Computational Methods of Optimization
First Midterm(7th Dec, 2020)

Time: 60 minutes

Instructions

- Answer all questions
- See upload instructions in the form

In the following, assume that f is a $\mathcal{C}^1$ function defined from $\mathbb{R}^d \to \mathbb{R}$ unless otherwise mentioned. Let $\mathbf{I} = [e_1, \dots, e_d]$ be a $d \times d$ matrix with e_j be the jth column. Also $\mathbf{x} = [x_1, x_2, \dots, x_d]^\top \in \mathbb{R}^d$ 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 $\{1, 2, \dots, n\}$

1. (10 points) Please indicate True(T) or False(F) in the space given after each question. All questions carry equal marks
   (a) Let $a < b$ where $a, b \in \mathbb{R}$ and $h : [a, b] \to \mathbb{R}$ be differentiable and satisfies $h(a) = h(b)$. Then h has a critical point in (a, b). ____
   (b) Suppose the function defined in the previous question satisfies $|h(x) - h(y)| \leq 1|x - y|$ for all $x, y \in (a, b)$. There could exist a point in (a, b) such that $h'(x) \geq 2$, where $h'(x)$ is derivative of h at x. ____
   (c) If f is a coercive function then the global mimimum must lie at one of the critical points. Recall that a critical point is a point, $\mathbf{x}$, such that $\nabla f(\mathbf{x}) = 0$. ____
   (d) Consider $g : \mathbb{R} \to \mathbb{R}, g(u) = u^2 - \frac{1}{3}u^3$.The function has a global minimum. ____
   (e) The local maximum of g(defined in the previous question) is at $u = 0$. ____

2. Let $f : S \subset \mathbb{R}^d \to \mathbb{R} \in \mathcal{C}^2$ function. Let $H(\mathbf{x})$ be the Hessian of f with eigenvectors denoted by columns of $U \in \mathbb{R}^{d \times d}$
   (a) (2 points) Consider $S = \{\mathbf{x} | \mathbf{x} = \mathbf{x}^* + \mathbf{U}\mathbf{v}, \mathbf{v} \in \mathbb{R}^d\}$. $\bar{S}$, the complement of the set S, is not empty. True or False. Justify with reasons.

   (b) (4 points) The Hessian, $H(\mathbf{x}^*)$ at a stationary point $\mathbf{x}^*$ has one eigenvalue 0 and the rest positive. For any $\mathbf{x} \in S$ with $\mathbf{v} \neq 0$ $f(\mathbf{x}) > f(\mathbf{x}^*)$ is true. Prove or disprove

   (c) (4 points) Suppose $H(\mathbf{x}^*)$ has negative and positive eigenvalues. Construct two distinct points, $\mathbf{x}^1$ and $\mathbf{x}^2$ in terms of $\mathbf{U}$ so that
   $$f(\mathbf{x}^1) < f(\mathbf{x}) < f(\mathbf{x}^2)$$

3. Let $f(x) = \sqrt{1 + x^2}, x \in \mathbb{R}$.

(a) (2 points) Find the x^*, global minimum of the problem

(b) (4 points) Let $x^{(k)}$ be the output of the kth iteration of Newton's method applied to $f(x)$. Find a function g such that

$$|x^{(k+1)} - x^*| \leq |g(x^{(k)} - x^*)|$$

(c) (4 points) Using the above relationship find largest a so that the Newton's method is most effective for any $x^{(0)} \in (x^* - a, x^* + a)$.

4. Consider minimizing the function

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} - \mathbf{b}^\top \mathbf{x}$$

over $\mathbf{x} \in \mathbb{R}^d$ with $Q \in \mathcal{S}_d^+, \mathbf{b} \in \mathbb{R}^d$ using the steepest descent iterates,

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

executed with exact step-size selection strategy. Let $f^* = f(\mathbf{x}^*)$ be the global minimum. Let $g(\mathbf{y}) = f(A\mathbf{x})$ where A is a $d \times d$ natrix. Consider applying steepest descent procedure to g i.e.

$$\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} - \alpha_k \nabla g(\mathbf{y}^{(k)})$$

and g^* is the global minimum of g attained at $\mathbf{y}^*$.

(a) (4 points) State the Hessian of f and g. Derive the relationship between $\mathbf{x}^*$ and $\mathbf{y}^*$

(b) (3 points) What is the convergence rate of the steepest descent procedure for g?

(c) (3 points) What is the best value of A?

5. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a $C^{(2)}$ function. Consider the Quasi-newton update

   a. $s^{(k)} = -G^{(k)} \nabla f(\mathbf{x}^{(k)})$‘

   b. $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k s^{(k)}$

   c. $G^{(k+1)} = G^{(k)} + A^{(k)} E A^{(k)\top}$

   where $\gamma_k = \nabla f(\mathbf{x}^{(k+1)}) - \nabla f(\mathbf{x}^{(k)})$, $\delta_k = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$, and $A^{(k)} = [\delta_k \quad G^{(k)} \gamma^{(k)}]$ and E is a matrix.

   (a) (1 point) What is your name

   (b) (4 points) Find E corresponding to DFP updates.

   (c) (5 points) Show that for exact line search DFP updates yield positive definite matrices.