Вычислительные Методы Решения

Kasymkhan Khubiev

September 2023

1 Пачка 1

1. Задача 1

Доказать, что при оптимизации квадратичной функции без ограничения общности матрицу A можно считать симметричной.

Док-во:

Пусть A - не симметричная матрица и квадратичная функция представима в виде:

$$f(x) = (Ax, x)/2 + (b, x) + c$$

Мы хотим найти такую матрицу $\tilde{A} | f(x) = (\tilde{A}x, x)/2 + (b, x) + c$

Начнем с описания прироста функции:

$$f(x+y) = \frac{1}{2}(A(x+y), (x+y)) + (b, (x+y)) + c$$

Раскроем скобки в скалярных произведениях и соберем то, что получится:

$$f(x+y) = [\frac{1}{2}(Ax, x) + (b, x) + c] + \frac{1}{2}(Ax, y) + \frac{1}{2}(Ay, x) + (b, y) + \frac{1}{2}(Ay, y) = f(x) + \frac{1}{2}(Ax + A^T x + 2b, y) + \frac{1}{2}(Ay, y)$$

если положить $\frac{1}{2}(Ay, y) = o(y)$, то вектор $\frac{1}{2}(Ax + A^T x + 2b) = grad[f(x)]$

теперь возьмем вторую производную:

$$grad^2[f(x)] = \frac{1}{2}(A + A^T)$$

теперь если расписать полученную сумму матриц, получим следующее. Пусть матрица $B = \frac{1}{2}(A + A^T)$ с элементами b_{ij} , где $i, j = 1, ..., n$

тогда $b_{ij} = \frac{a_{ij}+a_{ji}}{2}$ при $i \neq j$ и $b_{ij} = a_{ij}$ при $i = j$

т.е. мы получили симметричную матрицу B, значит искомая матрица $\tilde{A} = B$ и любая квадратичная функция приводится к квадратичной функции с симметричной матрицей.

2. Задача 2. Доказать, что минимум квадратичной функции $f(x) = \frac{1}{2}x^TAx + b^Tx + c$ существует тогда и только тогда, когда матрица A положительно определена.

Док-во:

По определению матрица A называется положительно определенной, если для любого $x \in R^n$ выполняется неравенство: $(Ax, x) > 0$

Мы знаем, что A - симметричная матрица. Запишем прирост функции $f(x)$:

$$f(x+y) = \frac{1}{2}(A(x+y), (x+y)) + (b, (x+y)) + c =$$

$$\frac{1}{2}(Ax, x) + (b, x) + c + \frac{1}{2}(Ax, y) + \frac{1}{2}(Ay, x) + (b, y) + \frac{1}{2}(Ay, y) =$$

$$f(x) + ((Ax - b), y) + \frac{1}{2}(Ay, y)$$

откуда видим, что $Ax - b = grad[f(x)]$

теперь воспользуемся достаточным условием существования локального минимума:

$$f''(x) > 0$$

тогда для квадратичной функции имеем:

$$grad^2[f(x)] = A$$

- но гессиан это матрица, а второй градиент должен действовать так:

$$grad^2[f(x)] : R^{n \times n} - > R^1$$

значит с учетом условия минимума и типа возвращаемого значения имеем:

$$(Ax, x) > 0$$

, что в точности является признаком положительной определенности для матрицы. Ч.т.д.

3. Задача 3.

Найти оценку скорости сходимости для квадратичной функции $f(x) = \frac{1}{2}x^TAx + b^Tx + c$

Решение:

Будем рассматривать простейший (классический) вариант градиентного спуска:

$$x_{k+1} = x_k - \alpha grad[f(x_k)]$$

, положим шаг градиента постоянным.

ранее вычисляли, что $grad[f(x)] = Ax - b$, тогда имеем:

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

Пусть x^* - оптимальное решение, т.е. в этой точке градиент зануляется:

$$grad[f(x^*)] = Ax^* - b = 0 => Ax^* = b$$

(*)

Вычтем из обоих частей уравнения градиентного спуска x^*:

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

итеративно заменим все x_k до x_0:

$$||x_{k+1} - x^*|| \leq ||I - \alpha A||^{k+1}||x_0 - x^*||$$

что и требовалось найти.

4. Задача 4.

Вычислить $\delta f(x)$ для $f(x) = |x|$. Показать, что для $f(x) = |x|$ при постоянном шаге субградиентный метод не сходится.

Решение:

По определению дифференцируемая функция $f(x)$ выпукла, когда

$$f(y) \geq f(x) + grad[f(x)]^T(y - x)|\forall x, y$$

Субградиент это такой вектор для которого верно неравенство:

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

Множество всех субградиентов в т. x называется субдифференциалом x и обозначается $\delta f(x)$

Функция $f(x) = |x|$ дифференцируема на всем пространстве R кроме точки $x = 0$,

Но в точке нуля мы можем взять один из градиентов слева или справа.

для $x > 0$ можно взять вектор -1, для $x < 0$ вектор 1, а в точке ноль $-1, 1$

О несходимости субградиентного метода при постоянном шаге: Помним, что классический метод градиентного спуска с применением субградиента выглядит так:

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

пусть $\alpha = const$, тогда имеем: $x_{k+1} = x_k - \alpha sing(x_k)$

5. Задача 5.

Вычислить производную по направлению $f'(0, y)$ для $f(x) = ||x||$.

Решение:

По определению производной по направлению называется такая функция:

$$f^{'}(x,y) = \lim_{\epsilon->0} \frac{f(x+\epsilon y) - f(x)}{\epsilon}$$

Тогда в нашем случае имеем:

$$f^{'}(0,y) = \lim_{\epsilon->0} \frac{f(0+\epsilon y) - f(0)}{\epsilon} = \lim_{\epsilon->0} \frac{||\epsilon y||}{\epsilon} = ||y||$$

Ответ: $f^{'}(0,y) = ||y||$ при $f(x) = ||x||$

6. Задача 6.
Найти $\nabla[f(x)]$ для $f(x) = ||x||$ при $x \neq 0$
Решение:
Рассматриваем Евклидову норму: $||x|| = \sqrt{\Sigma x_i^2}$
Тогда $\frac{df}{dx} = (\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}) = \frac{1}{2||x||}(2x_1, ..., 2x_n) = \frac{x}{||x||}$, при $x \neq 0$
в точке $x = 0$ функция не дифференцируема

7. Задача 7.
Найти $\nabla f(x)$ для $f(x) = ||x_+||^2$
Решение:

$$f(x) = ||x_+||^2 = ||x||^2, x > 0, ||x|| = \Sigma x_i^2$$
$$f(x) = 0, x \leq 0$$

Если построить график в одномерном случае, то видно что на отрицательный промежуток градиент не распространяется, там он будет нулевым - функция постоянна, следовательно скорость равна нулю.

Тогда имеем $\nabla f(x) = \nabla||x_+||^2 = 2x$ при $x > 0$ и $\nabla f(x) \equiv 0$ при $x \leq 0$

Тогда объединив оба равенства получаем:

$$\nabla f(x) = \nabla||x_+||^2 = 2x_+$$

8. Задача 8.
Найти $\nabla||(Ax-b)_+||$
Решение:
Вспомним правило дифференцирования сложной функции:

$$[h(g(x))]^{'} = h'(g)g'(x)$$

Поэтому разобъем задачу на два этапа: сначала найдем градиент от функции $f(x) = ||x_+||$, а потом рассчитаем градиент $Ax - b$

Из рассмотрения линейного случая и взяв Евклидову норму получим, что

$$\nabla||x_+|| = \frac{x}{||x||}, x > 0$$

$$\nabla||x_+|| \equiv 0, x < 0$$

Тогда обобщая оба случая имеем:

$$\nabla||x_+|| = \frac{x_+}{||x_+||}$$

И получаем с учетом нашей функции:

$$\nabla f(x) = \frac{(Ax-b)_+}{||(Ax-b)_+||}\nabla(Ax-b)$$

В то же время $\nabla Ax - b = A$ - матрица
теперь нам нужно понять с какой стороны поставить градиент от $Ax-b$
если A - матрица размера $m \times n$, x - вектор размера $n \times 1$
Тогда $Ax - b$ - вектор размера $(m \times n)(n \times 1) = (m \times 1)$
Ну а мы должны получить вектор, значит матрицу A нужно транспонировать и умножить на вектор в числителе слева:

$$\nabla||(Ax-b)_+|| = \frac{A^T(Ax-b)_+}{||(Ax-b)_+||}$$

2 Пачка 2

1. Задача 1.

Доказать, что функция $f : Q -> R$ имеет глобальное решение

$$Q = \{(x_1, x_2) \in R^2 | x_1 + x_2 > 0\},$$

$$f(x_1, x_2) = x_1^2 + x_2 + \frac{1}{x_1 + x_2}.$$

Решение:

У нас задача с ограничением, условие на $x_1 + x_2 > 0$ вырезает полуплоскость без границы

Можем рассчитать: $\nabla f = (2x_1 - \frac{1}{(x_1+x_2)^2}, 1 - \frac{1}{(x_1+x_2)^2}) = (0, 0)$

Тогда приходим к системе уравнений:

$$(1) 2x_1 - \frac{1}{(x_1 + x_2)^2} = 0$$

$$(2) 1 - \frac{1}{(x_1 + x_2)^2} = 0$$

Вычтем (2) из (1):

$$2x_1 - 1 = 0 => x_1 = 1/2$$

Тогда $x_2 = -1/2$ или $x_2 = 3/2$

Точка (1/2, -1/2) - лежит на границе и не попадает в область допустимых значений

Поэтому исследуем стационарную точку (1/2, 3/2) на локальный минимум:

запишем градиент:

$$A = \begin{pmatrix} 2 + \frac{2}{(x_1+x_2)^3} & \frac{2}{(x_1+x_2)^3} \\ \frac{2}{(x_1+x_2)^3} & \frac{2}{(x_1+x_2)^3} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1/4 & 1/4 \end{pmatrix}$$

Матрица А положительно определена в этой точке по критерию Сильвестра - положительны главные миноры матрицы

Теперь поговорим о глобальности решения. Нам нужно показать, что для любого вектора $y \in Q$ верно неравенство $f(y) > f(1/2, 3/2)$ //

То есть условие глобального минимума говорит об ограниченности функции снизу значением в минимальной точке

В то же время, чтобы говорить о наличии глобального минимума мы должные сравнить значения функции в точках локального минимума и вдоль границы.

Точка локального минимума единственная -> проверим границы

если мы возъмем значения x_1, x_2 на бесконечности положительной, то функция уйдет в бесконечность, если мы идем вдоль границы - функция резко растет вверх на бесконечность.

Таким образом имеем одну точку локального минимума и функция стремится к бесконечности вдоль границы, значит найденная точка минимума является глобальной в рассматриваемой области.

2. Задача 2.

С помощью условий первого и второго порядка найти экстемальные точки функции $f : R^2 - > R$

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

//

Чтобы найти экстремальные точки найдем экстремум функции и занулим:

$$\nabla f(x) = (3x_1^2 - 3x_2, 3x_2^2 - 3x_1) = (0, 0)$$

получаем экстремальные точки: $x^{(1)} = (0, 0)$ и $x^{(2)} = (1, 1)//$

Тогда исследуем на локальность минимума:

$$\nabla^2 f(x) = \begin{pmatrix} 6x_1 & -3 \\ -3 & 6x_2 \end{pmatrix}$$

в точке (0,0) не подходит

$$\nabla^2 f(x) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$$

в точке (1, 1) имеем:

$$B = \nabla^2 f(x) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$

матрица B положительна определена по критерию Сильвестра, т.е. точка (1, 1) является точкой локального минимума.

3. Задача 3.

Найти глобальные решения для

$$Q = \{(x_1, x_2, x_3) \in R^3 | x_1 > 0, x_2 > 0, x_3 > 0\}$$

$$f(x_1, x_2, x_3) = x_1 + \frac{x_2^2}{4x_1} + \frac{x_3^2}{x_2} + \frac{2}{x_3} - - > min$$

посчитаем производную функции:

$$\nabla f(x) = (1 - \frac{x_2^2}{4x_1^2}, \frac{x_2}{2x_1} - \frac{x_3^2}{x_2^2}, \frac{2x_3}{x_2} - \frac{2}{x_3^2}) = (0,0,0)$$

решая эту систему уравнений получаем:

$$x_3^6(x_3^8 - 1) = 0,$$

$$x_2 = x_3^3,$$

$$x_1 = x_3^7/2$$

Получаем 3 точки: $x^(1) = (0,0,0), x^(2) = (1/2,1,1), x^(3) = (-1/2,-1,-1)$, при этом $x_3||x_2$

Точка $x^(1) = (0,0,0)$ не подходит, т.к. лежит вне области допустимых значений функции $f(x)$

Давайте рассчитаем вторую производную:

$$\nabla^2 f(x) = \begin{pmatrix} \frac{x_2^2}{2x_1^3} & \frac{x_2}{2x_1^2} & 0 \\ -\frac{x_2}{2x_1^2} & \frac{1}{2x_1} + \frac{2x_3^2}{x_2^3} & \frac{2}{x_2} + \frac{4}{x_3^3} \\ 0 & -\frac{2x_2}{x_2^2} & \frac{2}{x_2} + \frac{4}{x_3^3} \end{pmatrix}$$

$$\nabla^2 f(1/2,1,1) = \begin{pmatrix} 4 & 2 & 0 \\ -2 & 3 & 6 \\ 0 & -2 & 6 \end{pmatrix} > 0$$

т.е. Гессиан положительно определен по критерию Сильвестра. Вторая точка не подходит по второму критерию лок минимума - матрица Гессе не положительно определена.

Теперь проверим поведение функции на границе:

1. при $x_1, x_2 \approx 0$, и $x_3 \to +\infty$ имеем: $f(x) \to +\infty$
2. при $x_1, x_3 \approx 0$, и $x_2 \to +\infty$ имеем: $f(x) \to +\infty$
3. при $x_3, x_2 \approx 0$, и $x_1 \to +\infty$ имеем: $f(x) \to +\infty$
4. при $x_1, x_2, x_3 \approx 0$ имеем: $f(x) \to +\infty$
5. при $x_1, x_2, x_3 \to +\infty$ имеем: $f(x) \to +\infty$

Таким образом всюду вблизи границы функция стремится к бесконечности, значит найденное решение локального минимума является решением глобальной задачи.

4. Задача 4.

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

Решение:

Рассмотрим функцию $f(x) = e^x - ||x+1||$
для n = 1 имеем простой случай: $f(x) = e^x - |x+1|$
Для построения графика функции рассмотрим следующие промежутки:

1. $|x| < 1 \to f(x) \approx 1 + x + \frac{x^2}{2} - x - 1 = \frac{x^2}{2}$ - ведет себя как параболла с минимумом в нуле 2. $x >> 1 \to f(x) \approx e^x$ 3. $x = -1 | f(x) = e^-1$ - тут зануляется модуль и функции "загибается"но в этой точке по прежнему дифференцируема 4. $x << -1 \to f(x) \approx -|x+1|$ - экспонента уже занулилась и только отрицательный модуль дает линейный вклад.

Видим, что функция всюду дифференцируема, и минимум достигается на минус бесконечности.

Возьмем и найдет стационарные точки и проверим на локальный минимум:

$$\nabla f(x) = e^x - 1 = 0$$
$$\nabla^2 f(x) = e^x$$

$x = 0$ - точка локального минимума, но из анализа поведения функции следует, что эта точка не является глобальным минимумом.

5. Задача 5.

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

Решение:

Условие Липшица говорит о следуюзщем:

$$|f(y) - f(x)| \leq L||y - x||$$

перепишем так:

$$\frac{|f(y) - f(x)|}{||y - x||} \leq L$$

Заметим, что:

$$\frac{|f(y) - f(x)|}{||y - x||} \xrightarrow{||y-x|| \to 0} ||\nabla f(x)|| \leq L$$

Тогда рассмотрим такую функцию: $f(x) = \sqrt{|x|}$, $f'(x) = \frac{1}{2\sqrt{|x|}}$, но

$$f'(x) \xrightarrow{x \to 0} \infty$$

Следовательно функция не удовлетворяет условию Липшица!

6. Задача 6.

Доказать, что при любом значении параметра $\sigma > 1$ система уравнений относительно $x \in R^2$ имеет ненулевое решение.

$$\begin{cases} \sigma cos x_1 sin x_2 + x_1 e^{x_1^2+x_2^2} = 0 \\ \sigma sin x_1 cos x_2 + x_2 e^{x_1^2+x_2^2} = 0 \end{cases}$$

Решение:

Сложим первое и второе уравнение и воспользуемся тригонометрическим тожеством:

$$\sigma sin(x_1 + x_2) + (x_1 + x_2) e^{x_1^2+x_2^2} = 0$$

При этом знаем, что $\sigma > 0$ по условию и $e^y \neq 0 \forall y \in R$
Значит должно выполниться следующее:

$$\begin{cases} sin(x_1 + x_2) = 0 \\ x_1 + x_2 = 0 \end{cases}$$

Тогда получаем следующее: $x_1 = -x_2$
теперь рассмотрим систему при таких значениях:

$$\begin{cases} -\sigma cos x_2 sin x_2 - x_2 e^{2x_2^2} = 0 \\ \sigma sin x_2 cos x_2 + x_2 e^{2x_2^2} = 0 \end{cases}$$

3 Пачка 3

1. Задача 1. Пусть $\phi : R^m \to R^m, f : R^m \to R$ дифференцируемые функции и $h(x) = f(\phi(x))$. Выписать: $dh(x)$ и $d^2h(x)$ через $d\phi_j$ и dx.

Решение:

из условий задачи следует, что функция $h(x) : R^n \to R$ Воспользуемся формулой дифференцирования сложной функции: $\frac{dh(g(x))}{dx} = h'(g(x))g'(x)$ и $[h'(g)g'(x)]' = h''(x)g'^2(x) + h'(x)g''(x)$

$$A = d\phi = \begin{pmatrix} \frac{d\phi_1}{dx_1} & ... & \frac{d\phi_m}{dx_1} \\ ... & ... & ... \\ \frac{d\phi_1}{dx_n} & ... & \frac{d\phi_m}{dx_n} \end{pmatrix}$$

$$g = df = (\frac{df}{d\phi_1}, ..., \frac{df}{d\phi_m})$$

$$B = \nabla^2 f = \begin{pmatrix} \frac{df}{(d\phi_1)^2} & ... & \frac{df}{d\phi_1 d\phi_m} \\ ... & ... & ... \\ \frac{df}{d\phi_m d\phi_1} & ... & \frac{d\phi_m}{(d\phi_m)^2} \end{pmatrix}$$

$$C = \nabla\phi = \begin{pmatrix} \frac{d\phi_1}{(dx_1)^2} & ... & \frac{d\phi_m}{dx_1 dx_n} \\ ... & ... & ... \\ \frac{d\phi_1}{dx_n dx_1} & ... & \frac{d\phi_m}{(dx_n)^2} \end{pmatrix}$$

Тогда получаем ответ:

$$dh = A\nabla f,$$

$$d^2h = A^T BA + C\nabla f$$

2. Задача 2.

Пусть $f \in C^2(R^n)$. Доказать, что

$$f(y) = f(x) + \int_0^1 (f'(x + t(y - x)), (y - x))dt;$$

$$f'(y) = f'(x) + \int_0^1 f''(x + t(y - x))(y - x)dt.$$

Доказательство:

1) т.к. $C^1(R^n) \subset C^2(R^n)$ то $f(x) \in C^1(R^n)$, поэтому можем воспользоваться формулой Ньютона-Лейбница:

$$\phi(1) = \phi(0) + \int_0^1 \phi'(t)dt$$

В качестве функции $\phi(t)$ рассмотрим следующую функцию:

$$\phi(t) = f(x + t(y - x))$$

Тогда имеем:

$$\phi(1) = f(x + y - x) = f(y)$$
$$\phi(0) = f(x)$$
$$\phi'(t) = f'(x + t(y - x))(y - x)$$

Тогда с учетом формулы Н-Л имеем:

$$f(y) = f(x) + \int_0^1 (f'(x + t(y - x)), (y - x))dt$$

ч.т.д.

2) т.к. $f \in C^2(R^n)$ восполльзуемся формулой Ньютона лейбница для второй производной:

$$\phi'(1) = \phi'(0) + \int_0^1 \phi''(t)dt$$

по прежнему будем рассматривать $\phi(t) = f(x + t(y - x))$

$$\phi'(t) = f'(x + t(y - x))(y - x)$$
$$\phi''(t) = (f''(x + t(y - x))(y - x))(y - x)$$

Тогда с учетом формулы Н-Л имеем:

$$f'(y)(y - x) = f'(x)(y - x) + \int_0^1 (f''(x + t(y - x))(y - x), (y - x))dt$$

сократим общий множитель и получим:

$$f'(y) = f'(x) + \int_0^1 (f''(x + t(y - x)), (y - x)dt$$

ч.т.д.

3. Задача 3.

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

Решение:

$$f'(x) = \frac{f(x + y) - f(y)}{y} \xrightarrow{y \to 0} Ax + a$$

По определению градиента:

$$\nabla f(x)|f(y) = f(x) + (\nabla f(x), y - x) + 0(||y - x||)$$

Than lets consider the following equation:

$$f(x + y) = \alpha + (a, x + y) + \frac{1}{2}(A(x + y), x + y) =$$

$$\alpha + (a, x) + \frac{1}{2}(Ax, x) + (Ax + a, y) + \frac{1}{2}(Ay, y) =$$

$$f(x) + (Ax + a, y) + \frac{1}{2}(Ay, y) =< z = x + y >$$

Finally we get:

$$f(z) = f(x) + (Ax + a, z - x) + \frac{1}{2}(A(z - x), x - z)$$

Comparing the derived equation with the gradient definition we get:

$$\nabla f(x) = Ax + a$$

Similar to the previous section

$$f''(x) = A$$

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

$$a)(x, Ay) = (A^T x, y);$$
$$b)\phi'(x) = A^T f'(Ax + b).$$

Решение:

1. Вспомним определение скалярного произведения в виде матричного произведения: $(x, y) = y^T x$

Преобразуем скалярные произведения:

$$(x, Ay) = (Ay)^T x = y^T A^T x$$

$$(A^T x, y) = y^T A^T x$$

т.о. получаем, что $(x, Ay) = (A^T x, y$, ч.т.д.

2. $\nabla\phi(x) = A^T \nabla f(Ax + b)$ т.к. $\nabla f(Ax + b)$ - вектор размера $m \times 1$, а матрица A размера $m \times n$, то для того, чтобы произведение прошло правильно,

нужно поставить матрицу A слева и транспонировать. $\nabla\phi(x)$ - это вектор размера $n \times 1$ так как зависит только от x: $n \times 1 = (m\ timesn)^T * m \times 1 = n \times m * m \times 1$, ч.т.д.

5. Задача 5.

Пусть $f(x) \in C_M^{2,2}(R^n)$. Тогда имеет место неравенство:

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

Доказательство:

Из формулы Ньютона-Лейбница следует:

$$\phi(1) = \phi(0) + \int_0^1 \phi'(t)dt = \phi(0) + \phi'(0) - \phi'(0) + \int_0^1 \phi'(t)dt =$$

$$\phi(0) + \phi'(0) + \int_0^1 (\phi'(t) - \phi'(0))dt = \phi(0) + \phi'(0) + \int_0^1 \int_0^\tau \phi''(\tau)d\tau dt$$

Рассмотрим такую функцию: $\phi(t) = f(x + t(y - x))$, тогда получим:

$$f(y) = f(x) + (f'(x), y - x) + \frac{1}{2}\int_0^1 \int_0^\tau (f''(x - \tau(y - x))(y - x), y - x))d\tau dt$$

Имеем:

$$|f(y) - f(x) - (f'(x), y - x) - \frac{1}{2}(f''(x)(y - x), y - x)| =$$

$$|\int_0^1 \int_0^\tau (f''(x + \tau(y - x) - f(x))(y - x), y - x)d\tau dt| \leq$$

$$\int_0^1 \int_0^\tau ||(f''(x + \tau(y - x) - f(x))(y - x), (y - x)||d\tau dt =$$

$$\int_0^1 \int_0^\tau M\tau||y - x||^3 d\tau dt = \frac{M}{6}||y - x||^3$$

Таким образом получаем:

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

ч.т.д.

4 Пачка 4

1. Задача 1.
Из начального приближения $x^0 = (-2, 1)$ сделать два шага метода скорейшего спуска для задачи безусловной оптимизации

$$f(x_1, x_2) = x_1^2 + x_1 x_2 + 2x_2^2 + x_1 \to min$$

Решение:

Нам нужно показать, что делая шаги оптимизации методом градиентного спуска мы с каждым шагом приближаемся к оптимальной точке x^*

Для этого найдем x^*:

$$\nabla f(x) = (2x_1 + x_2 + 1, x_1 + 4x_2) = (0, 0)$$

получаем точки: $x_1 = -4/7$ и $x_2 = 1$
Теперь запишем формулу градиентного спуска:

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

Найдем расстояние между оптимальной точкой и начальной точкой в начале алгоритма:

$$||x^{(0)} - x^*|| = ||(-2 + 4/7, 1 - 1/7)|| = \frac{1}{7}||(-10, 6)|| = \frac{1}{7}\sqrt{136}$$

Положим $\alpha = 0.5$ и сделаем первый шаг градиентного спуска:

$$x^{(1)} = x^{(0)} - 0.5\nabla f(x^{(0)}) = (-2, 1) - 0.5(-2, 2) = (-1, 0)$$

Рассчитаем расстояние до оптимальной точки после первого шага:

$$||x^{(1)} - x^*|| = ||(-1 + 4/7, -1/7)|| = \frac{1}{7}||(-3, -1)|| = \frac{1}{7}\sqrt{10}$$

Теперь сделаем второй шаг оптимизации:

$$x^{(2)} = x^{(1)} - 0.5\nabla f(x^{(1)}) = (-1, 0) - 0.5(-1, -1) = (-0.5, 0.5)$$

Рассчитаем расстояние до оптимальной точки после второго шага:

$$||x^{(2)} - x^*|| = ||(-\frac{1}{2} + \frac{4}{7}, \frac{1}{2} - \frac{1}{7})|| = ||(\frac{1}{14}, \frac{5}{14})|| = \frac{1}{14}\sqrt{26}$$

Таким образом получаем:

$$\frac{1}{7}\sqrt{136} > \frac{1}{7}\sqrt{10} > \frac{1}{14}\sqrt{26}$$

$$||x^{(0)} - x^*|| > ||x^{(1)} - x^*|| > ||x^{(2)} - x^*||$$

т.е. с каждым шагом мы приближаемся к оптимальной точке.

2. Задача 2

Из начального приближения $x^{(0)} = (1, -1)$ сделать шаг градиентного метода по правилу Голдштейна-Армийо (подобрать параметры α, β) для задачи беусловной оптимизации

$$f(x_1, x_2) = 2x_1^2 + x_1x_2 + 3x_2^2 \to min$$

Решение:

Сначала найдем оптимальную точку, для этого рассчитаем градиент:

$$\nabla f(x_1, x_2) = (4x_1 + x_2, 6x_2 + x_1) = (0, 0)$$

Получаем точки: $x^* = (1/23, -4/23)$ - точка глоб мин, т.к. $\nabla^2 f(x) > 0$

Правило Голдштейна-Армийо помогает рассчитать шаг для градиентного спуска для текущей итерации: $x_{k+1} = x_k - h\nabla f(x_k)$:

$$\begin{cases} \alpha(f'(x_k), x_k - x_{k+1}) \leq f(x_k) - f(x_{k+1}) \\ \beta(f'(x_k), x_k - x_{k+1}) \geq f(x_k) - f(x_{k+1}) \end{cases}$$

Заметим, что $x_k - x_{k+1} = h\nabla f(x_k)$, тогда получим:

$$\begin{cases} \alpha h||\nabla f(x_k)||^2 \leq f(x_k) - f(x_{k+1}) \\ \beta h||\nabla f(x_k)||^2 \geq f(x_k) - f(x_{k+1}) \end{cases}$$

Будем выбирать гиперпараметры так: $\alpha = \rho$, $\beta = (1 - \rho)$

Тогда получаем условие на изменение функции:

$$\rho h||\nabla f(x_k)||^2 \leq f(x_k) - f(x_{k+1}) \leq (1 - \rho)h||\nabla f(x_k)||^2$$

Рассчитаем $x^{(1)} = x^{(0)} - h\nabla f(x^{(0)}) = (1, -1) - h(3, -5) = (1 - 3h, -1 + 5h)$

Тогда $f(x^{(1)} = 2(1 - 3h)^2 + (1 - 3h)(5h - 1) + 3(5h - 1)^2 = 93h^2 - 49h + 4$

при этом $f(x^{(0)} = 4$ и $||\nabla f(x^{(0)})||^2 = ||(3, -5)||^2 = 34$

Тогда получаем:

$$34\rho h \leq h(49 - 93h) \leq 34(1 - \rho)h$$

Положим $\rho = 0.3$:

$$0.271 = \frac{49 - 0.7 * 34}{93} \leq h \leq \frac{49 - 0.3 * 34}{93} = 0.41$$

Возьмем $h = 0.35$ тогда $x^{(1)} = (-0.05, 0.75)$

Сравним расстояния до оптимальной точки:

$$1.26 = ||x^{(0)} - x^*|| = \frac{||(22, -19)||}{23} < \frac{||(-43, 115)||}{460} = ||x^{(1)} - x^*|| = 0.26$$

т.е. после первого шага мы приблизились к оптимальной точке

3. Задача 3.

Из начального приближения $x^{(0)} = 1$ сделать два шага методом Ньютона для уравнения $x^p = 0$, где $p \geq 2$ - целочисленный параметр. Проанализировать скорость сходимости, объяснить наблюдаемый эффект. Модифицировать метод Ньютона, введя параметр длины шага, равный p. Проанализировать скорость сходимости.

Решение:

4. Задача 4.

Из начального приближения $x^{(0)} = (1, 1)$ сделать два шага метода Ньютона для задачи безусловной оптимизации:

$$f(x_1, x_2) = x_1^2 + e^{x_2^2} \to min$$

Проанализировать сходимость. Показать, что в данном случае метод Ньютона сходится глобально.

Решение:

5. Задача 5.

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

Решение:

5 Пачка 5