Вопросы к экзамену по методам оптимизации

Ilya Yaroshevskiy

13 мая 2023 г.

Содержание

1 1 Вариант

1. Какая функция называется целевой?

   Ответ: Целевая функция — функция которую надо оптимизировать (найти минимум)

2. Сформулировать два необходимых и достаточных дифференциальных условия выпуклости функций.

   Ответ: Если $\forall x: \ f''(x) \geq 0$, тогда $f(x)$ выпукла вниз.
   Если $\forall x: \ f''(x) \leq 0$, тогда $f(x)$ выпукла вверх.
   Это как достаточные, так и необходимые условия

3. Является ли условие $\frac{df}{dx}\Big|_{x=\bar{x}} = 0$ достаточным для того, чтобы число $\bar{x}$ было точкой минимума унимодальной, но невыпуклой функции $f(x)$?

   Ответ: Нет. Функция может не строго убывать, поэтому производная может обращаться в 0 не только в точке минимума.

4. Вычислить и нарисовать градиенты, а также вычислить матрицу Гессе функции $f(x) = x_1^2 - x_2^2$ в точках $x_1 = (1, 1)^T$ и $x_2 = (1, -1)^T$.

   Ответ:

$$\nabla f(x) = (2x_1 \quad -2x_2)$$

$$\forall x: \ H(x) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

$$\nabla f(x_1) = (2 \ \ -2) \quad \nabla f(x_2) = (2 \ 2)$$

5. Для функции $f(x) = 64x_1^2 + 126x_1x_2 + 64x_2^2 - 10x_1 + 30x_2 + 13$ выпишите квадратичную форму, линейную и постоянную часть.

Ответ:

$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} x^2 + (b_1\ b_2)x + c$$

$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} x + (b_1\ b_2)x + c$$

$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} x + (b_1\ b_2)x + c$$

$$f(x) = \frac{1}{2}(a_{11}x_1^2 + a_{12}x_2x_1 + a_{21}x_1x_2 + a_{22}x_2^2) + b_1x_1 + b_2x_2 + c$$

$$f(x) = \frac{1}{2}a_{11}x_1^2 + a_{12}x_2x_1 + \frac{1}{2}a_{22}x_2^2 + b_1x_1 + b_2x_2 + c$$

$$\frac{1}{2}a_{11} = 64 \implies a_{11} = 128$$

$$a_{12} = a_{21} = 126$$

$$\frac{1}{2}a_{22} = 64 \implies a_{22} = 128$$

$$b_1 = -10 \quad b_2 = 30 \quad c = 13$$

$$f(x) = \frac{1}{2}\begin{pmatrix} 128 & 126 \\ 126 & 128 \end{pmatrix} x^2 + (-10\ 30)x + 13$$

6. Дайте определение направления спуска. Какие методы многомерной минимизации функций называются методами спуска?

   Ответ: Вектор p^k — направление убывания(спуска) функции $f(x)$ в точке x^*, если при всех достаточно малых положительных α выполняется неравенство:

   $$f(x^k + \alpha p^k) < f(x^k)$$

   или по другому $\left\langle \nabla f(x^k), p^k \right\rangle < 0$
   Методами спуска называются методы для которых итерационная последовательность имеет вид:

   $$x^{k+1} = x^k + \alpha p^k$$

   То есть на каждой итерации обеспечивается движение в сторону наибольшего убывания функции.

7. Какова особенность траектории поиска для метода наискорейшего спуска?

   Ответ: В направлении убывания делается не маленький шаг, а такой, что в этом направлении достигается минимум функции. Значит два последовательных веткора спуска будут перпендикулярны друг другу, так как p^k будет касательным к линии уровня, а p^{k+1} нормалью к ней.

8. При каких условиях в итерационном процессе с наискорейшим спуском и в процессе с исчерпывающим спуском найденные точки x_k (k — номер итерации) совпадают, различаются?

   Ответ: Для исчерпывающего спуска точка α_k — стационарная точка функции $\varphi_k(\alpha)$, тогда как для наискорейшего спуска α_k — точка минимума функции $\varphi_k(\alpha)$. Т.е. точки в этих методах совпадают если α_k — стационарная и точка минимума для функции $\varphi_k(\alpha)$.

9. Каким условиям должна удовлетворять минимизируемая функция и начальные условия, чтобы метод градиентного спуска/метод наискорейшего спуска сошелся за одну итерацию?

   Ответ: Если функия имеет вид $f(x) = \frac{1}{2}\langle Ax, x\rangle$, а начальное приближение x^0 лежит на собственном векторе матрицы A.

10. Дайте определение метода сопряженных градиентов. Для каких матриц доказана теорема о сходимости метода сопряженных градиентов?

Ответ: Метод сопряженных градиентов находит точку минимума использую A-ортогональные направления спуска p^k и находит ее не более чем за n шагов. Метод сопряженных градиентов работает для квадратичных функций с положительно(отрицательно) определенной симметричной матрицей A.

11. Можно ли в процессе ортогонализации в методе сопряженных градиентов использовать произвольный базис?

    Ответ: Нет, т.к. требуется A-ортогнальность векторов p^k, т.е.:

$$\left\langle Ap^i, p^j \right\rangle = 0$$

12. Какое преимущество дает использование в процессе ортогонализации последовательности антиградиентов?

    Ответ: Использование системы антиградиентов возволяет в процессе ортогонализации также производить процесс спуска.

13. Опишите, вследствие чего возникают вычислительные ошибки в итерационном процессе метода сопряженных градиентов.

    Ответ: В общем случае точное вычисление α_k невозможно. Из-за накопления погрешности вектора p_k могут оказаться не A-ортогональными и не указывать на направление спуска.

14. Опишите метод Ньютона с одномерным поиском. Как находят шаг в методе Ньютона с одномерным поиском?

    Ответ: На k-той итерации:

    (a) Вычисляется градиент ∇f и Гессиан H функции в точке x^{k-1}
    (b) Для нахождения направления спуска решается СЛАУ: $Hp^k = -\nabla f$
    (c) С помощью алгоритма одномерного поиска находится $\alpha_k = \min f(x^{k-1} + \alpha p^k)$
    (d) Новое приближение: $x^k = x^{k-1} + \alpha_k p^k$
    (e) Проверка условия останова $\left\| x^k - x^{k-1} \right\| < \varepsilon$

15. Как вычисляется величина шага в методе Марквардта? Какого порядка метод Марквардта?

    Ответ: Величина и направление шага определяется вектором p^k, который находится из решения СЛАУ: $(H(x) + \tau I)p^k = -\nabla f(x)$, где нахождение τ определяется конкретой вариацией метода. Метод Марквардта использует вторые производные, поэтому это метод второго порядка.

16. Запишите уравнение коррекции для аппроксимации обратной матрицы Гессе в квазиньютоновских методах.

    Ответ:

$$G_{k+1} = G_k + \Delta G_k \quad k \in \mathbb{N}$$

    , где ΔG_k — положительно определнная матрица, которая наызвается поправочной, ее вычисление зависит от конекретного метода.

17. Погрешность, невязка, число обусловленности и их связь. Приведите примеры.

    Ответ: Пусть требуется решить СЛАУ $Ax = b$. Тогда **число обусловленности** $\operatorname{cond} A = \frac{L}{l}$, где l, L — соответсвенно минимальное и максимальное собственные значения матрицы A. Пусть x^* — точное решение СЛАУ, а x — приближенное. Тогда **величина ошибки(относительная погрешность)** $\frac{\|x^* - x\|}{\|x^*\|}$, а **невязка** $\Gamma = b - Ax$
    Справедливо следущее неравенство:

$$\frac{\|x^* - x\|}{\|x^*\|} \leq \operatorname{cond} A \cdot \frac{\|b - Ax\|}{\|b\|}$$

Пример.

$$\begin{pmatrix} 0.780 & 0.563 \\ 0.457 & 0.330 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.217 \\ 0.127 \end{pmatrix}$$

$$x^* = (1\ -1)^T \quad x = (1.71\ -1.98)^T$$

$$\Delta x = (-0.71\ 0.98)^T$$

$$\Gamma = (0.217\ 0.127)^T - (0.21906\ 0.12807)^T = (-0.00206\ -0.00107)^T$$

18. Дана матрица:

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 101 & -10 & -10 \\ 0 & -10 & 2 & 1 \\ 0 & -10 & 1 & 101 \end{bmatrix}$$

и вектор $f = (1, 0, 1, 0)^T$. Решить СЛАУ $Ax = f$ методом квадратного корня (Холесского).

Ответ:

$$l_{11} = \sqrt{a_1 1} = 1$$

$$l_{j1} = \frac{a_{j1}}{l_{11}} \quad j = \overline{2, n}$$

$$l_{ii} = \sqrt{a_{ii} - \sum_{p=1}^{i-1} l_{ip}^2} \quad i = \overline{2, n}$$

$$l_{ji} = \frac{1}{l_{ii}} \left(a_{ji} - \sum_{p=1}^{i-1} l_{ip} l_{jp} \right) \quad i = \overline{2, n-1}, i = \overline{i+1, n}$$

$$l_{21} = -1;\ l_{31} = 0;\ l_{41} = 0$$

$$l_{22} = \sqrt{101 - 1} = 10$$

$$l_{32} = \frac{-10 - (-1 \cdot 0)}{10} = -1$$

$$l_{33} = \sqrt{2 - (0^2 + 1^2)} = 1$$

$$l_{42} = \frac{-10 - (-1 \cdot 0)}{10} = -1;\ l_{43} = \frac{1 - (0 \cdot 0 + (-1) \cdot (-1))}{1} = 0$$

$$l_{44} = \sqrt{101 - (0^2 + 1^2 + 0^2)} = 10$$

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 10 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 10 \end{pmatrix}$$

Тогда решением уравнения $Ly = (1\ 0\ 1\ 0)^T$ будет вектор $y = (1\ 0.1\ 1.1\ 0.01)$, а решением $L^T x = y$ будет $x = (1.1201\ 0.1201\ 1.1\ 0.001)$

2 2 Вариант

1. Дать определение локального и глобального минимумов функции.

   Ответ: Для функции $f : E \subset \mathbb{R}^m \to \mathbb{R}$, точка $x^* \in E$ называется **локальным минимумом**, если $\exists U(x^*) :\ \forall x \in U \cap E\ f(x^*) \le f(x)$ и **глобальным минимумом**, если $\forall x \in E\ f(x^*) \le f(x)$

2. Сформулировать условие Липшица для функции $f(x)$ на отрезке $[a, b]$.

   Ответ: $f(x)$ удовлетворяет условию Липшица, если

$$\exists L\ \forall x_1, x_2 \in [a, b] :\ |f(x_1) - f(x_2)| \le L|x_1 - x_2|$$

3. Для каких выпуклых дважды дифференцируемых функций метод золотого сечения приводит к цели за меньшее количество итераций, чем метод Ньютона?

Ответ: Если функция плохо аппроксимируема первыми тремя членами своего разложения в Тейлора, то метод Ньютона будет сходиться долго. А метод золотого сечения не зависит от разложения в Тейлора и его количество итераций от этого не зависит. Т.е. это должна быть неквадратичная функция.

4. Классифицировать квадратичные формы и соответствующие им матрицы Гессе:

$$H(x) = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, H(x) = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, H(x) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, H(x) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Ответ:

- $$H(x) = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$
  Угловые миноры $H - 2, 3 \implies$ положительно определена.

- $$H(x) = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$
  Угловые миноры $H - 1, 0 \implies$ положительно полуопределена.

- $$H(x) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$
  Угловые миноры $H - 2, 3 \implies$ положительно определена.

- $$H(x) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
  Угловые миноры $H - 1, 0 \implies$ положительно полуопределена.

5. Для функции $f(x) = 254x_1^2 + 506x_1x_2 + 254x_2^2 + 50x_1 + 130x_2 - 111$ выписать квадратичную форму, линейную и постоянную часть.

Ответ:

$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} x^2 + (b_1\ b_2)x + c$$

$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} x + (b_1\ b_2)x + c$$

$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} x + (b_1\ b_2)x + c$$

$$f(x) = \frac{1}{2}(a_{11}x_1^2 + a_{12}x_2x_1 + a_{21}x_1x_2 + a_{22}x_2^2) + b_1x_1 + b_2x_2 + c$$

$$f(x) = \frac{1}{2}a_{11}x_1^2 + a_{12}x_2x_1 + \frac{1}{2}a_{22}x_2^2 + b_1x_1 + b_2x_2 + c$$

$$\frac{1}{2}a_{11} = 254 \implies a_{11} = 508$$

$$a_{12} = a_{21} = 506$$

$$\frac{1}{2}a_{22} = 254 \implies a_{22} = 508$$

$$b_1 = 50 \quad b_2 = 130 \quad c = 111$$

$$f(x) = \frac{1}{2}\begin{pmatrix} 508 & 506 \\ 506 & 508 \end{pmatrix} x^2 + (50\ 130)x + 111$$

6. Приведите формулы итерации метода спуска. Составьте алгоритм метода спуска (в общем виде).

Ответ: Итерационный процесс:
$$x^{k+1} = x^k + \alpha_k p^k$$
, где p^k — определяется с учетом информации о частных производных, а величина $\alpha_k > 0$, такова что:
$$f(x^{k+1}) < f(x^k)$$
Останов итерационного процесса:
$$\left\|\nabla f(x^k)\right\| < \varepsilon$$

7. Какова особенность траектории поиска для метода градиентного спуска?

   Ответ: На каждой итерации делается шаг в направлении наибольшего убывания фукнции домноженный на α_k, при чем $\alpha_{k+1} = \frac{\alpha_k}{2}$. Получается зигзагообразная ломаная.

8. Как взаимно расположены два последовательных вектора спуска в итерационном процессе с наискорейшим спуском и в процессе с исчерпывающим спуском? Продемонстрируйте на рисунке.

   Ответ: Два последовательных вектора спуска в итерационном процессе с наискорейшим спуском и в процессе с исчерпывающим спуском пендикулярны друг другу, поскольку первый вектор – касательная к линии уровня, а градиент перпендикулярен ей.

9. Дайте определение сопряженных направлений.

   Ответ: Вектора $p^1, p^2 \neq 0$ называются **сопряженными направлениями** если $\left\langle p^1, p^2 \right\rangle_A = \left\langle Ap^1, p^2 \right\rangle = 0$

10. Как связаны сходимость градиентных методов и обусловленность матрицы квадратичной формы? Приведите примеры.

    Ответ: Градиентные методы быстро сходятся на хорошо обусловленных функциях($\mu \sim 1$), линии уровня функции в таком случае близки к окружностям. Для плохо обусловленных функций($\mu \gg 1$), когда линии уровня вытянуты и функция имеет так называемый овражный характер, сходимость становится зигзагообразной — то есть бысто сходится по одному направлению и медленно по другому. Эксперимétальным было обнаружено, что количество итераций линейно увеличивается с ростом числа обусловленности квадратичной формы.

11. Может ли метод сопряженных градиентов делать больше число итераций, чем размерность пространства, в каких случаях? Можно ли метод сопряженных градиентов использовать для не квадратичных функций? Ответ поясните.

    Ответ: Да, если произвойдет накопление погрешности и вектора p^k перестанут указывать в направлении убывания функции. Метод можно использовать для неквадратичных функций, заменив в формуле матрицу A на Гессиан фукции, но для них не доказана сходимость за количество итераций не превышающего размерности пространства.

12. Запишите итерационные формулы метода Полака-Рибьера.

    Ответ:
    $$x^{k+1} = x^k + \alpha_k p^k$$
    $$\alpha_k = \frac{\left\langle \omega^k, p^k \right\rangle}{\langle Ap^k, p^k \rangle}$$
    $$p^k = \omega^k + \beta_k p^{k-1}$$
    $$\beta_k = \frac{\left\langle w^k - w^{k-1}, w^k \right\rangle}{|w^{k-1}|^2}$$

13. Какой принцип задания направления шага в методе Ньютона?

    Ответ: Направление шага задается с использованием антиградиента и матрицы, обратной к Гессиану функции, т.е. $p^k = -H^{-1}(x^{k-1})\nabla f(x^{k-1})$

14. Как выполняется проверка того, что направление метода Ньютона является направлением спуска? По какому принципу задается направление одномерного поиска в методе Ньютона с направлением спуска?

    Ответ: Вектор p^k являвется направлением спуска, если $\left\langle \nabla f(x^{k-1}, p^k) \right\rangle < 0$. Одномерный поиск ищет минимальный α функции $f(x^{k-1} + \alpha p^k)$

15. Приведите формулу метода Марквардта.

    Ответ:
$$(H(x) + \tau I)p^k = -\nabla f(x)$$
    , где τ — параметр, I — единичная матрица

16. Какие достоинства и недостатки метода Ньютона с одномерным поиском привели к открытию квазиньютоновских методов?

    Ответ: Квазиньютоновские методы:

    - Не требуют обращения Гессиана(решения СЛАУ)
    - Объединяют достоинства методов Ньютона и метода нискорейшего спуска
      Методы Ньютона быстро сходятся в окрестности точки минимума x^*.

17. См. (17)

18. Дана матрица:
$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 101 & -10 & -10 \\ 0 & -10 & 2 & 1 \\ 0 & -10 & 1 & 101 \end{bmatrix}$$
    и вектор $f = (1, 0, 1, 0)^T$. Решить СЛАУ $Ax = f$ методом Гаусса с выбором главного элемента.

    Ответ:
$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 101 & -10 & -10 \\ 0 & -10 & 2 & 1 \\ 0 & -10 & 1 & 101 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 100 & -10 & -10 \\ 0 & -10 & 2 & 1 \\ 0 & -10 & 1 & 101 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 100 & -10 & -10 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 100 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1.1 \\ 0.1 \end{pmatrix}$$
    Тогда получаем решение $x = (1.1201\ 0.1201\ 1.1\ 0.001)$

3 3 Вариант

1. Какая функция называется унимодальной на отрезке $[a, b]$?

    Ответ: непрерывная на $[a, b]$ функция и при этом $\exists \alpha, \beta : a <= \alpha <= \beta <= b$ такие что:

    (a) если $a < \alpha$, то на $[a, \alpha]$ f строго монотонно убывает
    (b) если $\beta < b$, то на $[\beta, b]$ f строго монотонно возрастает
    (c) $\forall x \in [\alpha, \beta]$ (если $\alpha \neq \beta$) $f(x) = const = f^* = \min f(x)$ при $x \in [a, b]$

2. Сформулировать свойства унимодальных функций.

    Ответ:

(a) Любая из точек локального минимума унимодальной функции является и точкой ее глобального минимума на отрезке $[a, b]$

(b) Функция, унимодальная на отрезке $[a, b]$, унимодальна и на любом меньшем отрезке $[c, d] \subset [a, b]$

(c) Пусть $f(x)$ унимодальна на $[a, b]$ и $a \leq x_1 < x_2 \leq b$, тогда:

- если $f(x_1) \leq f(x_2)$, то $x^* \in [a, x_2]$
- если $f(x_1) > f(x_2)$, то $x^* \in [x_1, b]$

где x^* — одна из точек минимума $f(x)$ на отрезке $[a, b]$

3. Доказать, что последовательность $\left\{\frac{F_{2n}}{F_{2n+1}}\right\}$ сходится к $\frac{\sqrt{5}-1}{2}$, монотонно возрастая, а $\left\{\frac{F_{2n-1}}{F_{2n}}\right\}$ сходится к $\frac{\sqrt{5}-1}{2}$, монотонно убывая.

Ответ: Вспомним, что:

$$F_n = \left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right) \cdot \frac{1}{\sqrt{5}}$$

$$F_n \xrightarrow[n\to\infty]{} \left(\frac{1+\sqrt{5}}{2}\right)^n \cdot \frac{1}{\sqrt{5}}$$

Тогда получим:

$$\frac{F_{2n}}{F_{2n+1}} < \frac{F_{2n+2}}{F_{2n+3}}$$

$$F_{2n} \cdot F_{2n+3} < F_{2n+2} \cdot F_{2n+1}$$

$$\left(\left(\frac{1+\sqrt{5}}{2}\right)^{4n+3} + \left(\frac{1-\sqrt{5}}{2}\right)^{4n+3}\right) \cdot \frac{1}{5} - \frac{2^{-3}}{5} \cdot ((1+\sqrt{5})^3 + (1-\sqrt{5})^3) <$$

$$< \left(\left(\frac{1+\sqrt{5}}{2}\right)^{4n+3} + \left(\frac{1-\sqrt{5}}{2}\right)^{4n+3}\right) \cdot \frac{1}{5} + \frac{1}{5}$$

$$\frac{4}{5} > -\frac{1}{5} \implies \text{монотонно возрастает}$$

$$\frac{F_{2n}}{F_{2n+1}} \xrightarrow[n\to\infty]{} \frac{2}{1+\sqrt{5}} = \frac{2\sqrt{5}-2}{4} = \frac{\sqrt{5}-1}{2}$$

$$\frac{F_{2n-1}}{F_{2n}} > \frac{F_{2n+1}}{F_{2n+2}}$$

$$F_{2n-1} \cdot F_{2n+2} > F_{2n+1} \cdot F_{2n}$$

$$\left(\left(\frac{1+\sqrt{5}}{2}\right)^{4n+1} + \left(\frac{1-\sqrt{5}}{2}\right)^{4n+1}\right) \cdot \frac{1}{5} + \frac{2^{-3}}{5} \cdot ((1+\sqrt{5})^3 + (1-\sqrt{5})^3) >$$

$$> \left(\left(\frac{1+\sqrt{5}}{2}\right)^{4n+1} + \left(\frac{1-\sqrt{5}}{2}\right)^{4n+1}\right) \cdot \frac{1}{5} - \frac{1}{5}$$

$$-\frac{4}{5} < \frac{1}{5} \implies \text{монотонно убывает}$$

$$\frac{F_{2n-1}}{F_{2n}} \xrightarrow[n\to\infty]{} \frac{2}{1+\sqrt{5}} = \frac{2\sqrt{5}-2}{4} = \frac{\sqrt{5}-1}{2}$$

4. Дайте определение матрицы Гессе функции многих переменных.

Ответ: Матрицей Гессе $H(x)$ дважды непрерывно дифференцируемой в точке x функции $f(x)$ называется матрица частных производных второго порядка, вычисленных в данной точке

$$H(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

5. Для функции $f(x) = 211x_1^2 - 420x_1x_2 + 211x_2^2 - 192x_1 + 50x_2 - 25$ выпишите квадратичную форму, линейную и постоянную часть.

   Ответ:

$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} x^2 + (b_1\ b_2)x + c$$

$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} x + (b_1\ b_2)x + c$$

$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} x + (b_1\ b_2)x + c$$

$$f(x) = \frac{1}{2}(a_{11}x_1^2 + a_{12}x_2x_1 + a_{21}x_1x_2 + a_{22}x_2^2) + b_1x_1 + b_2x_2 + c$$

$$f(x) = \frac{1}{2}a_{11}x_1^2 + a_{12}x_2x_1 + \frac{1}{2}a_{22}x_2^2 + b_1x_1 + b_2x_2 + c$$

$$\frac{1}{2}a_{11} = 211 \implies a_{11} = 422$$

$$a_{12} = a_{21} = -420$$

$$\frac{1}{2}a_{22} = 211 \implies a_{22} = 422$$

$$b_1 = -192 \quad b_2 = 50 \quad c = -25$$

$$f(x) = \frac{1}{2}\begin{pmatrix} 422 & -420 \\ -420 & 422 \end{pmatrix} x^2 + (-192\ 50)x - 25$$

6. Дайте определение направления наискорейшего спуска. Опишите метод наискорейшего спуска.

   Ответ: Направление наискорейшего спуска - это вектор, равный антиградиенту $-\nabla f$. Алгоритм метода:

   (a) взять точность $\varepsilon > 0$, выбрать начальное приближение $x_0 \in E$, высчитать $f(x_0)$
   (b) вычислить $\nabla f(x)$, проверить условие остановки
   (c) найти α^*, минимизирующее $f(x - \alpha \cdot \nabla f(x))$ и положить $x = x - \alpha^* \cdot \nabla f(x)$, перейти к (6b)

7. Какова особенность траектории поиска для метода с исчерпывающим спуском?

   Ответ: Траектория идет по векторам, ортогональным линиям уровня

8. Доказать, что условие $\alpha_k \in \left(0, \frac{2}{L}\right)$, где L — константа Липшица, нарушается для метода с исчерпывающим спуском.

   Ответ: Исчерпывающий спуск на k-той итерации ищет стационарную точку $\varphi_k(\alpha) = f(x^{k-1} + \alpha\omega^k)$

$$\varphi_k(\alpha) = \frac{1}{2}\left\langle A(x^{k-1} + \alpha\omega^k), x^{k-1} + \alpha\omega^k\right\rangle = f(x^{k-1}) + \alpha\left\langle Ax^{k-1}, \omega^k\right\rangle + \frac{\alpha^2}{2}\left\langle A\omega^k, \omega^k\right\rangle$$

   Эта функция имеет положительный коэффициент при старшей степени, следовательно она имеет единственную стационарную точку

$$\alpha_k = -\frac{\left\langle Ax^{k-1}, \omega^k\right\rangle}{\left\langle A\omega^k, \omega^k\right\rangle} = \frac{|\omega^k|^2}{\left\langle A\omega^k, \omega^k\right\rangle}$$

   Если $\operatorname{cond} A = \frac{\lambda_n}{\lambda_1} > 2$ и если x^{k-1} — собственный вектор A с собственным значением λ_1, тогда $\alpha_k = \frac{1}{\lambda_1} \notin \left(0, \frac{2}{\lambda_n}\right)$,

9. Каким условиям должна удовлетворять минимизируемая функция и начальные условия, чтобы метод градиентного спуска/метод наискорейшего спуска сошелся за одну итерацию при любом выборе начального приближения?

   Ответ: Если $A = \lambda I$, где I — единичная матрица, то все собственные значения A совпадают, а каждый ненулевой вектор является собственным. Тогда минимум функции $f(x) = \frac{1}{2}\langle Ax, x\rangle$ достигается за одну итерацию при любом начальном приближении.

10. Какие преимущества дает минимизация квадратичной функции, имеющей канонический вид? Каковы минусы подхода приведения функции к каноническому виду и последующей минимизацией?

    Ответ: Минимизация такой функции производится за 1 шаг метода грдиентного спуска. Но сведение к такому виду вычислительно затратно в силу необходимости находить собственные значения. Кроме того, они находятся численно, а следовательно неточно, что может привести к сходимости за большее число шагов.

11. Сформулируйте основную теорему методов сопряжённых направлений.

    Ответ: Точка минимума функции $f(x) = \frac{1}{2}\langle Ax, x\rangle$ с положительно определенной матрицей A достигается за $\leq n$ итераций спуска по направлениям p^k, таким что они сопряжены относительно A, а α_k вычисляются как:

    $$\alpha_k = -\frac{\left\langle Ax^{k-1}, p^k\right\rangle}{\langle Ap^k, p^k\rangle}$$

    и $x^k = x^{k-1} + \alpha_k p^k$

12. Запишите итерационные формулы метода Флетчера-Ривса.

    Ответ:

    $$p^{k-1} = \omega^{k-1} + \beta_k p^{k-1}$$

    p^{k-1} и p^{k-2} ортогональны $\implies$

    $$\left\langle p^{k-1}, \omega^{k-1}\right\rangle = |\omega^{k-1}|^2 + \beta_k \left\langle p^{k-2}, \omega^{k-1}\right\rangle = |\omega^{k-1}|^2 \implies$$

    $$\implies \beta_k = \frac{|\omega^k|^2}{|\omega^{k-1}|^2} \quad k = 2, \ldots$$

13. Выведите формулу метода Ньютона. Какого порядка метод Ньютона.

    Ответ: Разложим оптимизируемую функцию в ряд Тейлора на k-той итерации:

    $$\varphi_k(x) = f(x^{k-1}) + \left\langle \nabla f(x^{k-1}, x - x^{k-1})\right\rangle + \frac{1}{2}\left\langle H(x^{k-1})\cdot(x - x^{k-1}), x - x^{k-1}\right\rangle$$

    $$\nabla\varphi_k(x) = \nabla f(x^{k-1}) + H(x^{k-1})\cdot(x - x^{k-1})$$
    $$x^k = x^{k-1} - H^{-1}(x^{k-1})\nabla f(x^{k-1})$$

    Метод Ньютона — метод второго порядка, т.к. использует вычисление Гессиана.

14. Опишите метод Ньютона с направлением спуска. Как вычисляется величина шага в методе Ньютона с направлением спуска?

    Ответ: На k-той итерации

    (a) Решаем СЛАУ относительно p^k:
    $$Hp^k = -\nabla f(x)$$
    (b) Если p^k не направление спуска, т.е. $\left\langle \nabla f(x^{k-1}), p^k\right\rangle > 0$, тогда $p^k := -\nabla f(x^{k-1})$
    (c) Используем одномерный поиск для поиска величины шага α_k — точки минимума функции $\varphi_k(\alpha) = f(x^{k-1} + \alpha p^k)$.
    (d) $x^k := x^{k-1} + \alpha_k p^k$
    (e) Проверяем критерий останова $\left\|x^k - x^{k-1}\right\| < \varepsilon$

15. Как решается проблема плохой обусловленности системы линейных алгебраических уравнений в методе Марквардта?

    Ответ:
    $$(H(x) + \tau I)p^k = -\nabla f(x)$$
    Если вначале взять $\tau \gg 1$, тогда матрицей $H(x)$ можно принебречь, и получим уравнение:
    $$\tau I p^k = -\nabla f(x)$$
    В нем нет число обусловленности левой части близко к 1.

16. Выведите квазиньютоновское условие, на котором основаны квазиньютоновские методы.

    Ответ: Пусть $f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle$, где A — положительно определенная. Тогда $\forall x: \; A = H, \; \nabla f(x) = Ax + b$. Пусть $\Delta w^k = w^k - w^{k-1}, \; \Delta x^k = x^k - x^{k-1}$, тогда
    $$\Delta w^k = \nabla f(x^{k-1}) - \nabla f(x^k) = A(x^{k-1} - x^k) = -A\Delta x^k$$
    Таким образом, для квадратичной функции
    $$A^{-1}\Delta w^k = -x^k$$
    То есть требуется чтобы приближенные матрицы G_k удовлетворяли условию:
    $$G_{k+1}\Delta w^k = -\Delta x^k$$

17. См. (17)

18. Указать элементы матрицы L, которые могут быть ненулевыми при построении LU-разложения матрицы:
    $$A = \begin{bmatrix} 12 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 14 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 7 & 3 & 0 & -2 & 0 \\ 0 & 1 & 0 & 3 & 8 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 9 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 & 1 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \end{bmatrix}$$
    Ответ: L — нижнетреугольная, а также ее профиль не может увеличиться, следовательно:
    $$A = \begin{bmatrix} \color{red}{12} & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \color{red}{-1} & \color{red}{14} & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & \color{red}{6} & 0 & 0 & 0 & 0 & 0 \\ 0 & \color{red}{1} & 0 & \color{red}{7} & 3 & 0 & -2 & 0 \\ 0 & \color{red}{1} & 0 & \color{red}{3} & \color{red}{8} & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \color{red}{2} & \color{red}{9} & 1 & 0 \\ 0 & 0 & 0 & \color{red}{-2} & \color{red}{0} & \color{red}{1} & \color{red}{12} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \color{red}{12} \end{bmatrix}$$

4 4 Вариант

1. Какая функция называется выпуклой на отрезке $[a, b]$?

   Ответ: $f(x)$ — выпуклая, если $\forall x, y; \; \forall \alpha \in [0, 1]: \; f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$

2. Каков геометрический смысл выпуклости функции?

   Ответ: Если функция $f(x)$ выпукла на отрезке $[a, b]$, то $\forall x_1, x_2 \in [a, b]$ график функции лежит ниже хорды, проведенной через точки графика с абциссами x_1, x_2

3. Показать, что отношение длин интервалов неопределенности метода золотого сечения к методу Фибоначчи составляет ≈ 1.1708.

Ответ: Рассмотрим длинну отрезка на n-той итерации

- Метод Фибоначчи
$$\Delta_F^n = \frac{1}{F_{n+2}} \cdot (b_0 - a_0)$$
- Метод золотого сечения
$$\Delta_G^n = \tau^n (b_0 - a_0) \quad \tau = \frac{\sqrt{5}-1}{2}$$

$$\frac{\Delta_G^n}{\Delta_F^n} = \tau^n \cdot F_{n+2} \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{5}} \cdot \left(\frac{\sqrt{5}-1}{2}\right)^n \cdot \left(\frac{\sqrt{5}+1}{2}\right)^{n+2} =$$
$$= \frac{(\sqrt{5}+1)^2}{4\sqrt{5}} \approx 1.1708$$

4. Объясните свойства матрицы Гессе функции многих переменных.

Ответ:

- Матрицы Гессе является квадратной симметричной матрицей.
- $\Delta x^T H(x) \Delta x$ — квадратичная форма
- По матрице Гессе можно определить выпклость функции
  - если $\forall x:\ H(x) \geq 0$, тогда функция выпуклая
  - если $\forall x:\ H(x) > 0$, тогда функция строго выпуклая
  - если $\forall x:\ H(x) \leq lE$, где E — единичная матрицы, тогда функция сильно выпуклая

  Аналогично для вогнутости

5. Для функции $f(x) = 99x_1^2 + 196x_1x_2 + 99x_2^2 - 95x_1 - 9x_2 + 91$ выпишите квадратичную форму, линейную и постоянную часть.

Ответ:
$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} x^2 + (b_1 \ b_2)x + c$$
$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} x + (b_1 \ b_2)x + c$$
$$f(x) = \frac{1}{2}\begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} x + (b_1 \ b_2)x + c$$
$$f(x) = \frac{1}{2}(a_{11}x_1^2 + a_{12}x_2x_1 + a_{21}x_1x_2 + a_{22}x_2^2) + b_1x_1 + b_2x_2 + c$$
$$f(x) = \frac{1}{2}a_{11}x_1^2 + a_{12}x_2x_1 + \frac{1}{2}a_{22}x_2^2 + b_1x_1 + b_2x_2 + c$$
$$\frac{1}{2}a_{11} = 99 \implies a_{11} = 198$$
$$a_{12} = a_{21} = 196$$
$$\frac{1}{2}a_{22} = 99 \implies a_{22} = 198$$
$$b_1 = -95 \quad b_2 = -9 \quad c = 91$$
$$f(x) = \frac{1}{2}\begin{pmatrix} 198 & 196 \\ 196 & 198 \end{pmatrix} x^2 + (-95 \ -9)x + 91$$

6. По какому принципу задается направление шага в методе наискорейшего спуска? Как вычисляется величина шага в методе наискорейшего спуска?

Ответ: В наискорейшем спуске направление движения минимизации берется как направление антиградиента $p^k = -\nabla f(x^k)$. Для выбора величины шага решается задача одномерной минимизации функции $\varphi_k(\alpha) = f(x^{k-1} + \alpha_k p^k)$. В таком случае это будет точка качания линии уровня.

7. Можно ли в методе наискорейшего спуска для квадратичной функции определить величину исчерпывающего спуска без решения одномерной задачи минимизации? Если да, то как?

   Ответ: Да, можно. Если дана квадратичная функция вида $f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c$, то α_k для спуска можно вычислить так:

$$\alpha_k = -\frac{\langle Ax^k + b, p^k \rangle}{\langle Ap^k, p^k\rangle}$$

8. Запишите формулу для градиента квадратичной функции. Как определить константу Липшица для произвольной аналитически заданной функции (для метода градиентного спуска)?

   Ответ: Для кадратичной функции $f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c$ градиент будет $\nabla f(x) = Ax + b$. L — константа Липшица, если $\forall x, y: \ |f(x) - f(y)| \leq L \cdot |x - y|$. Для градиента справедливо:

$$f(x + h) \approx f(x) + (\nabla f(x))^T h$$

   Значит можно оценить константу Липшица как $L = \sup|\nabla f(x)|$

9. Указать условия того, что $\{x^k\}$ — последовательность будет релаксационной в итерационном методе градиентного спуска: для квадратичной функции и для функции общего вида.

   Ответ: Чтобы последовательность $\{x^k\}$ была релаксационной для градиентного спуска, должно выполняться равенство:

$$x^k = x^{k-1} + \alpha\gamma^k \quad \alpha \in \left(0, \frac{2}{L}\right)$$

   , где $\gamma = -\nabla f(x)$, L — наибольшее собственное значение матрицы A квадратичной функции или константа липшицы произвольной функции.

10. В каком случае следует в методе градиентного спуска применять дробление шага α_k? Какие условия используются для принятия решения о дроблении шага на k-ой итерации?

    Ответ: На каждом шаге проверяем пока условие $f(x^k) > f(x^k + \alpha p^k)$ ложно, то делаем $\alpha \coloneqq \frac{\alpha}{2}$.

11. Может ли итерационный процесс метода сопряженных градиентов сходится для СЛАУ с отрицательно определенной матрицей? Ответ обоснуйте.

    Ответ: Для отрицательно определенных тоже сойдется. Антиградиент в этом случае будет смотреть от максимума, поэтому если домножить обе части СЛАУ на -1, то получим слева положительно определенную матрицу, а справа антиградиент, смотрящий в сторону минимума. Также во всех формулах метода сопряженных градиентов входит и градиент и матрица, поэтому если от A перейти к $-A$, то градиент поменяет знак, и минусы уничтожатся, а значит формулы останутся такими же как и для положительно определенных.

12. Обоснуйте необходимость применения рестартов и раскройте роль рестартов в повышении эффективности метода Флетчера-Ривса.

    Ответ: $\beta_k = 0$ через заданное число итераций (моменты рестарта кратные n). Это позволяет избежать накопления вычислительных погрешностей и уменьшить вероятность построения после каждых n итераций линейно зависимых направлений спуска, но приводит к росту общего числа итераций

13. В чем заключается основная идея метода Ньютона? Запишите итерационные формулы метода Ньютона.

    Ответ: Если $f(x)$ — дважды дифференцируемая функция, то через градиент и матрицу Гессе можно разложить в ряд Тейлора функцию:

$$f(x) = f(x^k) + \nabla f(x^k)^T\Delta x + \frac{1}{2}\Delta x^T H(x^k)\Delta x + o(\|\Delta x^k\|)$$

$$\Phi_k(x) = f(x^k) + \nabla f(x^k)^T\Delta x + \frac{1}{2}\Delta x^T H(x^k)\Delta x$$

    Найдем минимум функции $\Phi_k(x)$ через условие $\nabla\Phi_k(x) = 0$:

$$\nabla\Phi_k(x) = \nabla f(x^k) + H(x^k)(x - x^k) = 0$$

$$x^{k+1} = x^k - H^{-1}(x^k)\nabla f(x^k)$$

    — это итерационный процесс метода Ньютона

14. Как заканчиваются вычисления в методе Ньютона с направлением спуска? Какого порядка метод Ньютона с направлением спуска?

    Ответ: Условие завершения метода Ньютона $|\Delta x| < \varepsilon$. Метод Ньютона второго порядка.

15. Укажите достоинства и недостатки метода Марквардта.

    Ответ:

    \+ простота

    \+ высокая скорость сходимости в окрестности x^*

    \+ не нужен одномерный поиск

    \+ $f(x^k)$ убывает в итерационном процессе

    – решение СЛАУ

    – вычисление Гессиана

16. Как определяется направление одномерного поиска в квазиньютоновских методах?

    Ответ:
$$x^k = x^{k-1} + \alpha_k p^k$$
$$p^k = -G_k \cdot \nabla f(x^{k-1})$$

    Параметр α_k может находиться с помощью одномерного поиска по функции:
$$\varphi_k(\alpha) = f(x^{k-1} + \alpha p^k)$$

17. См. (17)

18. Записать структуры для хранения матрицы в разреженно строчно-столбцовом и профильном форматах:
$$A = \begin{bmatrix} 12 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 14 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 7 & 3 & 0 & -2 & 0 \\ 0 & 1 & 0 & 3 & 8 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 9 & 1 & 0 \\ 0 & 0 & 0 & -2 & 0 & 1 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \end{bmatrix}$$

    Ответ:

    - Строчно-столбцовый формат
$$di = [12, 14, 6, 7, 8, 9, 12, 12]$$
$$ia = [1, 1, 2, 2, 3, 5, 6, 8, 8]$$
$$ja = [1, 2, 2, 4, 5, 4, 6]$$
$$al = au = [-1, 1, 1, 3, 2, -2, 1]$$

    - Профильный формат
$$di = [12, 14, 6, 7, 8, 9, 12, 12]$$
$$ia = [1, 1, 2, 2, 4, 7, 8, 11, 11]$$
$$al = au = [-1, 1, 0, 1, 0, 3, 2, -2, 0, 1]$$