Экзамен. Вариант 4

Илья Ярошевский M3237

13 мая 2023 г.

1. 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
   $$\Delta_G^n = \tau^n (b_0 - a_0) \quad \tau = \frac{\sqrt{5}-1}{2}$$

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

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

   Ответ: $f(x)$ удовлетворяет условию Липшица, если
   $$\exists L \ \forall x_1, x_2 \in [a, b] : \ |f(x_1) - f(x_2)| \leq L|x_1 - x_2|$$

3. Увеличение используемого значения константы Липшица L при реализации метода ломаных приводит к замедлению сходимости метода. Объяснить этот факт с помощью геометрической иллюстрации.

   Ответ: Шаг в методе ломанных определяется как:
   $$\Delta_1 = \frac{1}{2L} f(x_k^* - p_1^*)$$

4. Сравните методы Ньютона (одномерная минимизация) и секущих. Каков порядок сходимости каждого?

   Ответ: Преимущества и недостатки методов Ньютона по сравнению с методом секущих:

   \+ высокая скорость сходимости

   − $p = -\frac{f'(x^k)}{f''(x^k)}$ определено только тогда, когда $f''(x^k) \neq 0$

   − требуется вычислять $f''(x)$

   Метод Ньютона имеет квадратичный порядок сходимости. Порядок сходимости метода секущих сравнителен с порядком сходимости метода золотого сечения, т.е. быстрее чем линейный, но не квадратичный.

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 задается антиградиентом функции $p^k = -\nabla f(x^k)$. Величина шага находится с помощью решения задача одномерной оптимизации функции $\varphi_k(\alpha) = f(x^{k-1} + \alpha_k p^k)$. В таком случае это будет точка касания 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:

$$\alpha_k = -\frac{\left\langle \nabla f(x^0), p^k \right\rangle}{\langle Ap^k, p^k\rangle}$$

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 $\forall x, y :\ |f(x) - f(y)| \le L \cdot |x - y|$. Для градиента справедливо:

$$f(x+h) \approx f(x) + (\nabla f(x))^T h \quad h > 0$$

$$|f(x+h) - f(x)| \le L \cdot |(x+h) - x|$$

$$h \cdot |\nabla f(x)| \le L \cdot h$$

$$|\nabla f(x)| \le L$$

   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:

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

   , где L — наибольшее собственное значение матрицы A квадратичной функции или константа Липшица 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 $\alpha := \frac{\alpha}{2}$.

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

    Ответ: Да, может. Можно заметить что во всех формулах, где встречается матрица A, также встречается и направление спуска. Следовательно, если домножить матрицу A на -1, то направление спуска также поменяет направление, а значит в формулах не играет роли положиельно или отрицательно определена матрица A. Значит метод будет по прежнему сходиться.

12. Матрицы СЛАУ с диагональным преобладанием: что это такое? Обязательно ли в методе сопряженных градиентов матрица должна быть с диагональным преобладанием для того, чтобы метод был сходящимся.

    Ответ: Матрицы с диагональным преобладанием — матрица у которой на главной диагонали стоят элементы большие чем внедиагональные. Не обязательно для метода сопряженных градиентов наличие диагонального преобладания, важную роль играет только ее обусловленность.

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 $f(x)$ получив квадратичную 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. Для каких функций эффективно применение методов второго порядка?

    Ответ: Методы второго поряда используют матрицу Гессе функции для вычисления приблизительной квадратичной формы. Следовательно для функций, которые хорошо аппроксимируются рядом Тейлора второго порядка такие методы эффективны.

15. Укажите достоинства и недостатки метода Марквардта.

    Ответ: Комбинация методов наискорейшего спуска и метода Ньютона

    + относительная простота
    + свойство убывания $f(x)$ при переходе от итерации к итерации
    + отсутствие процедуры одномерного поиска
    + высокая скоростью сходимости в окрестности x^*
    − 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. 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:

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

Пример. Пример решение СЛАУ на машине с высокой погрешностью вычислений чисел с плавающей точкой:

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

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

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

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

$$0.8557 \le \operatorname{cond} A \cdot 0.0092$$

$$1130.2444 = \operatorname{cond} A \ge 92.6864$$

18. Записать структуры для хранения матрицы в разреженно строчно-столбцовом и профильном форматах:

$$A = \begin{bmatrix} 10 & 1 & 0 & 2 & 0 & 0 & -1 & 0 & 0 \\ 1 & 20 & 0 & 0 & 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 30 & 0 & 0 & 4 & 5 & 0 & 2 \\ 2 & 0 & 0 & 40 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 50 & 1 & 0 & -1 & 0 \\ 0 & 0 & 4 & 0 & 1 & 60 & 0 & 0 & 7 \\ -1 & 0 & 5 & 0 & 0 & 0 & 70 & 0 & 0 \\ 0 & 6 & 0 & 0 & -1 & 0 & 0 & 80 & 0 \\ 0 & 0 & 2 & 0 & 0 & 7 & 0 & 0 & 90 \end{bmatrix}$$

Ответ:

- Строчно-столбцовый формат

$$\begin{aligned} di &= [10, 20, 30, 40, 50, 60, 70, 80, 90] \\ ia &= [1, 1, 2, 2, 3, 3, 5, 7, 9, 11] \\ ja &= [1, 1, 3, 5, 1, 3, 2, 5, 3, 6] \\ al = au &= [1, 2, 4, 1, -1, 5, 6, -1, 2, 7] \end{aligned}$$

- Профильный формат

$$\begin{aligned} di &= [10, 20, 30, 40, 50, 60, 70, 80, 90] \\ ia &= [1, 1, 2, 2, 5, 5, 8, 14, 20, 26] \\ al = au &= [1, 2, 0, 0, 4, 0, 1, -1, 0, 5, 0, 0, 0, 6, 0, 0, -1, 0, 0, 2, 0, 0, 7, 0, 0] \end{aligned}$$