Computational Methods of Optimization
Final Exam-Part 1(25th Jan,2021)

Start Time: 9:15 AM End Time: 10:25 AM

Instructions

- Answer all questions
- See upload instructions in the form

Question:	1	2	3	4	Total
Points:	10	10	10	10	40
Score:					

In the following $\|\mathbf{x}\| = \sqrt{\mathbf{x}^\top \mathbf{x}}$. For a real valued function, f, in one variable, f' denotes the first derivative, $f"$ denote the second derivative, and $f^{(3)}$ denotes the third derivative.

1. (a) Let $f : [-1, 2] \to \mathbb{R}$ be a differentiable function.
   i. (2 points) Suppose it was known that $f(1) = -f(0) = 1$. Which of the following is correct A. $|f'(x)| \leq 1.5$ for all $x \in (0, 1)$. B. $|f'(x)| \geq 1.5$ for all $x \in (0, 1)$. C. $|f'(x)| = 2$ for some $x \in (0, 1)$. D. None of the above
   ii. (2 points) Suppose $f(0.5) = f(0.8)$ and $f"(x) > 0\ \forall x \in (-1, 2)$. A. There are no minima in $[-1, 2]$. B. There is exactly one minimum in $[-1, 2]$ C. There is at-least one minimum in $[-1, 2]$ D. None of the above
   iii. (2 points) Let f attain minimum at $x = 2$. Which of the following is true A. $f(x) \geq f(2)$ for all $x \in \mathbb{R}$ B. $f'(x) \leq 0$ for all $x \in [-1, 2]$ C. $f'(x) \geq 0$ for all $x \in [-1, 2]$

   (b) (4 points) Consider minimizing a convex quadratic function whose Hessian has largest and smallest eigenvalue 3 and 1 respectively. Suppose we implement the steepest descent procedure starting at a point $\mathbf{x}^{(0)}$ such that $E(\mathbf{x}^{(0)} = 1$ where $E(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)H(\mathbf{x} - \mathbf{x}^*)$ where $\mathbf{x}^*$ is the global minimum. After how many iterations can you guarantee that $\|\mathbf{x}^{(T)} - \mathbf{x}^*\| \leq 10^{-2}$.

2. Let $f : (a, b) \to \mathbb{R}$ be thrice differentiable function such that $f'(a)f'(b) < 0$. Assume that for all $x \in (a, b), |f"(x)| \geq \beta, |f^{(3)}(x)| \leq \alpha$ where $\beta, \alpha > 0$.

   (a) i. (1 point) The Number of critical points in (a, b) is
   A. 1 B. 2 C. 3 D. 4
   ii. (2 points) Justify your answer

   (b) Consider Newton iterates

$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f"(x^{(k)})}$$

   It can be shown that there exists

$$e_{k+1} \leq Ce_k^2 \ \ e_k = |x^{(k)} - r|$$

i. (1 point) Choose the correct value of C from the following choices A. $\frac{\alpha}{\beta}$ B. $\frac{\beta}{\alpha}$ C. $\frac{\alpha}{2\beta}$ D. $\frac{\beta}{2\alpha}$

ii. (3 points) Justify your answer

(c) (3 points) Find $t > 0$ such that for any $x^{(0)} \in (r - t, r + t)$ the Newton iterates converge to r.

3. Let $f : C \subset \mathbb{R}^d \to \mathbb{R}$ be a differentiable function lowerbounded below and upperbounded by a function g as follows

$$f(\mathbf{y}) \leq g(\mathbf{y}; \mathbf{x}) \left(\equiv f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) + \frac{\beta}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right)$$

holds for all $\mathbf{x}, \mathbf{y} \in C$

(a) (2 points) Under what condition on $\mathbf{v}$

$$h(\alpha) = g(\mathbf{x}^{(k)} + \alpha \mathbf{v}; \mathbf{x}^{(k)}) - g(\mathbf{x}^{(k)}; \mathbf{x}^{(k)})$$

is strictly less than 0 for some $\alpha \geq 0$. For such a choice of $\mathbf{v}$ find

$$\alpha^* = min_{\alpha \geq 0} h(\alpha)$$

(b) (3 points) Set up an iterative scheme

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{v}^{(k)}$$

where $\mathbf{v}^{(k)}$ is chosen as in the previous question with $\mathbf{x} = \mathbf{x}^{(k)}$, and $\alpha_k = \alpha^*$. For such a choice find the smallest possible C_k such that

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

Your answer should mention C_k with a very brief justification.

(c) (5 points) Starting from arbitrary $\mathbf{x}^{(0)}$ and assuming that

$$\nabla f(\mathbf{x}^{(k)})^\top \mathbf{v}^{(k)} \geq \delta \|\nabla f(\mathbf{x}^{(k)}\|^2$$

holds for all $k = 0, 1, \ldots$, how many iterations will be required to find an $\hat{\mathbf{x}}$ such that $|\nabla f(\hat{\mathbf{x}})| \leq \epsilon$ for a given ϵ. (The answer should state the relationship between $T, \delta, \beta, f(\mathbf{x}^{(0)})$ and any other quantity you feel necessary.

4. Consider the Linear system equations $A\mathbf{x} = b$ where A is a $d \times d$ real valued matrix and b is a d-dimensional vector.

Define $\texttt{res}(\mathbf{x}) = b - A\mathbf{x}$. We wish to solve the system (A, b) with the following iterative procedure

- (Initialize)

$$\mathbf{u}^{(0)} = \texttt{res}(\mathbf{x}^{(0)})$$

- (Iterate)

$$\alpha_k = \frac{\|\texttt{res}(\mathbf{x}^{(k)})\|^2}{\mathbf{u}^{(k)\top} A \mathbf{u}^{(k)}}$$

- $$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{u}^{(k)}$$

- $$\beta_k = \frac{\|\texttt{res}(\mathbf{x}^{(k+1)})\|^2}{\|\texttt{res}(\mathbf{x}^{(k)})\|^2}$$

- $$\mathbf{u}^{(k+1)} = \texttt{res}(\mathbf{x}^{(k+1)}) + \beta_k \mathbf{u}^{(k)}$$

(a) (5 points) Why would this algorithm solve the linear system of equations?

(b) Consider two systems,(A_1, b) and (A_2, b) with same b but different matrices A_1 and A_2.

$$A_1 = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

i. (1 point) The Algorithm applies to only one of the systems. Which one and why?

ii. (4 points) Modify the algorithm so that it will apply to both the systems. State the modification and give a brief justification. Any modification should be under the same style of algorithms mentioned in a.