Computational Methods of Optimization
Third Midterm(9th Nov, 2021)

Instructions:

- This is a closed book test. Please do not consult any additional material.
- Attempt all questions
- Total time is 70 mins.
- Answer the questions in the spaces provided. Answers outside the spaces provided will not be graded.
- Rough work can be done in the spaces provided at the end of the booklet

Name: ______________________________

SRNO: Degree: Dept:

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

1. Our goal is to find the minimum norm solution of the following equations

$$2x_1 + x_2 + x_3 = 36 \quad \texttt{Eq1}, \quad -x_1 - x_2 + 3x_3 = 0, \quad \texttt{Eq2}$$

   (a) (1 point) Pose the problem as a quadratic optimization problem over the constraint set $A\mathbf{x} = b$

Solution:

$$min_{\mathbf{x}\in\mathbb{R}^3}\frac{1}{2}\|\mathbf{x}\|^2$$
$$\text{subject to} \quad A\mathbf{x}=b$$

where $A=\begin{bmatrix} 2 & 1 & 1 \\ -1 & -1 & 3 \end{bmatrix}$, and $b=[36,0]^\top$

(b) (2 points) State the KKT conditions of the problem in terms of the lagrange multipliers, A and b.

Solution:

$$\mathbf{x}=A^\top\mu, A\mathbf{x}=b$$

(c) (1 point) At optimality compute the Lagrange multipliers, μ_1^* corresponding to `Eq1` amd μ_2^* corresponding to `Eq2`.

Solution: $\mu_1^*=6, \mu_2^*=0$

(d) (2 points) Give reasons to justify your answer.

Solution:

$$\mu=\left(AA^\top\right)^{-1}b$$

. Direct compution leads to the answer.

(e) (3 points) Find the optimal solution, $\mathbf{x}^*$, of the optimization problem

Solution:

$$\mathbf{x}^*=A^\top\mu^*, \quad ,\mathbf{x}^*=[12,6,6]^\top$$

(f) (1 point) Find another solution of the system of equations which has the maximum number of zeros in the solution.

Solution: There is no solution with two zeros in the individual coordinates. Fix one coordinate to zero, and solve for the rest

- $x_1=0$ leads to $x_2=27, x_3=9$
- $x_2=0$ leads to $x_1=108/7, x_3=36/7$
- $x_3=0$ leads to $x_1=36, x_2=-36$

2. Let $a_1<a_2\ldots<a_{2m+1}$ be $2m+1$ real numbers sorted from smallest to largest, a_1 being the smallest and a_{2m+1} being the largest. Consider the problem

$$min_{x\in\mathbb{R},t\in\mathbb{R}^{2m+1}} \quad \sum_{i=1}^{m} t_i$$
$$\text{subject to} \qquad -t_i\le x-a_i\le t_i, \quad t_i\ge 0, i=1,\ldots,2m+1$$

where $t=[t_1,\ldots,t_{2m+1}]^\top$. Let the Lagrangian of the problem be

$$L(x,t,\lambda_1,\lambda_2,\lambda_3)=\sum_{i=1}^{2m+1}t_i-\sum_{i=1}^{2m+1}\lambda_{1i}(x-a_i+t_i)+\sum_{i=1}^{2m+1}\lambda_{2i}(x-a_i-t_i)-\sum_{i=1}^{2m+1}\lambda_{3i}t_i$$

(a) Suppose $x = a_l$ is the optimal solution.

i. (1 point) The value of $\lambda_{1i}, \lambda_{2i}$ when $i < l$ is

A. $\lambda_{1i} = \lambda_{2i} = 0$

B. $\lambda_{1i} = 1, \lambda_{2i} = 0$

C. $\lambda_{1i} = 0, \lambda_{2i} = 1$

D. $\lambda_{1i}, \lambda_{2i}$ are both positive

ii. (2 points) Justify your answer using KKT conditions

Solution: From KKT conditions we obtain

- for feasibility $t_i \geq a_l - a_i$.
- whenever $t_i > 0$, $\lambda_{1i} + \lambda_{2i} = 1$
- From complimentarity conditions it follows that $\lambda_{1i} = 0$ and hence from the above $\lambda_{2i} = 1$

iii. (1 point) The value of $\lambda_{1i}, \lambda_{2i}$ when $i > l$ is

A. $\lambda_{1i} = \lambda_{2i} = 0$

B. $\lambda_{1i} = 1, \lambda_{2i} = 0$

C. $\lambda_{1i} = 0, \lambda_{2i} = 1$

D. $\lambda_{1i}, \lambda_{2i}$ are both positive

iv. (2 points) Justify your answer using KKT conditions

Solution: Similar reasoning.

(b) i. (1 point) For what value of l, $x = a_l$ is optimal?

A. $l = 1$

B. $l = 2m + 1$

C. $l = m$

D. Cannot be determined

ii. (3 points) Jsutify your answer using KKT conditions

Solution: One of the KKT conditions require that

$$\sum_{i=1}^{2m+1} \lambda_{i1} = \sum_{i=1}^{2m+1} \lambda_{2i}$$

From before we know that $i < l$, $\lambda_{1i} = 0, \lambda_{2i} = 1$ and for $i > l$, $\lambda_{1i} = 1, \lambda_{2i} = 0$. Furthermore we can set $\lambda_{1m} = \lambda_{2m} = 0$. Thus for choice $l = m$,

$$\sum_{i=1}^{2m+1} \lambda_{i1} = \sum_{i=1}^{2m+1} \lambda_{2i} = m$$

and hence the $x = a_m$ is the global optima

3. (a) (5 points) Compute the projection, $P_C(\mathbf{z})$, of $\mathbf{z} \in \mathbb{R}^d$ on $C = [0, 1]^d$. You need to state the KKT conditions, and show that your answer satisifies these conditions

Solution: The Lagrangian is $L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \frac{1}{2}\|\mathbf{x} - \mathbf{z}\|^2 + \boldsymbol{\lambda}^T(\mathbf{x} - \mathbf{1}) - \boldsymbol{\mu}^T\mathbf{x}$ where $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^d_{\geq 0}$. Then the KKT conditions are as follows,

- **Primal feasibility:** $0 \leq x_i \leq 1$ for all $i \in [d]$.
- **Complementary slackness:** $\lambda_i(x_i - 1) = 0$ and $\mu_i x_i = 0$ for all $i \in [d]$.
- **Dual feasibility:** $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^d_{\geq 0}$.
- **Stationarity:** $\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = 0 \implies x_i - z_i + \lambda_i - \mu_i = 0, \forall i \in [d]$.

If could you guess the projection as the following,

$$P_C(\mathbf{z})_i = \begin{cases} 1 & z_i > 1 \\ z_i & 0 \leq z_i \leq 1 \\ 0 & z_i < 0 \end{cases}, \tag{1}$$

you need to show that in each of the three cases there exist some values of λ and μ such that all the KKT conditions are satisfied.

Case 1: $z_i > 1$. Then setting $x_i = 1$, $\lambda_i = z_i - 1$ and $\mu_i = 0$ satisfies all the KKT conditions.
Case 2: $0 \leq z_i \leq 1$. Then setting $x_i = z_i$, $\lambda_i = 0$ and $\mu_i = 0$ satisfies all the KKT conditions.
Case 3: $z_i < 0$. Then setting $x_i = 0$, $\lambda_i = 0$ and $\mu_i = -z_i$ satisfies all the KKT conditions.

Alternatively, we can also compute the projection using the KKT conditions. There are four possible cases for each set of λ_i, μ_i as follows,

Case 1: $\lambda_i = 0, \mu_i = 0$. Dual feasibility and complementary slackness is satisfied for any value of z_i. To satisfy stationarity we set $x_i = z_i$. Now we get that primal feasiblity can be satisfied only when $0 \leq z_i \leq 1$.

Note. If primal feasiblity constraints are not mentioned, for any value of $\mathbf{z}$ the rest of the constraints are satisfied and the projection becomes $P_C(\mathbf{z}) = \mathbf{z}$ for any $\mathbf{z}$.

Case 2: $\lambda_i > 0, \mu_i = 0$. Dual feasibility is satisfied. To satisfy complementary slackness we set $x_i = 1$. This also satisfies primal feasibility. Then the stationarity can be satisfied whenver $z_i - 1 = \lambda_i$, that is, whenever $z_i > 1$ there exists a value of $\lambda_i > 0$ that satisfies all the constraints. (Note that for $z_i \leq 1$ there is no value of $\lambda_i > 0$ that satisfies stationarity).

Case 3: $\lambda_i = 0, \mu_i > 0$. Dual feasibility is satisfied. To satisfy complementary slackness we set $x_i = 0$. This also satisfies primal feasibility. Then the stationarity can be satisfied whenver $-z_i = \mu_i$, that is, whenever $z_i < 0$ there exists a value of $\mu_i > 0$ that satisfies all the constraints.

Case 4: $\lambda_i > 0, \mu_i > 0$. We can not satisfy complementary slackness as we need $x_i = 0$ and $x_i = 1$ simultaneously. Therefore this case does not correspond to any feasible solution.

These four cases give the projection in (1).

(b) (5 points) Suppose it is given that $\mathbf{z} \in \mathbb{R}^3$, $\mathbf{z} \notin [0,1]^3$ with $P_C(\mathbf{z}) = [0, 0.5, 0.8]^\top$. It is given that the largest absolute value of the coordinates of $\mathbf{z}$ is 1. From this information can you reconstruct $\mathbf{z}$. Justify your answer.

Solution: Let $P_C(\mathbf{z}) = \mathbf{x}^*$. In this case a key property of projection is $\mathbf{x}_i^* = \mathbf{z}_i$ whenever $0 \le \mathbf{z}_i \le 1$. Thus this determines $\mathbf{z}_2 = 0.5, \mathbf{z}_3 = 0.8$. Since $\mathbf{x}_1^* = 0$, it implies that $\mathbf{z}_1 \le 0$. Note that the largest absolute value of coordinates of $\mathbf{z}$ is 1, this neccessarily means that $\mathbf{z}_1 = -1$. Hence the point is $\mathbf{z} = [-1, 0.5, 0.8]^\top$.

4. Consider the following problem

$$min_{\mathbf{x}} f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top Q \mathbf{x} - b^\top \mathbf{x}, \quad \text{subject to } \|\mathbf{x}\|^2 \le 1$$

The eigenvalues of Q lie between -1 and 1.

(a) (5 points) Write one step of the gradient projection algorithm starting at $\mathbf{x}^{(k)}$ for this problem? Assume that the stepsize is fixed to α.

Solution:

$$\begin{aligned}
\mathbf{x}^{(k+1} &= P_C\left[\mathbf{x}^k - \alpha \nabla f(\mathbf{x}^{(k)})\right] \\
P_C(\mathbf{z}) &= \frac{\mathbf{z}}{\|\mathbf{z}\|} \text{ whenever } \|\mathbf{z}\| > 1, z \text{ otherwise.} \\
P_C(\mathbf{z}) &= \frac{\mathbf{z}}{\max(1, \|\mathbf{z}\|)} \\
\mathbf{x}^{(k+1)} &= \frac{[I - \alpha Q]\,\mathbf{x}^{(k)} + \alpha b}{\max\left(1, \|\,[I - \alpha Q]\,\mathbf{x}^{(k)} + \alpha b\|\right)}
\end{aligned}$$

(b) (5 points) What is the range of α? Justify your answer

Solution: Direct computation shows that

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

$$\leq \frac{1}{2}\lambda\|\mathbf{y} - \mathbf{x}\|^2 = \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|^2$$

where $\lambda = 1$ is the largest eigenvalue of Q. Let $\mathbf{y} = P_C(\mathbf{x} - \alpha\nabla f(\mathbf{x}))$ Thus

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

Thus

$$f(\mathbf{y}) \leq f(\mathbf{x}) - \left(\frac{1}{\alpha} - \frac{1}{2}\right)\|\mathbf{y} - \mathbf{x}\|^2$$

Thus $0 \leq \alpha \leq 2$