Computational Methods of Optimization
Final Exam(Nov 22, 2024)

Time: 120 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: ____________________

Index: SRNO: Degree:

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

Rough Work

In the following, let $\mathbf{I} = [\mathbf{e}_1, \ldots, \mathbf{e}_d]$ be a $d \times d$ matrix with e_j be the jth 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 $\{1, 2, \ldots, n\}$ Let

$$h(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top Q \mathbf{x} + b^\top \mathbf{x} + c \qquad \texttt{(QUAD)}$$

$\mathbf{x} \in \mathbb{R}^d, Q \in \mathcal{S}_d, b \in \mathbb{R}^d, c \in \mathbb{R}$. For a function $f : \mathbb{R}^d \to \mathbb{R}, f \in \mathcal{C}_L^1$ the Algorithm `DESCENT` 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. True or False.

   (a) (1 point) Let $h(\mathbf{x})$ be defined as in (`QUAD`). For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$

   $$h(\mathbf{y}) - h(\mathbf{x}) = \nabla h(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^\top Q (\mathbf{y} - \mathbf{x})$$

   is true only if Q is positive semi-definite or positive definite. **F**

   (b) (1 point) Let $\mathbf{u} \in \mathbb{R}^d, \mathbf{u} \neq 0, a > 0$. Consider minimizing $\mathbf{u}^\top \mathbf{x}$ subject to $\|\mathbf{x}\| \leq a$. Optimum value is reached at $\mathbf{x}^*$ where $\|\mathbf{x}^*\| = \frac{a}{2}$ for some $\mathbf{u}$. **F**

   (c) (1 point) Consider running the descent algorithm in (`DESCENT`) by setting $\mathbf{u}^{(k)} = -\nabla f(\mathbf{x}^{(k)}), \alpha_k = \frac{1}{L}$ with $f(\mathbf{x}) = h(\mathbf{x})$ where $h(\mathbf{x})$ is defined by `QUAD`. Starting from an arbitrary point $\hat{\mathbf{x}}$ after 5 iterations, we can be assured that

   $$h(\mathbf{x}^{(5)}) - h(\mathbf{x}^*) \leq (\frac{3}{4})^5 \, (h(\hat{\mathbf{x}}) - h(\mathbf{x}^*))$$

   Assume the eigenvalues of Q are all greater than 1 but less than 7. And assume $\mathbf{x}^*$ is the global minimum. **F**

   (d) (1 point) Conjugate Gradient algorithm finds the global unconstrained minimum of h in atmost d iterations **T**.

   (e) (1 point) Convergence of Newton's method for minimizing h when the Hessian is positive definite depends on the eigen-values of Q. **F**.

   (f) (1 point) Let $\mathbf{z} \notin C$ where $C \subset \mathbb{R}^d$ is a closed convex set. There is only one $\mathbf{x}^* \in C$ such that $\|\mathbf{x}^* - \mathbf{z}| < \|\mathbf{x} - \mathbf{z}\|$ for all $\mathbf{x} \in C$. **T**

   (g) (1 point) For a constrained optimization problem with objective function $f(\mathbf{x})$ it was found that the at $\mathbf{x}^*$,the optimal point, $\nabla f(\mathbf{x}^*) \neq 0$. Then at least one of the constraints must be active **T**

   (h) (1 point) The set of Feasible Descent Directions at $\hat{\mathbf{x}}$ is not empty if $\hat{\mathbf{x}}$ is a KKT point.

   (i) (1 point) Gradient projection algorithm in finite number of iterations will find the global minimum for a convex programming problem **F**

   (j) (1 point) For a given constrained problem with objective function f it was found that there exists a dual feasible point with dual objective function value -1 and a primal feasible point, $\bar{\mathbf{x}}$ with primal point $f(\bar{\mathbf{x}}) = .5$. Then $f(\bar{\mathbf{x}}) - f(\mathbf{x}^*) \leq 1.5$ **T**

Rough Work

Rough Work

2. Let $f \in \mathcal{C}_L^1$ over $\mathbb{R}^d$ and convex with $L = 5$. For some $\mu > 0$, the function f satisfies the following for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$.

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) + \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|^2$$

Consider minimizing f over $\mathbb{R}^d$ using (`DESCENT`) using $\mathbf{u}^{(k)} = -\nabla f(\mathbf{x}^{(k)}), \alpha_k = \alpha$. Let $\mathbf{x}^*$ be the global minimum

(a) (2 points) Suppose $\mathbf{x}^{(0)}$ be a point such that $\|\nabla f(\mathbf{x}^{(0)})\| = \sqrt{2}$. Find an upper bound on $\|\mathbf{x}^{(0)} - \mathbf{x}^*\|^2$. Credit will be given to the lowest upper bound.

Solution: Minimizing both sides we obtain

$$f(\mathbf{x}^{(0)}) - f(\mathbf{x}^*) \leq \frac{1}{2}\|\nabla f(\mathbf{x}^{(0)}\|^2 = 1$$

Also since $\nabla f(\mathbf{x}^*) = 0$,

$$f(\mathbf{x}^{(0)}) - f(\mathbf{x}^*) \geq \frac{1}{2}\|\mathbf{x}^{(0)} - \mathbf{x}^*\|^2.$$

(b) (2 points) At each iteration find a lower bound on the decrease, $f(\mathbf{x}^{(k)} - f(\mathbf{x}^{(k+1)})$ in terms of $\|\nabla f(\mathbf{x}^{(k)}\|^2$. Find the range of α where decrease is guaranteed?

Solution: The relationship is

$$f(\mathbf{x}^{(k)} - f(\mathbf{x}^{(k+1)}) \geq \alpha(1 - \frac{L}{2}\alpha)\|\nabla f(\mathbf{x}^{(k)}\|^2$$

Thus range of $\alpha \in (0, \frac{2}{5})$.

Rough Work

(c) (3 points) Let $\Delta_k = f(\mathbf{x}^{(k)}) - f(\mathbf{x}^*)$. After T iterations derive a bound, more specifically find A and B

$$\log \frac{\Delta_T}{\Delta_0} \leq A + BT?$$

Solution:

$$\Delta_k - \Delta_{k+1} \geq \alpha(1 - \frac{5}{2}\alpha)\|\nabla f(\mathbf{x}^{(k)})\|^2$$

$$\Delta_k \leq \frac{1}{2}\|\nabla f(\mathbf{x}^{(k)})\|^2$$

Define $\beta = 2\alpha(1 - \frac{5}{2}\alpha)$

$$\Delta_{k+1} \leq (1-\beta)\Delta_k \leq e^{-\beta}\Delta_k$$

$$\Delta_T \leq e^{-T\beta}\Delta_0$$

Thus $A = 0$ and $B = -\beta$

(d) (3 points) Similarly find A, B such that $\log \|\mathbf{x}^{(T)}) - \mathbf{x}^*\|^2 \leq A + BT$. ?

Solution: For all k

$$\frac{1}{2}\|\mathbf{x}^{(k)} - \mathbf{x}^*\|^2 \leq \Delta_k$$

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\|^2 \leq 2e^{-T\beta}\Delta_0 \leq 2e^{-T\beta}\frac{1}{2}\|\nabla f(\mathbf{x}^{(0)}\|^2$$

Substituting the value of the norm of the gradient we obtain $A = \log 2, B = -\beta$ as defined earlier.

Rough Work

3. Consider $f : \mathbb{R}^3 \to \mathbb{R} \in \mathcal{C}_L^1$ a convex function and the problem,

$$min_{\mathbf{x}\in\mathbb{R}^3} \quad f(\mathbf{x})$$

$$\text{s.t } \frac{1}{3}(x_1x_2 + x_3x_2 + x_1x_2) \leq 1.5, \quad x_1 + ax_2 \geq b, \quad x_2 + x_3 \leq 10, x_3 \geq 0.5$$

Suppose you are given a point $\hat{\mathbf{x}} = [1,1,1]^\top$ and oracle access to f, namely you have the value $f(\hat{\mathbf{x}}) = 3, \nabla f(\hat{\mathbf{x}}) = [1,10,0]^\top$.

(a) (5 points) Can you choose $a, b \in \mathbb{R}$ so that $\hat{\mathbf{x}}$ is a KKT point.

Solution: $\hat{\mathbf{x}}$ is feasible. Since $\nabla f(\hat{\mathbf{x}})$ is not zero it means that there must be atleast one active constraint. Only possible candidate is the constraint $x_1 + ax_2 \geq b$. To satisfy KKT conditions

$$\nabla f(\hat{\mathbf{x}}) = \lambda[1, a, 0]^\top$$

Thus $\lambda = 1, \lambda a = 10$. Hence $a = 10$. Since the constraint is active $b = 1 + 10 \cdot 1 = 11$.

(b) (5 points) The global optimal value of f is strictly less than 3. __**F**__. Justify

Solution: Consider the problem

$$min_{\mathbf{x}\in\mathbb{R}^3} \quad f(\mathbf{x}), \quad \text{s.t. } x_1 + 10x_2 \geq 11$$

Since $\hat{\mathbf{x}}$ is a KKT point for this problem and hence it is globally optimal. Since $\hat{\mathbf{x}}$ is also feasible for the original problem which has more constraints, it must be the case $f(\hat{\mathbf{x}})$ must be less than or equal to the global optimum of the original problem.

Rough Work

Rough Work

4. (5 points) Let $f, g_1, g_2, g_3 : \mathbb{R}^3 \to \mathbb{R}$ be $\mathcal{C}^{(1)}$ functions. Consider the following problem.

$$min_{\mathbf{x}} f(\mathbf{x}) \text{ subject to } g_i(\mathbf{x}) \geq 0, i \in \{1, 2, 3\}$$

However the inequalities maybe in the wrong direction. Given the following partial information about a point $\mathbf{x}^{(0)}$ we need to determine if the inequalities are in the proper directions.

- $\mathbf{x}^{(0)}$ is a KKT point
- $g_1(\mathbf{x}^{(0)}) = 0,\ g_2(\mathbf{x}^{(0)}) = 0,\ g_3(\mathbf{x}^{(0)}) = -4$
- $\nabla f(\mathbf{x}^{(0)}) = [4, -12, 4]^\top, \nabla g_1(\mathbf{x}^{(0)}) = [1, -2, 0]^\top, \nabla g_2(\mathbf{x}^{(0)}) = [0, 1, -2]^\top, \nabla g_3(\mathbf{x}^{(0)} = [-1, 1, 10]^\top$

From the above information can you decide if the inequalities are in the proper direction? If they are not please correct them? Give reasons for your answer.

Solution: The inequalities g_2, g_3 are not in the proper direction. Since $\mathbf{x}^{(0)}$ is a KKT point, it is also feasible. This implies that $g_3(\mathbf{x}) \leq 0$. Hence

$$\nabla f(\mathbf{x}^{(0)}) = \lambda_1 \nabla g_1(\mathbf{x}^{(0)}) + \lambda_2 \nabla g_2(\mathbf{x}^{(0)}) + \lambda_3 \nabla g_3(\mathbf{x}^{(0)})$$

$\lambda_3 = 0$ since g_3 is not active. Assuming that the inequalities are in the correct direction both the λ_i are positive. However the given values indicate $\lambda_1 = 4, \lambda_2 = -2$ is the only possibility. This indicates that g_2 is also in the wrong direction.

Rough Work

Rough Work

5. Let $\mathbf{u} \in \mathbb{R}^d$ such that all coordinates are distinct non-zero values with $a = min_i u_i, b = max_i u_i$. Consider the linear programming problem

$$max_{\mathbf{x}\in\mathbb{R}^d} \mathbf{u}^\top \mathbf{x}, \text{ subject to } \sum_{i=1}^{d} x_i \leq 1,\ x_i \geq 0, i \in \{1, \ldots, d\}$$

(a) (5 points) Suppose it is given that the constraint $\sum_{i=1}^{d} x_i \leq 1$ is not active at optimality. What conclusions can you draw about a, b and the optimal value.

Solution: This is a convex problem and hence a KKT point is sufficient for optimality. Since it is given that the constraint is not active then from the KKT conditions

$$-u_i = \lambda_i \quad \forall i \in [d]$$

where $\lambda_i \geq 0$. This implies that $u_i \leq 0$. Since none of the coordinates are 0 this implies that $a < b < 0$ and furthermore due to KKT conditons $\lambda_i x_i = 0 \implies x_i = 0$. Hence $\mathbf{u}^\top \mathbf{x} = 0$ at optimality.

(b) (5 points) Repeat the same question supposing that the constraint $\sum_{i=1}^{d} x_i \leq 1$ is active at optimality.

Solution: The KKT conditions yield the following

$$-u_i = -\lambda + \lambda_i \quad \forall i \in [d]$$

where $\lambda, \lambda_i \geq 0$. All $\lambda_i > 0$ for all i is not possible as it will force all $x_i = 0$ which render the constraint to be inactive. Thus there exists at least one l such that $\lambda_l = 0$. For such l, $u_l = \lambda$ and hence $b > 0$. Also note that $\lambda \geq u_i$ for all i. Thus at KKT point $\lambda = b, \lambda_l = 0, \lambda_i = 0, i \neq l$. At optimality then $x_l = 1$ and $x_i = 0, i \neq l$. Nothing can be said about a.

Rough Work

Rough Work

Rough Work

Rough Work

Rough Work

Rough Work