# Computational Methods of Optimization
# First Midterm(Sep 18 , 2024)

Time: 70 minutes

## Instructions

- Answer all questions
- Please write your answer in the spaces provided. Answers written outside will not be graded.
- Rough work can be done in designated spaces.
- This is a closed book exam.

Name: ______________________________

SRNO: Degree: Dept:

| Question: | 1 | 2 | 3 | 4 | 5 | Total |
|---|---|---|---|---|---|---|
| Points: | 10 | 10 | 10 | 10 | 5 | 45 |
| Score: | | | | | | |

In the following, let $\mathbf{I} = [\mathbf{e}_1, \ldots, \mathbf{e}_d]$ be a $d \times d$ matrix with $e_j$ be the $j$th column. Also $\mathbf{x} = [x_1, x_2, \ldots, x_d]^\top \in \mathbb{R}^d$ and $\|\mathbf{x}\| = \sqrt{\mathbf{x}^\top \mathbf{x}}$. Set of real symmetric $d \times d$ matrices will be denoted by $\mathcal{S}_d$. $[n]$ will denote the set $\{1, 2, \ldots, n\}$ For a function $f : \mathbb{R}^d \to \mathbb{R}, f \in \mathcal{C}^1_L$ the Algorithm `DESCEENT` will refer to the following iteration

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{u}^{(k)} \qquad (\texttt{DESCENT})$$

where $\mathbf{u}^{(k)\top} \nabla f(\mathbf{x}^{(k)}) < 0$.

1. Let $f : \mathbb{R} \to \mathbb{R}, f(x) = \frac{3}{2}x^4 - 2x^3 + 1$.
   (a) (4 points) Find all points such that $f'(x) = 0$.

**Solution:**

$$f'(x) = 6x^2(x-1) = 0, \implies x = 0, x = 1$$

   (b) (6 points) For each of the points found in the above question identify if it is a local or global minima or maxima.

**Solution:** Check that $f''(x) = 18x^2 - 12x = 0$. Thus $f''(0) = 0, f''(1) = 6$. It is clear that 1 is a local minimum. It is observed that $f(x) = x^3(\frac{3}{2}x - 2) + 1$ is positive for $x < 0$ and $x > 2$. Thus $\underset{|x|\to\infty}{lim} f(x) = \infty$. Hence there is no global maximizer. Also it is a coercive function. Since $x = 1$ is a local minimum it is also the global minimum. But $x = 0$ is an inflection point which is neiteher a local minimum or a local maximum. $f(0) = 1$. For any $0 < x < 1, f(x) < 1$ but for $x < 0, f(x) > 1$.

**Rough Work**

2. It is given that $\mathbf{a} = [a_1, \ldots, a_d]^\top, a_i \in \mathbb{R}, a_i \neq 0, i \in \{1, \ldots, d\}$ be such that

$$\sum_{i=1}^{d} a_i = 10, \quad \frac{a_{max}}{a_{min}} = 5,$$

where $a_{max} = max_i a_i, a_{min} = min_i a_i$ respectively. Consider $f : \mathbb{R}^d \to \mathbb{R}$,

$$f(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^{d} \frac{x_i^2}{a_i} - \sum_{i=1}^{d} x_i$$

Let $\hat{\mathbf{x}}$ be such that $\nabla f(\hat{\mathbf{x}}) = 0$ and let $f(\widetilde{\mathbf{x}}) = -20$, where $\widetilde{\mathbf{x}} = \mathbf{a} + \alpha \mathbf{e}_2$ for some positive $\alpha$.

(a) (1 point) Find $f(\hat{\mathbf{x}}) =$ **-5**

(b) (2 points) Mark all correct answers

A. $\hat{\mathbf{x}}$ is the global minimum
**B. $\hat{\mathbf{x}}$ is not the global minimum**
C. $a_2 > 0$
**D.** $a_2 < 0$

(c) (2 points) Suppose $\widetilde{\mathbf{x}}$ did not exist, but it was given that $f$ was bounded from below. How will your answer change to the previous question. More specifically, mark all correct answers.

**A. $\hat{\mathbf{x}}$ is the global minimum**
B. $\hat{\mathbf{x}}$ is not the global minimum
C. $a_2 < 0$
**D.** $a_2 > 0$

(d) (1 point) Assuming that $f$ is bounded from below, find $\rho$ such that

$$f(\mathbf{x}^{(k+1)}) - f(\hat{\mathbf{x}}) \leq \rho(f(\mathbf{x}^{(k)}) - f(\hat{\mathbf{x}}))$$

holds for any $k$. $\rho =$ $\frac{4}{9}$ (Answer must be in fraction).

(e) (4 points) Justify your answers to the previous questions

**Solution:**
Equating the gradient of $f$ to zero we obtain $\hat{\mathbf{x}} = \mathbf{a}$ and using $\sum_{i=1}^{d} a_i = 10$.

$$f(\hat{\mathbf{x}}) = -\frac{1}{2} \sum_{i=1}^{d} a_i = -5.$$

Therefore if there exists $\tilde{\mathbf{x}} = -20$ then $\hat{\mathbf{x}}$ cannot be a global minimum. Since now that the Hessian is constant, in fact it is a diagonal matrix, with the $i$th entry $\frac{1}{a_i}$. If it was positive definite then $\hat{\mathbf{x}}$ would have been the global minimum. Since $\nabla f(\mathbf{a}) = 0$, it follows that for any $\mathbf{x} \in \mathbb{R}^d$,

$$f(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^{d} \frac{1}{a_i} (x_i - a_i)^2 + f(\mathbf{a})$$

Note that this can be proved from first principles ($x_i^2 = (x_i - a_i)^2 + 2a_i x_i - a_i^2$). Hence, $f(\tilde{\mathbf{x}}) = f(\mathbf{a}) + \frac{1}{2} \frac{(\alpha - a_2)^2}{a_2}$. Given the values this can only be achieved if $a_2 < 0$. Instead of existence of $\hat{\mathbf{x}}$, if we knew $f$ is bounded from below and since $f$ is a quadratic function all eigenvalues must be non-negative. Otherwise we can find a direction for which $a_i$ is negative for $f$ to go to $-\infty$. Last question follows from the condition number of the Hessian which is 5.

3. Let $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} - \mathbf{v}^\top\mathbf{x} + c$ where $\mathbf{v} \in \mathbb{R}^d, c = 10, Q \succ 0$. The condition number of $Q$ is 10. It is given that $Q\mathbf{v} = 2\mathbf{v}, \quad \|\mathbf{v}\| = 10$. Let $\mathbf{x}^*$ be the global minimum point and $f^* = f(\mathbf{x}^*)$ is the global minimum value

   (a) i. (1 point) $\mathbf{x}^* =$ $\underline{\quad \frac{1}{2}\mathbf{v} \quad}$

   ii. (1 point) $f^* =$ $\underline{\quad \textbf{-15} \quad}$

   iii. (4 points) Justify your answer

   **Solution:**

$$Q\mathbf{x}^* = \mathbf{v}, \quad \implies \mathbf{x}^* = Q^{-1}\mathbf{v} = \frac{1}{2}\mathbf{v}$$

$$f(\mathbf{x}^*) = \frac{1}{2} \times \frac{1}{4}\mathbf{v}^\top Q\mathbf{v} - \frac{1}{2}\mathbf{v}^\top\mathbf{v} + 10$$

   $f(\mathbf{x}^*) = 10 - \frac{1}{2}\|\mathbf{v}\|^2 = 10 - 25 = -15$

   (b) (4 points) Let us execute the Algorithm `DESCENT` with $\mathbf{u}^{(k)} = -\nabla f(\mathbf{x}^{(k)})$ and $\alpha_k$ is determined with exact line search starting at $\mathbf{x}^{(0)} = 0$. In how many iterations will the algorithm reach the global minimum point. Justify your answer.

   **Solution:** It will reach in 1 iteration.

$$\nabla f(0) = -\mathbf{v}, \quad \mathbf{x}^{(1)} = -\alpha\nabla f(0) = \alpha\mathbf{v}.$$

$$\alpha_k = \underset{\alpha \geq 0}{\operatorname{argmax}} \, f(\alpha\mathbf{v})$$

$$\alpha_k = \underset{\alpha \geq 0}{(} \ \alpha^2 - \alpha)\|\mathbf{v}\|^2 = \frac{1}{2}$$

   (By using the fact that $Q\mathbf{v} = 2\mathbf{v}$.) Thus $\mathbf{x}^{(1)} = \frac{1}{2}\mathbf{v}$.

**Rough Work**

4. Let $f : \mathbb{R}^d \to [0, \infty), f \in \mathcal{C}_L^1$ with $L = 2$. Consider applying Algorithm `DESCENT` to minimize $f$ where $\alpha_k$ is obtained through inexact line search in such a way that the following holds

$$\frac{1}{\|\nabla f(\mathbf{x}^{(k)}\|^2}(f(\mathbf{x}^{(k)}) - f(\mathbf{x}^{(k+1)})) \geq 0.2, \text{ for all } k$$

It is given that $f(\mathbf{x}^{(0)}) = 20$.

(a) (4 points) We wish to determine the value of $T$ for which we are sure to find $\hat{\mathbf{x}}$ such that $\|\nabla f(\hat{\mathbf{x}})\| \leq 0.1$. Pick all the correct choices about value of $T$

A. $T = 1000$

B. $T = 2000$

**C.** $T = 10000$

**D.** $T = 20000$

**Solution:** 5cm

$$\|\nabla f(\mathbf{x}^{(T)}\|^2 \leq \frac{f(\mathbf{x}^{(0)})}{0.2T} \leq (0.1)^2$$

$$\frac{100}{(0.1)^2} \leq T$$

(b) Instead of the above we decide to use the constant stepsize, $\alpha_k = \alpha$ with $u^{(k)} = -\nabla f(\mathbf{x}^{(k)})$.

i. (2 points) Find $h(\alpha)$ such that at the $k$th iteration Decrease is atleast greater than $h(\alpha)\|\nabla f(\mathbf{x}^{(k)}\|^2$

**Solution:**

$$h(\alpha) = \alpha(1 - \alpha)$$

ii. (2 points) Find $\alpha$ which gives the maximum decrease in every iteration. $\alpha =$ <u>$\frac{1}{2}$</u>.

iii. (2 points) How many iterations are required to find a point $\hat{\mathbf{x}}$ such that $\|\nabla f(\hat{\mathbf{x}}\| \leq 0.1$

**Solution:**

$$\frac{1}{4}T\|\nabla f(\hat{\mathbf{x}})\|^2 \leq 20$$

Thus $\frac{80}{T} \leq (0.1)^2$

5. Let $f : \mathbb{R}^d \to \mathbb{R}, f \in C_L^1$ and $\mathbf{x}^*$ be a point such that $\nabla f(\mathbf{x}^*) = 0$. We apply Algorithm `DESCENT` to this function with inexact line search to obtain iterates $\mathbf{x}^{(k)}$ Let $\Delta_k = f(\mathbf{x}^{(k)} - f(\mathbf{x}^*)$. Answer True or False

(a) (1 point) For the Algorithm `DESCENT` there exists $C > 0$ such that $\Delta_k - \Delta_{k+1} \geq C\|\nabla f(\mathbf{x}^{(k)}\|^2$ **F**

(b) (1 point) In the above question if $\mathbf{u}^{(k)} = -\nabla f(\mathbf{x}^{(k)}$ then the value of $C$ increases **T**

(c) (1 point) If $f$ is coercive then all minimum must be global minimum **F**

(d) (1 point) If $f$ is strongly convex it can be shown that $\Delta_k - \Delta_{k+1} \geq C\Delta_k^2$ **F**

(e) (1 point) If $f \in C^2$ and $H(\mathbf{x}^*)$, the Hessian, is positive definite then $\mathbf{x}^*$ must be global minimum **F**

**Rough Work**

**Rough Work**

**Rough Work**

**Rough Work**

**Rough Work**