Computational Methods of Optimization
Third Midterm(Nov 13, 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: ______________________________

Index: SRNO: Degree: Dept:

Question:	1	2	3	4	Total
Points:	5	15	15	5	40
Score:					

1. Answer True or False

   (a) (1 point) The set $\{\mathbf{x} \in \mathbb{R}^d | \|\mathbf{x} - \mathbf{a}\| > r^2\}$ is convex. Given $\mathbf{a} \in \mathbb{R}^d, r \in \mathbb{R}\}$ are some fixed constants. __**F**__.

   (b) (1 point) Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function. The set $\{\mathbf{x} | f(\mathbf{x}) \leq f(\mathbf{x}^{(0)}\}$ is convex. Assume any $\mathbf{x}^{(0)} \in \mathbb{R}^d$ in __**T**__.

   (c) (1 point) Separating Hyperplane theorem applies only to halfspaces. __**F**__

   (d) (1 point) In the general setting of Active set method once the Working set size starts decreasing it cannot increase any more __**F**__

   (e) (1 point) Lagrange Dual and Wolfe dual gives the same Dual problem for convex optimization problem __**T**__

2. Consider the problem

$$min_{\mathbf{x} \in \mathbb{R}^d} f(\mathbf{x}) \ \text{ subject to } \mathbf{a}^\top \mathbf{x} \geq b$$

It is given that $f \in \mathcal{C}_L^1$ with $L = 5$. Consider solving the problem with gradient projection

$$\mathbf{x}^{(k+1)} = P_C\left(\mathbf{x}^k - 0.2\nabla f(\mathbf{x}^{(k)})\right) \tag{GRADPROJ}$$

where C is the constraint set.

(a) (4 points) For any $\mathbf{z} \in \mathbb{R}^d$ find $P_C(\mathbf{z})$ with brief justification? Assume $\mathbf{a}^\top \mathbf{z} < b$.

Solution:

$$P_C(z) = argmin_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \mathbf{z}\|^2 \ \text{ subject to } \mathbf{a}^\top \mathbf{x} \geq b$$

From optimality $\mathbf{x} = \mathbf{z} + \lambda \mathbf{a}$, $\lambda(\mathbf{a}^\top \mathbf{x} - b) = 0, \lambda \geq 0$. Thus $P_C(\mathbf{z}) = \mathbf{z} + \left(\frac{b - \mathbf{a}^\top \mathbf{z}}{\|\mathbf{a}\|^2}\right)\mathbf{a}$

(b) Assuming $\mathbf{a}^\top \mathbf{x}^{(k)} = b$, $\mathbf{a}^\top \nabla f(\mathbf{x}^{(k)} > 0$ answer the following

   i. (3 points) find $\mathbf{x}^{(k+1)}$ in (GRADPROJ).

Solution: In the previous question set $\mathbf{z} = \mathbf{x}^{(k)} - 0.2\nabla f(\mathbf{x}^{(k)})$ Using the condition in the question, $\mathbf{z} \notin C$ and hence

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - 0.2\nabla f(\mathbf{x}^{(k)}) + 0.2(\mathbf{a}^\top \nabla f(\mathbf{x}^{(k)})\frac{\mathbf{a}}{\|\mathbf{a}\|^2}$$

   ii. (3 points) Answer True or False.

$\nabla f(\mathbf{x}^{(k)\top}(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) > 0$ holds.__**F**__ Justify.

Solution:

$$(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)})^\top \nabla f(\mathbf{x}^{(k)}) = -0.2\left(\|\nabla f(\mathbf{x}^{(k)})\|^2 - \frac{(\nabla f(\mathbf{x}^{(k)\top}\mathbf{a})}{\|\mathbf{a}\|^2}\right)$$

By Cauchy Schwartz inequality the expression is negative.

(c) (5 points) In (`GRADPROJ`) is known that the scheme will converge to a point $\hat{\mathbf{x}}$. What can you say about the optimality of $\hat{\mathbf{x}}$.

Solution: At convergence

$$\hat{\mathbf{x}} = P_C(\hat{\mathbf{x}} - 0.2\nabla f(\hat{\mathbf{x}}))$$

From the property of Projections, for any $\mathbf{x} \in C$,

$$(\hat{\mathbf{x}} - \hat{\mathbf{x}} + 0.2\nabla f(\hat{\mathbf{x}}))^\top (\mathbf{x} - \hat{\mathbf{x}}) \geq 0$$

Thus $\nabla f(\hat{\mathbf{x}}))^\top (\mathbf{x} - \hat{\mathbf{x}}) \geq 0$ and hence $\hat{\mathbf{x}}$ can be considered as satisfying the necessary condition for local minimum.

Rough Work

3. Consier the problem

$$min_{\mathbf{x}\in\mathbb{R}^3} f(\mathbf{x}) \qquad f(\mathbf{x}) = x_1^2 - 2x_2^2 + 4x_3^2 + 2\sum_{i=1}^{3} r_i x_i$$

Subject to $\sum_{i=1}^{3} x_i^2 = 1$.
The following conditions are given

$$\text{Condition C1: } \frac{r_1^2}{6} + \frac{r_2^2}{3} + \frac{r_3^2}{9} = 5, \qquad \text{Condition C2 } \left(\frac{r_1}{6}\right)^2 + \left(\frac{r_2}{3})\right)^2 + \left(\frac{r_3}{9}\right)^2 = 1$$

(a) (1 point) True or False. f is convex **F**.

(b) (2 points) Find the Dual function, $g(\mu)$, of the problem where μ is the dual variable. State the Dual feasible set.

Solution:

$$L(\mathbf{x}, \mu) = x_1^2(1+\mu) + x_2^2(\mu - 2) + x_3^2(4+\mu) + 2\sum_{i=1}^{3} r_i x_i - \mu$$

$$g(\mu) = min_{\mathbf{x}\in\mathbb{R}^3} f(\mathbf{x}) \qquad g(\mu) = -\mu - \left(\frac{r_1^2}{(1+\mu)} + \frac{r_2^2}{(\mu-2)} + \frac{r_3^2}{(4+\mu)}\right)$$

This is defined for all $\mu \geq 2$. The minimization is attained at $x_1 = \frac{r_1}{1+\mu}, x_2 = \frac{r_2}{\mu-2}, x_3 = \frac{r_3}{4+\mu}$.

(c) (3 points) Evaluate the value of Dual objective at optimality?

i. 1 State the Dual problem? Both the objective and constraints if any need to be mentioned.

Solution: The Dual problem is

$$max_{\mu \geq 2} - \mu - \left(\frac{r_1^2}{(1+\mu)} + \frac{r_2^2}{(\mu-2)} + \frac{r_3^2}{(4+\mu)}\right)$$

Rough Work

ii. (4 points) The Dual objective function value is ___**-10**___ and it is attained at $\mu =$ ___**5**___. Justify

Solution: Since $-g$ is convex in the domain, so a sufficient condition for optimality is given by $\frac{d}{d\mu}g(\mu) = 0$. Thus

$$1 = \frac{r_1^2}{(1+\mu)^2} + \frac{r_2^2}{(\mu-2)^2} + \frac{r_3^2}{(4+\mu)^2}$$

From C2 it appears that $\mu = 5$ satisfies the sufficient condition and hence $g(\mu) \leq g(5)$. From C1 one can derive that $g(5) = -5 - 5 = -10$

(d) (5 points) The global optimal value of the primal problem is ___$= -10$___. Justify

Solution: For any $\mu \geq 0$, the minimization is attained at $x_1 = -\frac{r_1}{1+\mu}, x_2 = -\frac{r_2}{\mu-2}, x_3 = -\frac{r_3}{4+\mu}$. For $\mu = 5$ it satisfies $\|\mathbf{x}\|^2 = 1$ and hence it is feasible solution to the original problem. The function value at that point can be obtained as

$$\begin{aligned} f(\mathbf{x}) &= -\left(\frac{r_1^2}{\mu+1}(1+\frac{\mu}{\mu+1}) + \frac{r_2^2}{\mu-2}(1+\frac{\mu}{\mu-2}) + \frac{r_3^2}{\mu+4}(1+\frac{\mu}{\mu+4})\right) \\ &= -\left(\frac{r_1^2}{\mu+1} + \frac{r_2^2}{\mu-2} + \frac{r_3^2}{\mu+4}\right) - \mu \\ &= -5 - 5 = -10 \end{aligned}$$

Thus strong duality holds and -10.

Rough Work

4. Consider the problem

$$\min_{\mathbf{x}\in\mathbb{R}^d} \frac{1}{2}\|\mathbf{x}|^2 \quad \text{subject to } \mathbf{a}^\top\mathbf{x} \leq -b, \quad \mathbf{c}^\top\mathbf{x} \geq b$$

where $b > 0, \mathbf{a}, \mathbf{c} \in \mathbb{R}^d$

(a) (1 point) State the Wolfe Dual of the problem? The statement should not have primal variables.

Solution: The Dual problem is $max_{\lambda_1\geq 0,\lambda_2\geq 0} g(\lambda_1, \lambda_2)$

$$g(\lambda_1, \lambda_2) = -\frac{1}{2}\|\lambda_1\mathbf{c} - \lambda_1^2\mathbf{a}\|^2 + b(\lambda_1 + \lambda_2)$$

(b) (1 point) Assuming the following $\mathbf{a}^\top\mathbf{c} = 0, \|\mathbf{a}\| = |\mathbf{c}\| = r$, solve the Wolfe Dual and compute both the optimal values of the dual variables and the dual objective function value

Solution: $g(\lambda_1, \lambda_2) = -\frac{r^2}{2}\left(\lambda_1^2 + \lambda_2^2\right) + b(\lambda_1 + \lambda_2)$ The optimal values are thus $\lambda_1 = \lambda_2 = \frac{b}{r^2}$. Dual objective function value is $\frac{b^2}{r^2}$.

(c) (3 points) Assuming the conditons of the previous question solve the primal problem? You need to state the objective function value and the optimal $\mathbf{x}$.

Solution: Wolfe Dual is obtained at $\mathbf{x}^* = \frac{b}{r}(\mathbf{c} - \mathbf{a})$ and the objective function value of the primal is $\frac{b^2}{r^2}$. We can check that the Wolfe Dual problem is same as that obtained through Lagrange Dual. Hence strong duality holds.

Rough Work

Rough Work

Rough Work

Rough Work

Rough Work