STAT 280: OPTIMIZATION
SPRING 2022
PROBLEM SET 2

1. (a) Show that for any $X \in \mathbb{R}^{n\times n}$, the series
$$I + X + \frac{X^2}{2} + \cdots + \frac{X^k}{k!} + \cdots$$
is always convergent. The limit of this series is usually denoted as $\exp(X)$ and called matrix exponential of X.

(b) Show that the set of invertible matrices
$$\mathrm{GL}(n) := \{X \in \mathbb{R}^{n\times n} : \det(X) \neq 0\}$$
is an open set in $\mathbb{R}^{n\times n}$.

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

(d) Show that the set of orthogonal matrices
$$\mathrm{O}(n) := \{X \in \mathbb{R}^{n\times n} : X^\mathsf{T} X = I\}$$
is a compact set in $\mathbb{R}^{n\times n}$.

(e) Show that the set of symmetric positive definite matrices
$$\mathbb{S}^n_{++} := \{X \in \mathbb{S}^n : \mathbf{y}^\mathsf{T} X \mathbf{y} > 0 \text{ for all } \mathbf{0} \neq \mathbf{y} \in \mathbb{R}^n\}$$
is an open set in $\mathbb{S}^n := \{X \in \mathbb{R}^{n\times n} : X^\mathsf{T} = X\}$ and that its closure is the set of symmetric positive semidefinite matrices
$$\mathbb{S}^n_{+} := \{X \in \mathbb{S}^n : \mathbf{y}^\mathsf{T} X \mathbf{y} \geq 0 \text{ for all } \mathbf{y} \in \mathbb{R}^n\}.$$

2. (a) Let $f : \Omega \to \mathbb{R}$ where $\Omega \subseteq \mathbb{R}^{m\times n}$ is unbounded. Show that all sublevel sets of f are bounded if and only if
$$\lim_{k\to\infty} \|X_k\| = +\infty \quad \Longrightarrow \quad \lim_{k\to\infty} f(X_k) = +\infty,$$
assuming $X_k \in \Omega$ for all $k \in \mathbb{N}$. Deduce that if such an f is continuous and Ω is closed, then f has a global minimizer $X_* \in \Omega$.

(b) Let $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ be in $C^1(\mathbb{R}^{m\times n})$. Suppose f satisfies
$$\left|\frac{\partial f}{\partial x_{ij}}(X)\right| \leq K, \quad i = 1, \ldots, m,\ j = 1, \ldots, n,$$
for all $X \in \mathbb{R}^{m\times n}$ where $K > 0$ is a constant. Show that
$$|f(X) - f(Y)| \leq \sqrt{mn} K \|X - Y\|$$
for all $X, Y \in \mathbb{R}^{m\times n}$ (Hint: Apply the univariate mean-value theorem mn times).

(c) Define $f : \mathbb{R}^{2\times 2} \to \mathbb{R}$ by
$$f\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{cases} \left[1 - \cos\left(\dfrac{x^2+y^2+z^2}{t}\right)\right]\sqrt{x^2+y^2+z^2+t^2} & \text{if } t \neq 0, \\ 0 & \text{if } t = 0. \end{cases}$$

Date: April 16, 2022 (Version 1.2); due: April 28, 2022.

Show that f is continuous but not differentiable at $\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$. For which $V \in \mathbb{R}^{2\times 2}$ does the directional derivative $\partial_V f\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$ exist?

3. 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 $f(x) = \operatorname{tr}\left(A(x)^3\right)$ is differentiable and that

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

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 $A(x) \in \mathbb{S}^2_{++}$ for all $x \in \mathbb{R}$ and define $B : \mathbb{R} \to \mathbb{S}^{2\times 2}$ 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 calculating a single partial derivative.

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

$$f(\mathbf{x}) = \|\mathbf{x}\|.$$

Find the set of points where f is not differentiable.

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

$$f(\mathbf{x}) = \frac{\mathbf{x}^\mathsf{T} A\mathbf{x}}{\mathbf{x}^\mathsf{T} B\mathbf{x}}$$

where $A, B \in \mathbb{S}^n_{++}$.

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

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

(d) $f : \mathbb{S}^n_{++} \to \mathbb{R}$ defined by

$$f(X) = \frac{\det(X)}{\operatorname{tr}(X)}.$$

(e) $f : \mathbb{R}^n \times \mathbb{S}^n_{++} \to \mathbb{R}$ defined by

$$f(\mathbf{x}, Y) = \mathbf{x}^\mathsf{T} Y^{-1}\mathbf{x}.$$

(f) $f : \Omega \to \mathbb{R}$ defined by

$$\Omega = \{X \in \mathbb{R}^{m\times n} : X^\mathsf{T} AX + B^\mathsf{T} X + X^\mathsf{T} B + C \succ 0\}$$

with $A \in \mathbb{S}^m$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{S}^n$ arbitrary; and

$$f(X) = \log\det(X^\mathsf{T} AX + B^\mathsf{T} X + X^\mathsf{T} B + C).$$

5. Find the derivatives of the following functions without calculating a single partial derivative.

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

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

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

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

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

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

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

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

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

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

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

$$f(X) = \exp(X),$$

as in Problem **1**(a).