilon$ :

$$K = \left\lceil \log_2 \frac{b - a}{\varepsilon} - 1 \right\rceil$$

Теперь получим выражение для сходимости по количеству вычислений значения функции. Знаем, что на каждой итерации вычисляем значение не более 2-х раз, значит количество вычислений значения функции возьмём  $N = 2k$ :

$$|x_k - x_*| \leq (0.5)^{\frac{N}{2}+1}(b-a)$$

$$|x_k - x_*| \leq (0.707)^N \frac{b-a}{2}$$

### Метод золотого сечения

Идея такая же, как и в методе дихотомии, но хотим уменьшить количество вычислений значения функции. Для этого будем вычислять значения в точках золотого сечения. Так на каждой итерации нам нужно будет вычислять значение только в одной точке, так как для нового отрезка в одной из точек золотого сечения значение будет уже посчитано:

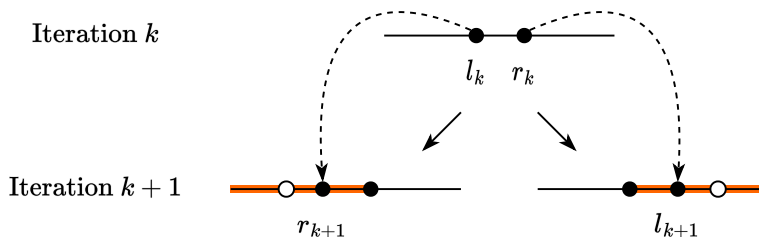


Рисунок 20: Золотое сечение

Алгоритм:

```
def golden_search(f, a, b, epsilon):
    tau = (sqrt(5) + 1) / 2
    y = a + (b - a) / tau**2
    z = a + (b - a) / tau
    while b - a > epsilon:
        if f(y) <= f(z):
            b = z
            z = y
            y = a + (b - a) / tau**2
        else:
            a = y
            y = z
            z = a + (b - a) / tau
    return (a + b) / 2
```

### Сходимость метода золотого сечения

На каждой итерации длина отрезка будет уменьшаться в  $\tau = \frac{\sqrt{5}+1}{2}$  раз. Тогда оценка сходимости (и по итерациям, и по вычислениям значений функции):

$$|x_k - x_*| \leq \frac{b_k - a_k}{2} = \left(\frac{1}{\tau}\right)^N \frac{b-a}{2} \approx 0.618^k \frac{b-a}{2}$$

Получили сходимость по итерациям хуже, чем у дихотомии, так как отрезки уменьшаются слабее на каждой итерации. Но по количеству вычислений значения функции сходимость у метода золотого сечения быстрее.

## 15. Теорема сходимости градиентного спуска для гладких выпуклых функций.

**i** Рассматриваем задачу

$$f(x) \rightarrow \min_{x \in \mathbb{R}^d}$$

и предполагаем, что  $f$  - выпуклая,  $L$ -гладкая,  $L > 0$ .

Пусть  $(x_k)_{k \in \mathbb{N}}$  это последовательность, созданная градиентным спуском с постоянным шагом  $\alpha$ ,  $0 < \alpha \leq \frac{1}{L}$ . Тогда градиентный спуск сходится сублинейно, то есть:

$$f(x_k) - f^* \leq \frac{\|x_0 - x^*\|^2}{2\alpha k}.$$

1. Формулировка метода:

$$x_{k+1} = x_k - \alpha \nabla f(x_k) \Rightarrow x_{k+1} - x_k = -\alpha \nabla f(x_k)$$

2.  $L$ -гладкость:  $\forall x, y : f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|^2$

$$y := x_{k+1}, x := x_k \Rightarrow f(x_{k+1}) \leq f(x_k) + \langle \nabla f(x_k), -\alpha \nabla f(x_k) \rangle + \frac{L}{2} \alpha^2 \|\nabla f(x_k)\|^2$$

$$f(x_{k+1}) \leq f(x_k) - \alpha \|\nabla f(x_k)\|^2 + \frac{L}{2} \alpha^2 \|\nabla f(x_k)\|^2 \quad (1)$$

3. Решим задачу оптимизации для поиска оптимального постоянного шага  $(\frac{L}{2} \alpha^2 - \alpha) \rightarrow \min_{\alpha}$ .

Получаем оптимальный шаг:  $\alpha = \frac{1}{L}$  и  $f(x_k) - f(x_{k+1}) \geq \frac{1}{2L} \|\nabla f(x_k)\|^2$

4. Выпуклость:  $f(y) \geq f(x) + \nabla f(x)^T (y - x)$

$$\begin{aligned} y := x^*, x := x_k &\Rightarrow f(x^*) \geq f(x_k) + \nabla f(x_k)^T (x^* - x_k) \Rightarrow \\ &\Rightarrow f(x_k) \leq f(x^*) + \nabla f(x_k)^T (x_k - x^*) \Rightarrow f(x_k) - f(x^*) \leq \nabla f(x_k)^T (x_k - x^*) \end{aligned}$$

5. Подставим выпуклость в (1):

$$\begin{aligned} f(x_{k+1}) &\leq f(x_k) - \frac{\alpha}{2} \|\nabla f(x_k)\|^2 \leq f^* + \langle \nabla f(x_k), x_k - x^* \rangle - \frac{\alpha}{2} \|\nabla f(x_k)\|^2 \\ &= f^* + \langle \nabla f(x_k), x_k - x^* - \frac{\alpha}{2} \nabla f(x_k) \rangle \\ &= f^* + \frac{1}{2\alpha} \left\langle \alpha \nabla f(x_k), 2 \left( x_k - x^* - \frac{\alpha}{2} \nabla f(x_k) \right) \right\rangle \end{aligned}$$

6. Пусть  $a = x_k - x^*$  и  $b = x_k - x^* - \alpha \nabla f(x_k)$ . Тогда  $a - b = \alpha \nabla f(x_k)$  и  $a + b = 2(x_k - x^* - \frac{\alpha}{2} \nabla f(x_k))$ .

$$\begin{aligned} f(x_{k+1}) &\leq f^* + \frac{1}{2\alpha} [\|x_k - x^*\|_2^2 - \|x_k - x^* - \alpha \nabla f(x_k)\|_2^2] \\ &\leq f^* + \frac{1}{2\alpha} [\|x_k - x^*\|_2^2 - \|x_{k+1} - x^*\|_2^2] \\ 2\alpha (f(x_{k+1}) - f^*) &\leq \|x_k - x^*\|_2^2 - \|x_{k+1} - x^*\|_2^2 \end{aligned}$$

7. Предположим, что последняя строка определена для некоторого индекса  $i$  и мы суммируем по  $i \in [0, k-1]$ . Большинство слагаемых будут равны нулю из-за телескопической природы суммы:

$$2\alpha \sum_{i=0}^{k-1} (f(x_{i+1}) - f^*) \leq \|x_0 - x^*\|_2^2 - \|x_k - x^*\|_2^2 \leq \|x_0 - x^*\|_2^2 \quad (1)$$

8. Из-за монотонного убывания на каждой итерации  $f(x_{i+1}) < f(x_i)$ :

$$kf(x_k) \leq \sum_{i=0}^{k-1} f(x_{i+1})$$

9. Подставим в (1):

$$\begin{aligned} 2\alpha kf(x_k) - 2\alpha kf^* &\leq 2\alpha \sum_{i=0}^{k-1} (f(x_{i+1}) - f^*) \leq \|x_0 - x^*\|_2^2 \\ f(x_k) - f^* &\leq \frac{\|x_0 - x^*\|_2^2}{2\alpha k} \leq \frac{L\|x_0 - x^*\|_2^2}{2k} \end{aligned}$$

То есть сходимость сублинейная.

16. Теорема сходимости градиентного спуска для гладких PL функций.

**i** Рассмотрим задачу

$$f(x) \rightarrow \min_{x \in \mathbb{R}^d}$$

и предположим, что  $f$  удовлетворяет условию Поляка-Лоясиевича с константой  $\mu$  и  $L$ -гладкости, для некоторых  $L \geq \mu > 0$ . Пусть  $(x_k)_{k \in \mathbb{N}}$  - последовательность, созданная градиентным спуском с постоянным шагом  $\alpha$ ,  $0 < \alpha \leq \frac{1}{L}$ . Тогда имеется линейная сходимость:

$$f(x_k) - f^* \leq (1 - \alpha\mu)^k (f(x_0) - f^*).$$

1. Используя  $L$ -гладкость, вместе с правилом обновления алгоритма, можно записать:

$$\begin{aligned} f(x_{k+1}) &\leq f(x_k) + \langle \nabla f(x_k), x_{k+1} - x_k \rangle + \frac{L}{2} \|x_{k+1} - x_k\|^2 \\ &= f(x_k) - \alpha \|\nabla f(x_k)\|^2 + \frac{L\alpha^2}{2} \|\nabla f(x_k)\|^2 \\ &= f(x_k) - \frac{\alpha}{2} (2 - L\alpha) \|\nabla f(x_k)\|^2 \\ &\leq f(x_k) - \frac{\alpha}{2} \|\nabla f(x_k)\|^2, \end{aligned}$$

В последнем неравенстве использовали предположение о шаге  $0 < \alpha L \leq 1$ .

2. Используя свойство Поляка-Лоясиевича, можно записать:

$$f(x_{k+1}) \leq f(x_k) - \alpha\mu(f(x_k) - f^*).$$

3. Вычитая  $f^*$  с обеих сторон и используя рекурсию, получаем:

$$f(x_k) - f^* \leq (1 - \alpha\mu)^k (f(x_0) - f^*).$$

17. Теорема сходимости градиентного спуска для сильно выпуклых квадратичных функций. Оптимальные гиперпараметры.

i

$$f(x) \rightarrow \min_{x \in \mathbb{R}^d}$$

$$f(x) = \frac{1}{2}x^T A x - b^T x + c, \quad A \in \mathbb{S}_{++}$$

Тогда градиентный спуск с шагом  $\alpha = \frac{2}{\mu+L}$  сходится линейно с показателем  $\frac{L-\mu}{L+\mu}$

$$f(x_k) - f^* \leq \left( \frac{L-\mu}{L+\mu} \right)^k (f(x_0) - f^*).$$

$$\nabla f(x) = Ax - b \stackrel{\nabla f(x^*)=0}{\Rightarrow} Ax^* = b$$

Тогда шаг градиентного спуска имеет вид

$$x_{k+1} = x_k - \alpha(Ax - b)$$

Найдем  $\alpha^*$ . Воспользуемся  $A = Q\Lambda Q^T$ , где  $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ ,  $Q = \|q_1, \dots, q_n\|$ ,  $\lambda_i, q_i$  - собственное значение и собственный вектор соответственно.

$$x_{k+1} = (I - \alpha A)x_k + \alpha Ax^* \mid - x^*$$

$$x_{k+1} - x^* = (I - \alpha A)(x_k - x^*)$$

$$x_{k+1} - x^* = (I - \alpha Q\Lambda Q^T)(x_k - x^*) \mid \cdot Q^T$$

$$Q^T(x_{k+1} - x^*) = (Q^T - \alpha\Lambda Q^T)(x_k - x^*) = (I - \alpha\Lambda)Q^T(x_k - x^*)$$

$$\text{Замена: } \tilde{x} = Q^T(x - x^*) \Rightarrow \tilde{x}_{k+1} = (I - \alpha\Lambda)\tilde{x}_k \Leftrightarrow \tilde{x}_i^{(k+1)} = (1 - \alpha\lambda_i)\tilde{x}_i^{(k)} \quad i = \overline{1, d}$$

$$\lambda_{\min} = \mu, \quad \lambda_{\max} = L$$

Сходимость есть  $\Leftrightarrow \max_i |1 - \alpha\lambda_i| < 1$

$$\left\{ \begin{array}{l} |1 - \lambda\mu| < 1 \Rightarrow 1 - \lambda\mu < 1 \Rightarrow \alpha > 0 \\ \phantom{|1 - \lambda\mu| < 1 \Rightarrow} \alpha\mu - 1 < 1 \Rightarrow \alpha < \frac{2}{\mu} \\ |1 - \alpha L| < 1 \Rightarrow 1 - \alpha L < 1 \Rightarrow \alpha > 0 \\ \phantom{|1 - \lambda\mu| < 1 \Rightarrow} \alpha L - 1 < 1 \Rightarrow \alpha < \frac{2}{L} \end{array} \right\} \Rightarrow \alpha < \frac{2}{L}$$

Радиус сходимости  $\rho = \max(|1 - \alpha\mu|, |1 - \alpha L|)$  и  $\rho \rightarrow \min \Leftrightarrow \alpha^* L - 1 = 1 - \alpha^* \mu \Rightarrow \alpha^* = \frac{2}{\mu+L}$  и  $\rho^* = \frac{L-\mu}{L+\mu}$

Итого получаем, что для градиентного спуска выполняется  $\|x_k - x^*\| \leq \left( \frac{L-\mu}{L+\mu} \right)^k \|x_0 - x^*\|$ .

18. Вывод ускоренного метода для квадратичной функции с помощью полиномов Чебышёва.

**i** Решаем квадратичную задачу:

$$f(x) = \frac{1}{2}x^T A x - b^T x \quad x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

Можно показать, что метод, полученный с помощью полиномов Чебышёва, имеет вид:

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) + \beta_k (x_k - x_{k-1})$$

обладает ускоренной линейной сходимостью.

1. Пусть  $x^*$  - единственное решение линейной системы  $Ax = b$  и пусть  $e_k = x_k - x^*$ , где  $x_{k+1} = x_k - \alpha_k (Ax_k - b)$  определяется рекурсивно, начиная с некоторого  $x_0$ , и  $\alpha_k$  - шаг, который мы определим позже.

$$e_{k+1} = (I - \alpha_k A) e_k.$$

1. Вышеуказанный расчет дает нам  $e_k = p_k(A) e_0$ , где  $p_k$  - полином

$$p_k(a) = \prod_{i=1}^k (1 - \alpha_i a).$$

Мы можем ограничить сверху норму ошибки как

$$\|e_k\| \leq \|p_k(A)\| \cdot \|e_0\|.$$

Поскольку  $A$  - симметричная матрица с собственными значениями в  $[\mu, L]$ ,

$$\|p_k(A)\| \leq \max_{\mu \leq a \leq L} |p_k(a)|.$$

Это приводит к интересной проблеме: среди всех полиномов, удовлетворяющих  $p_k(0) = 1$ , мы ищем полином, величина которого наименьшая в интервале  $[\mu, L]$ .

2. Наивный подход состоит в выборе равномерного шага  $\alpha_k = \frac{2}{\mu+L}$  в выражении. Этот выбор делает  $|p_k(\mu)| = |p_k(L)|$ .

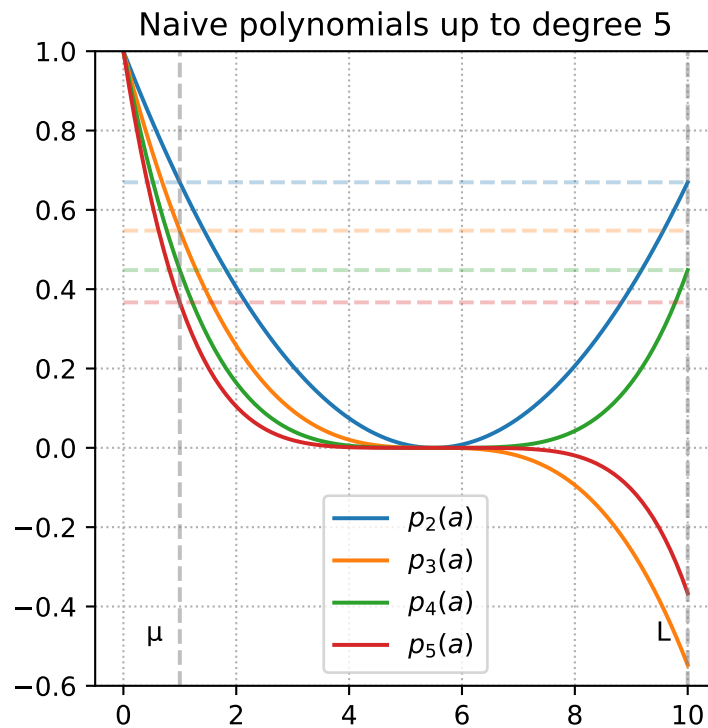
$$\|e_k\| \leq \left( \frac{L - \mu}{L + \mu} \right)^k \|e_0\|$$

Это точно такой же результат, который мы доказали для сходимости градиентного спуска в случае квадратичной функции.

Давайте взглянем на этот полином поближе. На правом рисунке мы выбрали  $\alpha = 1$  и  $\beta = 10$  так, что  $\kappa = 10$ . Соответствующий интервал, таким образом, равен  $[1, 10]$ .

Можем ли мы сделать лучше? Ответ - да.





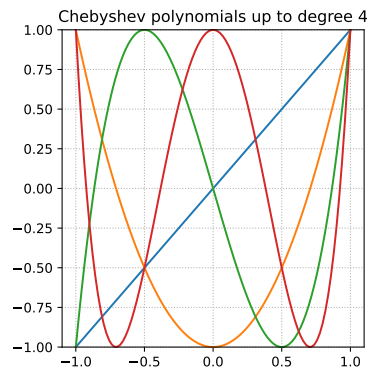
3. Полиномы Чебышёва оказываются оптимальным ответом на вопрос, который мы задавали. Соответствующим образом масштабированные, они минимизируют абсолютное значение в желаемом интервале  $[\mu, L]$  при условии, что значение равно 1 в начале.

$$T_0(x) = 1$$

$$T_1(x) = x$$

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x), \quad k \geq 2.$$

4. Давайте построим стандартные полиномы Чебышёва (без масштабирования):



5. Исходные полиномы Чебышёва определяются на интервале  $[-1, 1]$ . Чтобы использовать их для наших целей, нам нужно их масштабировать на интервал  $[\mu, L]$ .
6. Мы будем использовать следующую аффинную трансформацию:

$$x = \frac{L + \mu - 2a}{L - \mu}, \quad a \in [\mu, L], \quad x \in [-1, 1].$$

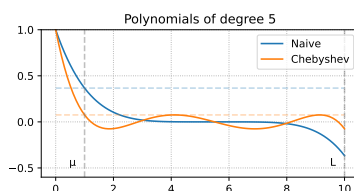
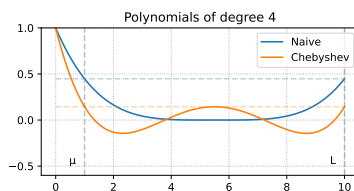
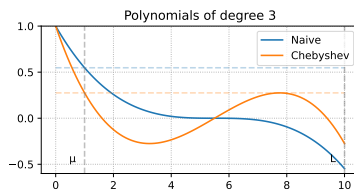
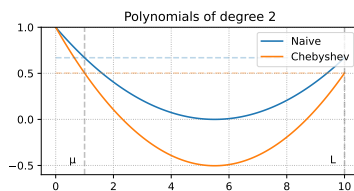
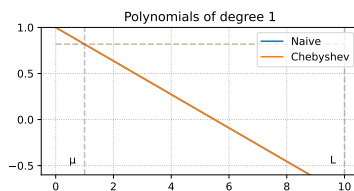
7. Обратите внимание, что  $x = 1$  соответствует  $a = \mu$ ,  $x = -1$  соответствует  $a = L$  и  $x = 0$  соответствует  $a = \frac{\mu+L}{2}$ . Эта трансформация гарантирует, что поведение полинома Чебышёва на интервале  $[-1, 1]$  отражается в интервал  $[\mu, L]$
8. В нашем анализе ошибок мы требуем, чтобы полином был равен 1 в 0 (т.е.,  $p_k(0) = 1$ ). После применения трансформации значение  $T_k$  в точке, соответствующей  $a = 0$ , может не быть 1. Таким образом, мы нормируем полином  $T_k$ , деля его на значение  $T_k\left(\frac{L+\mu}{L-\mu}\right)$ :

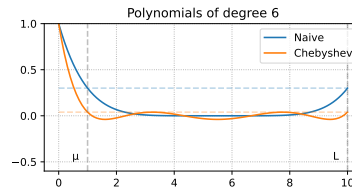
$$\frac{L + \mu}{L - \mu}, \quad \text{гарантируя, что} \quad P_k(0) = T_k\left(\frac{L + \mu - 0}{L - \mu}\right) \cdot T_k\left(\frac{L + \mu}{L - \mu}\right)^{-1} = 1.$$

9. Давайте построим масштабированные полиномы Чебышёва

$$P_k(a) = T_k\left(\frac{L + \mu - 2a}{L - \mu}\right) \cdot T_k\left(\frac{L + \mu}{L - \mu}\right)^{-1}$$

и наблюдаем, что они значительно лучше ведут себя в интервале  $[\mu, L]$  по сравнению с наивными полиномами.





10. Мы видим, что максимальное значение полинома Чебышёва на интервале  $[\mu, L]$  достигается в точке  $a = \mu$ . Следовательно, мы можем использовать следующую верхнюю границу:

$$\|P_k(A)\|_2 \leq P_k(\mu) = T_k\left(\frac{L + \mu - 2\mu}{L - \mu}\right) \cdot T_k\left(\frac{L + \mu}{L - \mu}\right)^{-1} = T_k(1) \cdot T_k\left(\frac{L + \mu}{L - \mu}\right)^{-1} = T_k\left(\frac{L + \mu}{L - \mu}\right)^{-1}$$

11. Используя определение числа обусловленности  $\kappa = \frac{L}{\mu}$ , мы получаем:

$$\|P_k(A)\|_2 \leq T_k\left(\frac{\kappa + 1}{\kappa - 1}\right)^{-1} = T_k\left(1 + \frac{2}{\kappa - 1}\right)^{-1} = T_k(1 + \epsilon)^{-1}, \quad \epsilon = \frac{2}{\kappa - 1}.$$

12. Следовательно, нам нужно только понять значение  $T_k$  в  $1 + \epsilon$ . Это то, откуда берется ускорение. Мы будем ограничивать это значение сверху величиной  $\mathcal{O}\left(\frac{1}{\sqrt{\epsilon}}\right)$ .

13. Чтобы ограничить  $|P_k|$  сверху, нам нужно оценить  $|T_k(1 + \epsilon)|$  снизу.

14. Для любого  $x \geq 1$ , полином Чебышёва первого рода может быть записан как

$$\begin{aligned} T_k(x) &= \cosh(k \operatorname{arccosh}(x)) \\ T_k(1 + \epsilon) &= \cosh(k \operatorname{arccosh}(1 + \epsilon)). \end{aligned}$$

15. Помним, что:

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \quad \operatorname{arccosh}(x) = \ln(x + \sqrt{x^2 - 1}).$$

16. Пусть  $\phi = \operatorname{arccosh}(1 + \epsilon)$ ,

$$e^\phi = 1 + \epsilon + \sqrt{2\epsilon + \epsilon^2} \geq 1 + \sqrt{\epsilon}.$$

17. Следовательно,

$$\begin{aligned} T_k(1 + \epsilon) &= \cosh(k \operatorname{arccosh}(1 + \epsilon)) \\ &= \cosh(k\phi) \\ &= \frac{e^{k\phi} + e^{-k\phi}}{2} \geq \frac{e^{k\phi}}{2} \\ &= \frac{(1 + \sqrt{\epsilon})^k}{2}. \end{aligned}$$

18. Наконец, мы получаем:

$$\begin{aligned} \|e_k\| &\leq \|P_k(A)\| \|e_0\| \leq \frac{2}{(1 + \sqrt{\epsilon})^k} \|e_0\| \\ &\leq 2 \left(1 + \sqrt{\frac{2}{\kappa - 1}}\right)^{-k} \|e_0\| \\ &\leq 2 \exp\left(-\sqrt{\frac{2}{\kappa - 1}} k\right) \|e_0\| \end{aligned}$$

19. Из-за рекурсивного определения полиномов Чебышёва, мы непосредственно получаем итерационную схему ускорения. Переформулируя рекуррентное соотношение в терминах наших масштабированных полиномов Чебышёва, мы получаем:

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x)$$

Поскольку  $x = \frac{L+\mu-2a}{L-\mu}$ , и:

$$\begin{aligned} P_k(a) &= T_k\left(\frac{L+\mu-2a}{L-\mu}\right) T_k\left(\frac{L+\mu}{L-\mu}\right)^{-1} \\ T_k\left(\frac{L+\mu-2a}{L-\mu}\right) &= P_k(a) T_k\left(\frac{L+\mu}{L-\mu}\right) \\ T_{k-1}\left(\frac{L+\mu-2a}{L-\mu}\right) &= P_{k-1}(a) T_{k-1}\left(\frac{L+\mu}{L-\mu}\right) \\ T_{k+1}\left(\frac{L+\mu-2a}{L-\mu}\right) &= P_{k+1}(a) T_{k+1}\left(\frac{L+\mu}{L-\mu}\right) \\ P_{k+1}(a) t_{k+1} &= 2 \frac{L+\mu-2a}{L-\mu} P_k(a) t_k - P_{k-1}(a) t_{k-1}, \text{ where } t_k = T_k\left(\frac{L+\mu}{L-\mu}\right) \\ P_{k+1}(a) &= 2 \frac{L+\mu-2a}{L-\mu} P_k(a) \frac{t_k}{t_{k+1}} - P_{k-1}(a) \frac{t_{k-1}}{t_{k+1}} \end{aligned}$$

20. Поскольку мы имеем  $P_{k+1}(0) = P_k(0) = P_{k-1}(0) = 1$ , мы можем записать метод в следующей форме:

$$P_{k+1}(a) = (1 - \alpha_k a) P_k(a) + \beta_k (P_k(a) - P_{k-1}(a)).$$

21. Перегруппируя члены, мы получаем:

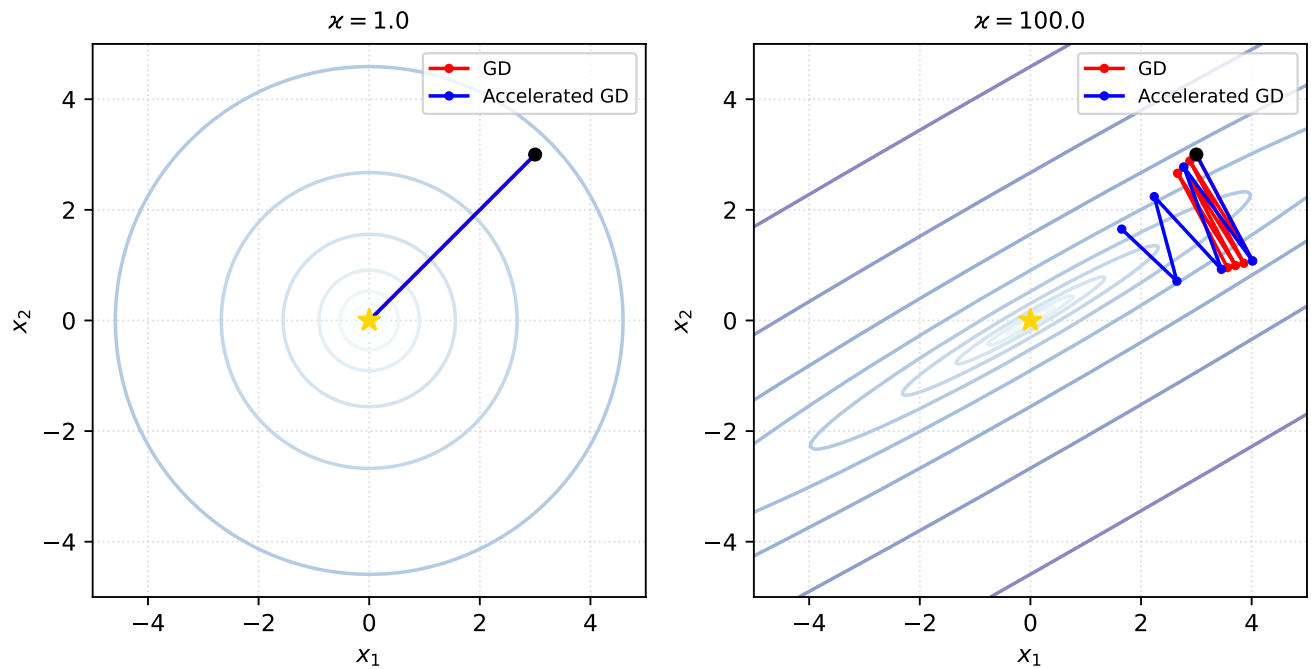
$$\begin{aligned} P_{k+1}(a) &= (1 + \beta_k) P_k(a) - \alpha_k a P_k(a) - \beta_k P_{k-1}(a), \\ P_{k+1}(a) &= 2 \frac{L+\mu}{L-\mu} \frac{t_k}{t_{k+1}} P_k(a) - \frac{4a}{L-\mu} \frac{t_k}{t_{k+1}} P_k(a) - \frac{t_{k-1}}{t_{k+1}} P_{k-1}(a) \\ &\begin{cases} \beta_k = \frac{t_{k-1}}{t_{k+1}}, \\ \alpha_k = \frac{4}{L-\mu} \frac{t_k}{t_{k+1}}, \\ 1 + \beta_k = 2 \frac{L+\mu}{L-\mu} \frac{t_k}{t_{k+1}} \end{cases} \end{aligned}$$

22. Мы почти закончили 😊. Помним, что  $e_{k+1} = P_{k+1}(A)e_0$ . Также отметим, что мы работаем с квадратичной задачей, поэтому мы можем предположить  $x^* = 0$  без потери общности. В этом случае  $e_0 = x_0$  и  $e_{k+1} = x_{k+1}$ .

$$\begin{aligned} x_{k+1} &= P_{k+1}(A)x_0 = (I - \alpha_k A)P_k(A)x_0 + \beta_k (P_k(A) - P_{k-1}(A))x_0 \\ &= (I - \alpha_k A)x_k + \beta_k (x_k - x_{k-1}) \end{aligned}$$

23. Для квадратичной задачи мы имеем  $\nabla f(x_k) = Ax_k$ , поэтому мы можем переписать обновление как:

$$\boxed{x_{k+1} = x_k - \alpha_k \nabla f(x_k) + \beta_k (x_k - x_{k-1})}$$



19. Доказательство сходимости метода сопряженных градиентов и вывод формул метода (В этом вопросе необходимо доказать за какое количество шагов сходится метод, как выбираются направления и почему в  $A$ -ортогонализации достаточно хранить лишь предыдущий шаг метода, а не все предыдущие).

i

$$\mathbf{r}_0 := \mathbf{b} - \mathbf{A}\mathbf{x}_0$$

Если  $\mathbf{r}_0$  достаточно мала, то вернуть  $\mathbf{x}_0$  как результат

$$\mathbf{d}_0 := \mathbf{r}_0$$

$$k := 0$$

повторить

$$\alpha_k := \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{d}_k^T \mathbf{A} \mathbf{d}_k}$$

$$\mathbf{x}_{k+1} := \mathbf{x}_k + \alpha_k \mathbf{d}_k$$

$$\mathbf{r}_{k+1} := \mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{d}_k$$

Если  $\mathbf{r}_{k+1}$  достаточно мала, то выйти из цикла

$$\beta_k := \frac{\mathbf{r}_{k+1}^T \mathbf{r}_{k+1}}{\mathbf{r}_k^T \mathbf{r}_k}$$

$$\mathbf{d}_{k+1} := \mathbf{r}_{k+1} + \beta_k \mathbf{d}_k$$

$$k := k + 1$$

конец повторения

вернуть  $\mathbf{x}_{k+1}$  как результат

Рассмотрим следующую квадратичную задачу оптимизации:

$$\min_{x \in \mathbb{R}^n} f(x) = \min_{x \in \mathbb{R}^n} \frac{1}{2} x^T A x - b^T x + c, \text{ где } A \in \mathbb{S}_{++}^n. \quad (2)$$

Берём в арсенал процесс Грам-Шмидта:

$$d_k = u_k + \sum_{i=0}^{k-1} \beta_{ik} d_i \quad \beta_{ik} = -\frac{\langle d_i, u_k \rangle}{\langle d_i, d_i \rangle} \quad (3)$$

**Лемма 1. Линейная независимость  $A$ -ортогональных векторов.**

Если множество векторов  $d_1, \dots, d_n$  - попарно  $A$ -ортогональны (каждая пара векторов  $A$ -ортогональна), то эти векторы линейно независимы.  $A \in \mathbb{S}_{++}^n$ .

**Доказательство**

Покажем, что если  $\sum_{i=1}^n \alpha_i d_i = 0$ , то все коэффициенты должны быть равны нулю:

$$\begin{aligned} 0 &= \sum_{i=1}^n \alpha_i d_i \\ \text{Умножаем на } d_j^T A \cdot &= d_j^T A \left( \sum_{i=1}^n \alpha_i d_i \right) = \sum_{i=1}^n \alpha_i d_j^T A d_i \\ &= \alpha_j d_j^T A d_j + 0 + \dots + 0 \end{aligned}$$

Таким образом,  $\alpha_j = 0$ , для всех остальных индексов нужно проделать тот же процесс

Введем следующие обозначения:

- $r_k = b - Ax_k$  - невязка
- $e_k = x_k - x^*$  - ошибка
- Поскольку  $Ax^* = b$ , имеем  $r_k = b - Ax_k = Ax^* - Ax_k = -A(x_k - x^*)$

$$r_k = -Ae_k. \quad (4)$$

- Также заметим, что поскольку  $x_{k+1} = x_0 + \sum_{i=1}^k \alpha_i d_i$ , имеем

$$e_{k+1} = e_0 + \sum_{i=1}^k \alpha_i d_i. \quad (5)$$

**Лемма 2. Сходимость метода сопряженных направлений.**

Предположим, мы решаем  $n$ -мерную квадратичную сильно выпуклую задачу оптимизации (2). Метод сопряженных направлений

$$x_{k+1} = x_0 + \sum_{i=0}^k \alpha_i d_i$$

с  $\alpha_i = \frac{\langle d_i, r_i \rangle}{\langle d_i, A d_i \rangle}$  взятым из точного линейного поиска, сходится за не более  $n$  шагов алгоритма.

**Доказательство**

1. Нужно доказать, что  $\delta_i = -\alpha_i$ :

$$e_0 = x_0 - x^* = \sum_{i=0}^{n-1} \delta_i d_i$$

2. Умножаем обе части слева на  $d_k^T A$ :

$$d_k^T A e_0 = \sum_{i=0}^{n-1} \delta_i d_k^T A d_i = \delta_k d_k^T A d_k$$

$$d_k^T A \left( e_0 + \sum_{i=0}^{k-1} \alpha_i d_i \right) = d_k^T A e_k = \delta_k d_k^T A d_k \quad (A - \text{ортогональность})$$

$$\delta_k = \frac{d_k^T A e_k}{d_k^T A d_k} = -\frac{d_k^T r_k}{d_k^T A d_k} \Leftrightarrow \delta_k = -\alpha_k$$

**Лемма 3. Разложение ошибки.**

$$e_i = \sum_{j=i}^{n-1} -\alpha_j d_j \quad (6)$$

**Доказательство**

По определению

$$e_i = e_0 + \sum_{j=0}^{i-1} \alpha_j d_j = x_0 - x^* + \sum_{j=0}^{i-1} \alpha_j d_j = -\sum_{j=0}^{n-1} \alpha_j d_j + \sum_{j=0}^{i-1} \alpha_j d_j = \sum_{j=i}^{n-1} -\alpha_j d_j$$

**Лемма 4. Невязка ортогональна всем предыдущим направлениям для CD.**

Рассмотрим невязку метода сопряженных направлений на  $k$  итерации  $r_k$ , тогда для любого  $i < k$ :

$$d_i^T r_k = 0 \quad (7)$$

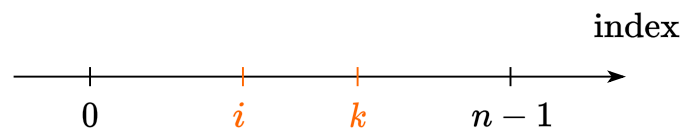
**Доказательство**

Запишем (6) для некоторого фиксированного индекса  $k$ :

$$e_k = \sum_{j=k}^{n-1} -\alpha_j d_j$$

Умножаем обе части на  $-d_i^T A$ .

$$-d_i^T A e_k = \sum_{j=k}^{n-1} \alpha_j d_i^T A d_j = 0$$



Таким образом,  $d_i^T r_k = 0$  и невязка  $r_k$  ортогональна всем предыдущим направлениям  $d_i$  для метода CD.

**Идея метода сопряженных градиентов (CG)**

- Это буквально метод сопряженных направлений, в котором мы выбираем специальный набор  $d_0, \dots, d_{n-1}$ , позволяющий значительно ускорить процесс Грама-Шмидта.
- Используется процесс Грама-Шмидта с  $A$ -ортогональностью вместо Евклидовой ортогональности, чтобы получить их из набора начальных векторов.
- На каждой итерации  $r_0, \dots, r_{n-1}$  используются в качестве начальных векторов для процесса Грама-Шмидта.
- Основная идея заключается в том, что для произвольного метода CD процесс Грама-Шмидта вычислительно дорогой и требует квадратичного числа операций сложения векторов и скалярных произведений  $\mathcal{O}(n^2)$ , в то время как в случае CG мы покажем, что сложность этой процедуры может быть уменьшена до линейной  $\mathcal{O}(n)$ .

CG = CD +  $r_0, \dots, r_{n-1}$  как начальные векторы для процесса Грама-Шмидта +  $A$ -ортогональность.

### Лемма 5. Невязки ортогональны друг другу в методе CG

Все невязки в методе CG ортогональны друг другу:

$$r_i^T r_k = 0 \quad \forall i \neq k \quad (8)$$

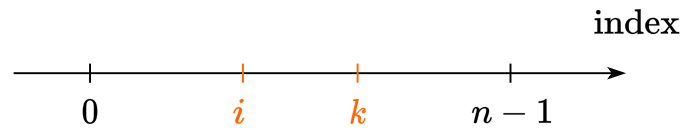
### Доказательство

Запишем процесс Грама-Шмидта (3) с  $\langle \cdot, \cdot \rangle$  замененным на  $\langle \cdot, \cdot \rangle_A = x^T A y$

$$d_i = u_i + \sum_{j=0}^{i-1} \beta_{ji} d_j \quad \beta_{ji} = -\frac{\langle d_j, u_i \rangle_A}{\langle d_j, d_j \rangle_A} \quad (9)$$

Тогда, мы используем невязки в качестве начальных векторов для процесса и  $u_i = r_i$ .

$$d_i = r_i + \sum_{j=0}^{i-1} \beta_{ji} d_j \quad \beta_{ji} = -\frac{\langle d_j, r_i \rangle_A}{\langle d_j, d_j \rangle_A} \quad (10)$$



Умножаем обе части (9) на  $r_k^T$  для некоторого индекса  $k$ :

$$r_k^T d_i = r_k^T u_i + \sum_{j=0}^{i-1} \beta_{ji} r_k^T d_j$$

Если  $j < i < k$ , то имеем лемму 4 с  $d_i^T r_k = 0$  и  $d_j^T r_k = 0$ . Имеем:

$$r_k^T u_i = 0 \quad \text{для CD} \quad r_k^T r_i = 0 \quad \text{для CG}$$

Более того, если  $k = i$ :

$$r_k^T d_k = r_k^T u_k + \sum_{j=0}^{k-1} \beta_{jk} r_k^T d_j = r_k^T u_k + 0,$$



и мы имеем для любого  $k$  (из-за произвольного выбора  $i$ ):

$$r_k^T d_k = r_k^T u_k. \quad (11)$$

#### Лемма 6. Пересчет невязки

$$\begin{aligned} r_{k+1} &= r_k - \alpha_k Ad_k \\ r_{k+1} &= -Ae_{k+1} = -A(e_k + \alpha_k d_k) = -Ae_k - \alpha_k Ad_k = r_k - \alpha_k Ad_k \end{aligned} \quad (12)$$

Наконец, все эти вышеуказанные леммы достаточны для доказательства, что  $\beta_{ji} = 0$  для всех  $i, j$ , кроме соседних.

#### Грам-Шмидт в методе CG

Рассмотрим процесс Грам-Шмидта в методе CG:

$$\beta_{ji} = -\frac{\langle d_j, u_i \rangle_A}{\langle d_j, d_j \rangle_A} = -\frac{d_j^T A u_i}{d_j^T A d_j} = -\frac{d_j^T A r_i}{d_j^T A d_j} = -\frac{r_i^T A d_j}{d_j^T A d_j}.$$

Рассмотрим скалярное произведение  $\langle r_i, r_{j+1} \rangle$  используя (12):

$$\begin{aligned} \langle r_i, r_{j+1} \rangle &= \langle r_i, r_j - \alpha_j A d_j \rangle = \langle r_i, r_j \rangle - \alpha_j \langle r_i, A d_j \rangle \\ \alpha_j \langle r_i, A d_j \rangle &= \langle r_i, r_j \rangle - \langle r_i, r_{j+1} \rangle \end{aligned}$$

- Если  $i = j$ :  $\alpha_i \langle r_i, A d_i \rangle = \langle r_i, r_i \rangle - \langle r_i, r_{i+1} \rangle = \langle r_i, r_i \rangle$ . Этот случай не интересен по построению процесса Грам-Шмидта.
- Соседний случай  $i = j + 1$ :  $\alpha_j \langle r_i, A d_j \rangle = \langle r_i, r_{i-1} \rangle - \langle r_i, r_i \rangle = -\langle r_i, r_i \rangle$
- Для любого другого случая:  $\alpha_j \langle r_i, A d_j \rangle = 0$ , потому что все невязки ортогональны друг другу.

Наконец, мы имеем формулу для  $i = j + 1$ :

$$\beta_{ji} = -\frac{r_i^T A d_j}{d_j^T A d_j} = \frac{1}{\alpha_j} \frac{\langle r_i, r_i \rangle}{d_j^T A d_j} = \frac{d_j^T A d_j}{d_j^T r_j} \frac{\langle r_i, r_i \rangle}{d_j^T A d_j} = \frac{\langle r_i, r_i \rangle}{\langle r_j, r_j \rangle} = \frac{\langle r_i, r_i \rangle}{\langle r_{i-1}, r_{i-1} \rangle}$$

И для направления  $d_{k+1} = r_{k+1} + \beta_{k,k+1} d_k$ ,  $\beta_{k,k+1} = \beta_k = \frac{\langle r_{k+1}, r_{k+1} \rangle}{\langle r_k, r_k \rangle}$ .