Computational Methods of Optimization
Second Midterm(30th 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 $\mathbf{I} = [e_1, \ldots, e_d]$ 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 $\{1, 2, \ldots, n\}$

1. (5 points) Please indicate True(T) or False(F) in the space given after each question. All questions carry equal marks
   (a) The set $\{\mathbf{x} \in \mathbb{R}^d | f(\mathbf{x}) \leq b\}$ is convex where $f : C \subset \mathbb{R}^d \to \mathbb{R}$ is a convex function and $b \in \mathbb{R}$ **T**
   (b) The set $\{\mathbf{x} \in \mathbb{R}^d | \|\mathbf{x}\| = 1\}$ is not convex **T**
   (c) The set $\{\mathbf{x} \in \mathbb{R}^d | 1 \leq \|\mathbf{x}\| \leq 2\}$ is not convex **T**
   (d) The projection of $\mathbf{z} \in \mathbb{R}^d$ on a non-convex set C does not exist **F**
   (e) Let $\mathbf{x}^*$ be the global minimum of

   $$min_{\mathbf{x} \in C} f(\mathbf{x}) \left(= \|\mathbf{x} - \mathbf{a}\|^2\right)$$

   and $\mathbf{z}^*$ be the minimum of $\sqrt{f(\mathbf{x})}$. The two minima are different **F**.

2. (4 points) Pick the correct choice. All questions carry equal marks
   (a) Consider the following problem

   $$\mathbf{x}^* = argmin_{\mathbf{x} \in C} f(\mathbf{x})$$

   where $f \in \mathcal{C}^1$. Let $\mathbf{z}^*$ be the unconstrained minimum of $f(\mathbf{x})$. When is $\mathbf{z}^* = \mathbf{x}^*$?
   A. There is no relationship B. $\mathbf{x}^*$ is not an interior point of C **C. $\mathbf{x}^*$ is an interior point of C**
   (b) Let the columns of $d \times d$ matrix $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_d]$ be Q conjugate for a $d \times d$ matrix Q. The off-diagonal entries of matrix $B = \mathbf{U}^\top Q \mathbf{U}$ are A. $\mathbf{u}_i^\top \mathbf{u}_j$ **B. 0** C. Numerical value cannot be determined

3. (a) (3 points) Let $\mathbf{e}_1, \ldots, \mathbf{e}_d$ be the columns of $\mathbf{I}_{d \times d}$ matrix and $Q \in \mathbb{R}^{d \times d}$ be a positive semi-definite matrix. Find A_{ij} such that $\mathbf{u}_i$ are Q conjugate where $\mathbf{u}_1 = \mathbf{e}_1$, $\mathbf{u}_i = \mathbf{e}_i + \sum_{j=1}^{i-1} A_{ij} \mathbf{u}_j$ for $i \geq 2$.

   Solution: Since $\mathbf{u}_i$ are Qconjugate $\mathbf{u}_l^\top Q \mathbf{u}_i = 0$ holds for all $l < i$. As a consequence $\mathbf{u}_l^\top Q \mathbf{u}_i = \mathbf{e}_i^\top Q \mathbf{u}_l + A_{il} \mathbf{u}_l^\top Q \mathbf{u}_l = 0$. Hence $A_{il} = -\frac{(Q\mathbf{u}_l)_i}{\mathbf{u}_l^\top Q \mathbf{u}_l}$ will ensure that $\mathbf{u}_i$ are $Q-$conjugate. Here $(\mathbf{x})_i$ denote the ith coordinate of the vector $\mathbf{x}$.

   (b) Let $B\mathbf{x} = b$ be a linear system of equations where $B \in \mathbb{R}^{d \times d}$, a symmetric matrix which is positive definite and $b \in \mathbb{R}^d$. Using $\mathbf{u}_i$ defined in the previous question we wish to solve the linear system of equations using Conjugate direction algorithm
      i. (3 points) State the objective function to be used and argue why it will lead to solving the linear system of equations.

      Solution:

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

      The optimal solution is $\nabla f(\mathbf{x}) = 0$, which implies that $B\mathbf{x} = b$

      ii. (4 points) Starting at $\mathbf{x}^{(0)} = 0$ find $\mathbf{x}^{(1)}$ using the direction $\mathbf{u}_1$

Solution:

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha \mathbf{u}_1,\ \alpha = -\frac{\nabla f(\mathbf{x}^{(0)})^\top \mathbf{u}_1}{\mathbf{u}_1^\top B \mathbf{u}_1}$$

Since $\nabla f(\mathbf{x}^{(0)}) = -b$ and $\mathbf{u}_1 = \mathbf{e}_1$ which implies $\mathbf{x}^{(1)} = -\frac{b_1}{B_{11}}\mathbf{e}_1$

4. Bunty and Babli were arguing over the following problem

$$min_{x_1 \in \mathbb{R}, x_2 \in \mathbb{R}} f(x_1, x_2) = (x_1 - 2)^2 + x_2^2 \text{ subject to } (x_1 - 2)^2 = (x_2 - 3)^5 \ \ (\mathcal{P})$$

Bunty substitutes $(x_1 - 2)^2$ in the objective by $(x_2 - 3)^5$ and transforms $(\mathcal{P})$ into the following unconstrained problem

$$min_{x_2} (x_2 - 3)^5 + x_2^2 \ \ (\mathcal{Q})$$

The objective function of $(\mathcal{Q})$ is not bounded from below and global minimum does not exist. Hence Bunty concludes that global minimum of $(\mathcal{P})$ does not exist. Babli disagrees with Bunty and says that $(\mathcal{Q})$ is not equivalent to $(\mathcal{P})$.

(a) (1 point) Who is correct, Bunty or Babli? Give reasons.

Solution: Babli is correct. The objective function of $(\mathcal{P})$ is bounded from below, it cannot be less than zero.

(b) (3 points) Babli further says that $(\mathcal{P})$ can be solved by solving a equivalent convex optimization problem. What should Bunty do to make $(\mathcal{Q})$,a convex optimization problem? State the optimization problem.

Solution: The constraint $(x_1 - 2)^2 = (x_2 - 3)^5$ is not feasible for $x_2 < 3$. Thus $\mathcal{Q}$ is not equivalent to $\mathcal{P}$ as it allows for all values of x_2. Bunty needs to add the constraint $x_2 \geq 3$. The new problem is thus

$$min_{x_2} (x_2 - 3)^5 + x_2^2 \ \ x_2 \geq 3$$

The Hessian for the objective function in $\mathcal{Q}$ is

$$\frac{5}{2}\frac{3}{2}(x_2 - 3) + 2$$

which is positive over $x_2 \geq 3$.

Thus the objective function is convex over a convex constraint set, and hence the problem is convex.

(c) i. (5 points) Find the global minimum point of the convex optimization problem. Jusitify your answer using KKT conditions.

Solution: Bunty now solves

$$min_z g(z) \left(= (z-3)^5 + z^2\right) \text{ subject to } z \geq 3$$

‘ The objective function is convex as $\frac{d}{dz}g(z) > 0$ for all $z \geq 3$. The feasible set is convex. For convex problems KKT conditions are sufficient for finding global minimum.
The Lagrangian of the problem is

$$L(z, \mu) = (z-3)^5 + z^2 - \mu(z-3)$$

The KKT conditions are

$$5(z-3)^4 + 2z - \mu = 0, \mu(z-3) = 0$$

Clearly $\mu = 6, z = 3$ is a KKT point. Hence the global optimum of $g(z)$ is obtained at $z = 3$ and $g(z) = 9$.

ii. (2 points) Find the global minimum point and optimal value of $(\mathcal{P})$. Justify your answer.

Solution: The feasible set can be described as

$$\{(x_1, x_2)^\top | x_1 = 2 \pm (x_2 - 3)^{\frac{5}{2}}, x_2 \geq 3\}$$

. Thus $f(x_1, x_2) \geq f(2, 3) = 9$ whenever $x_2 \geq 3$. The point $[2, 3]^\top$ is the global minimum and $f(x_1, x_2) \geq 9$.

5. We are interested in finding the projection of $\mathbf{z} \in \mathbb{R}^d$ on the set $C = \{\mathbf{x} \in \mathbb{R}^d | 0 \leq x_i \leq t\}$ where $t > 0$.

(a) (3 points) State the Lagrangain of the projection problem as $\sum_{i=1}^d g(x_i, \lambda_{1i}, \lambda_{2i})$

Solution:

$$L(\mathbf{x}, \lambda_1, \lambda_2) = \frac{1}{2}\|\mathbf{x} - \mathbf{z}\|^2 - \sum_{i=1}^d \lambda_{1i} x_i + \sum_{i=1}^d \lambda_{2i}(x_i - t)$$

$$L(\mathbf{x}, \lambda_1, \lambda_2) = \sum_{i=1}^d g(x_i, \lambda_{1i}, \lambda_{2i}) \left(= \frac{1}{2}(x_i - z_i)^2 - \lambda_{1i} x_i + \lambda_{2i}(x_i - t)\right)$$

(b) (6 points) Find a KKT point for the problem.

Solution:

$$(\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda_1, \lambda_2))_i = x_i - z_i - \lambda_{1i} + \lambda_{2i} = 0$$

The KKT point $(\mathbf{x}, \lambda_1, \lambda_2)$ is stated as follows

$$\mathbf{x}_i = \begin{cases} z_i & 0 \leq z_i \leq t & \lambda_{1i} = \lambda_{2i} = 0 \\ 0 & z_i < 0 & \lambda_{1i} = -z_i, \lambda_{2i} = 0 \\ t & z_i > 0 & \lambda_{1i} = 0, \lambda_{2i} = z_i - t \end{cases}$$

(c) (1 point) Find the projection.

> **Solution:** The problem is convex and hence the KKT point is global optimal. Thus $\mathbf{x}$ obtained in above problem is the projection

6. Let $B \in \mathbb{R}^{d\times d}$ is a symmetric positive definite matrix. We wish to solve

$$\mathcal{P} \quad min_{\mathbf{x}\in\mathbb{R}^d} \mathbf{x}^\top B\mathbf{x} \text{ subject to } \mathbf{a}^\top \mathbf{x} = 1$$

(a) (4 points) Show that $\mathcal{P}$ is solved if there exists $\mu \in \mathbb{R}$ so that $(\mathbf{z}^\top, \mu)^\top$ solves

$$\begin{bmatrix} 2B & \mathbf{a} \\ \mathbf{a}^\top & 0 \end{bmatrix} \begin{pmatrix} \mathbf{z} \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

> **Solution:** The Lagrangian is $L(\mathbf{x}, \mu) = \mathbf{x}^\top B\mathbf{x} + \mu(\mathbf{a}^\top \mathbf{x} - 1)$. KKT conditions are
>
> $$2B\mathbf{x} + \mu\mathbf{a} = 0, \;\; \mathbf{a}^\top \mathbf{x} = 1$$
>
> which leads to the desired system of equations. Since KKT conditions are sufficient for solving convex optimization problem the solution to the desired set of equations will also solve $\mathcal{P}$.

(b) (6 points) Solve $\mathcal{P}$. State the optimal objective function and the optimum point

> **Solution:** The system of equations can be solved as
>
> $$\mathbf{z} = -\frac{\mu}{2}B^{-1}\mathbf{a}, -\frac{\mu}{2} = \frac{\mathbf{a}^\top \mathbf{z}}{\mathbf{a}^\top B^{-1}\mathbf{a}} = \frac{1}{\mathbf{a}^\top B^{-1}\mathbf{a}}$$
>
> $$\mathbf{z} = \frac{1}{\mathbf{a}^\top B^{-1}\mathbf{a}}B^{-1}\mathbf{a}, \quad \mathbf{z}^\top B\mathbf{z} = \frac{1}{\mathbf{a}^\top B^{-1}\mathbf{a}}$$
>
> are the optimal point and optimum value respectively