STAT 280: OPTIMIZATION
SPRING 2019
PROBLEM SET 2

1. (a) Show that for any $A \in \mathbb{R}^{n \times n}$, the series

$$I + A + \frac{A^2}{2} + \cdots + \frac{A^k}{k!} + \cdots$$

is always convergent. The limit of this series is usually denoted as $\exp(A)$ and called matrix exponential of $A$.

(b) Show that the set of invertible matrices

$$\text{GL}(n) := \{ A \in \mathbb{R}^{n \times n} : \det(A) \neq 0 \}$$

is an open set in $\mathbb{R}^{n \times n}$.

(c) Show that the set of nilpotent matrices $\{ A \in \mathbb{R}^{n \times n} : A^k = 0 \text{ for some } k \in \mathbb{N} \cup \{0\} \}$ is a closed set in $\mathbb{R}^{n \times n}$.

(d) Show that the set of orthogonal matrices

$$\text{O}(n) := \{ A \in \mathbb{R}^{n \times n} : A^T A = I \}$$

is a compact set in $\mathbb{R}^{n \times n}$.

(e) Show that the set of symmetric positive definite matrices

$$\mathbb{S}_{++}^n := \{ A \in \mathbb{S}^n : x^T A x > 0 \text{ for all } 0 \neq x \in \mathbb{R}^n \}$$

is an open set in $\mathbb{S}^n := \{ A \in \mathbb{R}^{n \times n} : A^T = A \}$ and that its closure is the set of symmetric positive semidefinite matrices

$$\mathbb{S}_+^n := \{ A \in \mathbb{S}^n : x^T A x \geq 0 \text{ for all } x \in \mathbb{R}^n \}.$$

2. (a) Let $f : \mathbb{R}^n \to \mathbb{R}$ be in $C^1(\mathbb{R}^n)$. Suppose $f$ satisfies

$$\left| \frac{\partial f}{\partial x_i}(x) \right| \leq K, \quad i = 1, \ldots, n,$$

for all $x \in \mathbb{R}^n$ where $K > 0$ is a constant. Show that

$$|f(x) - f(y)| \leq \sqrt{n} K \|x - y\|_2$$

for all $x, y \in \mathbb{R}^n$ (Hint: Apply the univariate mean-value theorem $n$ times).

(b) Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x, y) = \begin{cases} 
1 - \cos \frac{x^2}{y} \sqrt{x^2 + y^2} & \text{if } y \neq 0, \\
0 & \text{if } y = 0.
\end{cases}$$

Show that $f$ is continuous but not differentiable at $(0,0)$. For which $v \in \mathbb{R}^2$ does the directional derivative $\partial_v f(0,0)$ exist?
Let $A : \mathbb{R} \to \mathbb{R}^{n \times n}$ be the matrix-valued function

$$A(x) = \begin{bmatrix} a_{11}(x) & \cdots & a_{1n}(x) \\ \vdots & & \vdots \\ a_{n1}(x) & \cdots & a_{nn}(x) \end{bmatrix}$$

where the functions $a_{ij} : \mathbb{R} \to \mathbb{R}$ are all in $C^1(\mathbb{R})$ for all $i, j = 1, \ldots, n$.

(a) Show that the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \text{tr} \left( A(x)^2 \right)$ is differentiable and that

$$f'(x) = 3 \text{tr} \left( A(x)^2 A'(x) \right)$$

where

$$A'(x) = \begin{bmatrix} a'_{11}(x) & \cdots & a'_{1n}(x) \\ \vdots & & \vdots \\ a'_{n1}(x) & \cdots & a'_{nn}(x) \end{bmatrix}.$$

(b) Let $n = 2$. Suppose $A(x) \in S^{2,+}_2$ for all $x \in \mathbb{R}$ and define $B : \mathbb{R} \to S^{2 \times 2}$ by $B(x) = A(x)^{-1}$. Show that

$$\frac{d}{dx} \log \det A(x) = \sum_{i,j=1}^2 a'_{ij}(x)b_{ij}(x).$$

This is actually true for arbitrary $n$. For $n = 1$, it reduces to $(\log a(x))' = a'(x)/a(x)$.

4. Find the gradients of the following functions without writing down a single partial derivative.

(a) $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) = \|x\|_2.$$

Find the set of points where $f$ is not differentiable.

(b) $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) = \frac{x^\top A x}{x^\top B x}$$

where $A, B \in S^{n,+}_n$.

(c) $f : S^{n,+}_n \to \mathbb{R}$ defined by

$$f(X) = \det(X).$$

5. Find the derivative of the following functions without writing down a single partial derivative.

(a) $f : \mathbb{R}^4 \to \mathbb{R}^{2 \times 2}$ defined by

$$f(x, y, z, w) = \begin{bmatrix} x & y \\ z & w \end{bmatrix}^2 + \begin{bmatrix} x & y \\ z & w \end{bmatrix}^\top = X^2 + X^\top.$$

(b) $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n \setminus \{0\}$ defined by

$$f(x) = \frac{x}{\|x\|}$$

where the norm is the Euclidean norm.

(c) $f : \mathbb{R}^{m \times n} \setminus \{0\} \to \mathbb{R}^{m \times n} \setminus \{0\}$ defined by

$$f(X) = \frac{X}{\|X\|}$$

where the norm is the Frobenius norm.

(d) $f : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ defined by

$$f(X) = X^4.$$

(e) $f : \text{GL}(n) \to \text{GL}(n)$ defined by

$$f(X) = X^{-1}.$$