Optimization HW1

Adi Amuzig
205928211

Grimovitch Ekaterina
314092826

27 November 2019

Question 1

Prove or disprove the following claims.
Note that for this questions we assume $x, y \in \mathbb{R}^n$.

1. The expression $\langle x, y\rangle_M = x^T My$ is an inner product if $M \in \mathbb{R}^{n\times n}$.

   By definition, an inner product maps $\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and for every $x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$ upholds:

   (a) Symmetry: $\langle x, y\rangle = \langle y, x\rangle$
   (b) Additivity: $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
   (c) Homogeneity: $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$
   (d) Positive-definite: $\langle x, x\rangle \geqslant 0$ and $\langle x, x\rangle = 0$ iff $x = 0$.

   We can easily show that the given product does not answer (d), meaning it is not positive definite. If M is the zero matrix (it contains only zeros) then: $\langle x, x\rangle_M = x^T Mx = 0$ even though $x \neq 0$. We got/ a contradiction, therefore the given expression is not an inner product.

2. The expression $\langle x, y\rangle_Q = x^T Qy$ is an inner product if $0 \prec Q \in \mathbb{R}^{n\times n}$.

   The definition of a positive definite matrix is that for a symmetrical matrix $Q \in \mathbb{R}^{n\times n}$ and for every $0 \neq x \in \mathbb{R}^n$ we get $x^T Qx > 0$. Marking $0 \prec Q$. Lets go over all 4 requirements for an inner product:

   (a) Symmetry: Q is symmetrical, therefore $Q = Q^T$. $\langle x, y\rangle = x^T Qy = (y^T Qx)^T$ The product of this equation is a single value (not a matrix) therefore it is equal to its transpose.
   $(y^T Qx)^T = y^T Qx = \langle y, x\rangle$
   (b) Additivity: $\langle x + y, z\rangle = (x + y)^T Qz = (x^T + y^T)Qz = x^T Qz + y^T Qz = \langle x, z\rangle + \langle y, z\rangle$
   (c) Homogeneity: $\langle \lambda x, y\rangle = (\lambda x)^T Qy = \lambda x^T Qy = \lambda\langle x, y\rangle$
   (d) Positive-definite: According to the positive definite matrix definition, the positive-definite requirement is met.

   The expression answers all requirements, therefore it is an inner product.

Question 2

Let $A \in \mathbb{R}^{n\times n}$

1. Prove that if $0 \prec A$ then it is invertible, and $0 \prec A^{-1}$

   By definition, a positive definite matrix is a symmetrical matrix $A \in \mathbb{R}^{n\times n}$ that for every $0 \neq x \in \mathbb{R}^n$ we get $x^T Ax > 0$. Marking it $0 \prec A$.

   According to tutorial 1, $x \prec A$ iff all of its eigenvalues are positive definite. It is known that 0 is not an eigenvalue of a square matrix iff it is invertible.

   For eigenvector $0 \neq v \in \mathbb{R}^n$ and its corresponding eigenvalue λ

   $$\begin{aligned} Av &= \lambda v \\ A^{-1}Av &= A^{-1}\lambda v \\ A^{-1}v &= \frac{1}{\lambda}v \end{aligned}$$

   Since $v \neq 0$ it implies that it is the eigenvector for A^{-1} with the corresponding eigenvalue $\frac{1}{\lambda}$. Given that, $\lambda > 0 \implies \frac{1}{\lambda} > 0 \implies 0 \prec A^{-1}$.

2. Prove that if $0 \preceq A$ and it is inevitable then $0 \prec A$

   As seen in tutorial 1, $x \preceq A$ iff all of its eigenvalues are positive semidefinite. Because it is also invertible that means that $\lambda \neq 0 \implies \lambda > 0$. This means that $x \prec A$ because all of its eigenvalues are positive definite.

3. $A = B^T B$ for some $B \in \mathbb{R}^{m\times n}$. We will prove that $0 \preceq A$.

   $$\begin{aligned} x^T Ax = x^T B^T Bx &= (Bx)^T(Bx) \\ Bx &\in \mathbb{R}^m \\ Bx &= y \\ (Bx)^T(Bx) &= y^T y = \sum_{i=1}^{m} y_i^2 \geq 0 \end{aligned}$$

   $0 \preceq A$ by definition.

   Note that $0 \nprec A$. For example

   $$\begin{aligned} x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}&, \ B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ A = B^T B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}&\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ x^T Ax = \begin{bmatrix} 0 & 1 \end{bmatrix}&\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \end{aligned}$$

4. Prove that if $A_{i,i} < 0$ for any $1 \leq i \leq n$ then $0 \npreceq A$.

   $$x^T Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_i A_{i,j} = \sum_{i=1}^{n}\sum_{j\neq i} x_i x_i A_{i,j} + \sum_{i=1}^{n} x_i^2 A_{i,i}$$

If there exists $A_{i,i} < 0$ then $x_i^2 A_{i,i} < 0$ for $x_i \neq 0$. By choosing a vector x containing only zeros except x_i that corresponds to $A_{i,i} < 0$ we get $x^T A x < 0$ meaning $0 \not\preceq A$.

It is not a sufficient condition. For example

$$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$x^T A x = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0$$

Question 3

Function $f : U \to \mathbb{R}$ is continuously differential while $U \subseteq \mathbb{R}^n$. Direction $0 \neq d$ is called the downhill direction of the function at point x if there exists $0 < T$ so that for every $0 < t < T$

$$f(x + td) < f(x)$$

1. Prove that if direction d uphold $f'(x, d) < 0$ for a certain x, then it is a downhill direction of f at x.

   By definition:

$$f'(x, d) = \lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t}$$

   We assume $0 < t < T$.

$$\begin{gathered} f'(x, d) < 0 \\ \lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t} < 0 \\ \frac{f(x + td) - f(x)}{t} < 0 \\ f(x + td) - f(x) < 0 \\ f(x + td) < f(x) \end{gathered}$$

   If direction d uphold $f'(x, d) < 0$ then $f(x + td) < f(x)$ which means it is a downhill direction.

2. Prove that if $\nabla f(x) \neq 0$ then the direction $d = -\nabla f(x)$ is a downhill direction of f at x.

   By definition:

$$\begin{gathered} \frac{\partial f}{\partial x_i}(x) = \lim_{t \to 0^+} \frac{f(x + te_i) - f(x)}{t} \\ \nabla f(x) = \left[\frac{\partial f}{\partial x_1}(x) \;\; \frac{\partial f}{\partial x_2}(x) \;\; ... \frac{\partial f}{\partial x_n}(x) \right]^T \end{gathered}$$

   As seen in tutorial 1, if f is continuously differential then

$$\begin{gathered} f'(x, d) = \nabla f(x)^T d, \;\; \forall d \in \mathbb{R}^n \\ d = -\nabla f(x) \to \nabla f(x) = -d \\ \implies f'(x, d) = \nabla f(x)^T d = -d^T \times d = -\sum d_i^2 < 0 \end{gathered}$$

   As seen in part 1, if $f'(x, d) < 0$ for a certain x then it is a downhill direction of f at x.

3. Function $f(x_1, x_2) = x_1^2 + x_2^2$ is continuously differential. Given point $x^{(0)} = [x_1^{(0)} \; x_2^{(0)}]^T = [1 \; 1]^T$:

$$\nabla f(x) = \left[\frac{\partial f}{\partial x_1}(x) \;\; \frac{\partial f}{\partial x_2}(x) \right]^T = \left[2x_1^{(0)} \;\; 2x_2^{(0)} \right]^T = [2 \;\; 2]^T$$

According to part 2, if $\nabla f(x) \neq 0$ then the direction $d = -\nabla f(x)$ is a downhill direction of f at x. In our example $d_1 = [-2 \quad -2]^T$. Because $x^{(0)}$ is not a maximum or a minimun of the function $f(x_1, x_2)$ the opposite from the downhill direction will not be a downhill direction as well. Meaning $d_2 = [2 \quad 2]^T$.

Question 4

$C \subseteq \mathbb{R}^m$ is a convex domain and $A \in \mathbb{R}^{m \times n}$.

$$B_1 = \{x \in \mathbb{R}^n | x = A^T z, z \in C\}$$
$$B_2 = \{x \in \mathbb{R}^n | Ax \in C\}$$

By definition, domain $C \in \mathbb{R}^n$ is convex if for all $x, y \in C$ and $\lambda \in [0,1] \to \lambda x + (1-\lambda)y \in C$.

$$\lambda \in [0,1]$$
$$x, y \in B_1$$
$$x = A^T z_x$$
$$y = A^T z_y$$
$$z_x, z_y \in C$$
$$q = \lambda x + (1-\lambda)y = \lambda A^T z_x + (1-\lambda)A^T z_y = A^T(\lambda z_x + (1-\lambda)z_y) = A^T z_q,\ z_q \in C$$

$\implies B_1$ is a convex domain by definition.

$$x, y \in B_2$$
$$Ax \in C$$
$$Ay \in C$$
$$Aq = A(\lambda x + (1-\lambda)y) = \lambda(Ax) + (1-\lambda)(Ay) \in C$$

$\implies B_2$ is a convex domain by definition.

Question 5

Definition 1 let there be a set $S \subseteq \mathbb{R}^n$

1. Point $a \in S$ is called an inner point of S if there exists $\varepsilon > 0$ for which $B(a, \varepsilon) \subseteq S$.
2. Point $x \in \mathbb{R}^n$ (not necessarily in S) are called condensation point if for all $\varepsilon > 0$ the open ball $B(x, \varepsilon)$ contains at least one point $y \in S$, $y \neq x$.

Definition 2 let there be a set $S \subseteq \mathbb{R}^n$

1. Set S is called open if every point in S is an inner point.
2. Set S is called closed if its complimentary set, S^c, is an open set.

We will prove the following claims:

1. Let $S \subseteq \mathbb{R}^n$ be some set and $x \in \mathbb{R}^n$ its condensation point. Therefore for all $\varepsilon > 0$ the open ball $B(x, \varepsilon)$ contains an infinite amount of points from S.

   Using proof by contradiction assuming for a certain $\varepsilon > 0$ the open ball $B(x, \varepsilon)$ contains a finite amount of points from S. Each one of the points from S in the ball have a different distance from point x. We will mark the closest point to x as a and the distance as ε_a where $\varepsilon_a \leq \varepsilon$. The open ball $B(x, \varepsilon_a/2)$ does not contain any points from S because the closest point to x which is a is outside of it. Therefore $B(x, \varepsilon_a/2)$ does not contains any points $y \in S$, $y \neq x$ meaning we got a contradiction with definition 1.2 of a condensation point.

2. Set S is closed iff it containers all its condensation points.

   Using proof by contradiction **given that set S is closed we assume it does not contain all of its condensation points**. Meaning there is at least one condensation point $x \notin S$. $x \in S^c$ where S^c is the complimentary set of S and, according to definition 2.2, an open set. According to definition 1.2 ,because x is a condensation point then for all $\varepsilon > 0$ the open ball $B(x, \varepsilon)$ contains at least one point $y \in S$, $y \neq x$ ($y \notin S^c$). Therefore it is not an inner point of S^c according to definition 1.1. Looking at definition 2.1 we get a contradiction where not all points in the open set S^c are inner points.

   Using proof by contradiction **given that S contains all of its condensation points we assume it is not a closed set**. According to definition 2.2 if S is not closed then S^c is not open. According to definition 2.1 if S^c is not open then it has at least one point $x \in S^c$ that is not an inner point. According to definition 1.1 for all $\varepsilon > 0$ the open ball $B(x, \varepsilon) \nsubseteq S^c$ meaning it has at least one point $y \in B(x, \varepsilon), y \in S$. If $y \in S$ then point x is a condensation point according to definition 1.2 and we got a contradiction because $x \notin S$.

3. If S is a closed set and the series $i = 1, 2, ...,\ x_i \in S$ converges to x (meaning $\lim_{i \to \infty} ||x_i - x|| = 0$) then $x \in S$.

   Using proof by contradiction assuming $x \notin S$. Given that S is closed it containers all its condensation points meaning x is not a condensation point. According to definition 1.2 there exists a $\varepsilon > 0$ for which the open ball $B(x, \varepsilon) \nsubseteq S$. Meaning $B(x, \varepsilon)$ does not contain any point of the series x_i and we get a contradiction from a conclusion from the definition of a limit that says that for all $\varepsilon > 0$ there exists an i_0 such that for all $i > i_0$ we get $x_i \in B(x, \varepsilon)$.

Question 6

$C \subseteq \mathbb{R}$ is a convex and closed domain. $x_0 \notin C$. Given the optimization problem

$$\min_{y \in C} ||y - x_0||$$

Assume there is a solution $y_0 \in C$ for the optimization problem. We will prove that it is singular. Let us assume there are 2 solutions $y_0, z_0 \in C$ such that

$$||y_0 - x_0||^2 = ||z_0 - x_0||^2 = (\min_{y \in C} ||y - x_0||)^2 = \delta^2$$

By definition of a convex domain

$$\begin{gathered} y_0\lambda + z_0(1 - \lambda) \in C \\ \lambda = 0.5 \\ \implies m_0 = (y_0 + z_0)/2 \in C \end{gathered}$$

$$\begin{gathered} ||m_0 - x_0||^2 = ||(y_0 + z_0)/2 - x_0||^2 = ||(y_0 - x_0 + z_0 - x_0)/2||^2 = \\ \langle (y_0 - x_0 + z_0 - x_0)/2, (y_0 - x_0 + z_0 - x_0)/2 \rangle = \\ \frac{1}{4}\langle y_0 - x_0, y_0 - x_0 \rangle + \frac{1}{2}\langle z_0 - x_0, y_0 - x_0 \rangle + \frac{1}{4}\langle z_0 - x_0, z_0 - x_0 \rangle = \\ \frac{1}{4}||y_0 - x_0||^2 + \frac{1}{2}\langle z_0 - x_0, y_0 - x_0 \rangle + \frac{1}{4}||z_0 - x_0||^2 = \\ \frac{1}{4}\delta^2 + \frac{1}{2}\langle z_0 - x_0, y_0 - x_0 \rangle + \frac{1}{4}\delta^2 = \\ \frac{1}{2}\delta^2 + \frac{1}{2}\langle z_0 - x_0, y_0 - x_0 \rangle \end{gathered}$$

$$\begin{gathered} ||m_0 - x_0||^2 \geq (\min_{y \in C} ||y - x_0||)^2 = \delta^2 \\ \implies \frac{1}{2}\delta^2 + \frac{1}{2}\langle z_0 - x_0, y_0 - x_0 \rangle \geq \delta^2 \\ \langle z_0 - x_0, y_0 - x_0 \rangle \geq \delta^2 \\ -2\langle z_0 - x_0, y_0 - x_0 \rangle + 2\delta^2 \leq 0 \\ ||y_0 - x_0||^2 - 2\langle z_0 - x_0, y_0 - x_0 \rangle + ||z_0 - x_0||^2 \leq 0 \\ ||(y_0 - x_0) - (z_0 - x_0)||^2 \leq 0 \\ ||y_0 - z_0||^2 \leq 0 \end{gathered}$$

By definition $|| \cdot || \geq 0$ meaning

$$\begin{gathered} y_0 - z_0 = 0 \\ y_0 = z_0 \end{gathered}$$