Math 350 Problem Set 2 (Part I) (due Friday 9/17 by 3pm)

You may work on the problems in groups, but please write up your answers individually and cite your collaborators on the top of your assignment. If you consult sources other than the textbook to help you find answers to these questions, please cite these sources in your answer.

1. Mean Value Theorems:

   (a) (10pts) Use the Mean Value Theorem from single variable calculus to prove the Mean Value Theorem below for scalar functions of several variables. A **convex set** $A \subset \mathbb{R}^n$ is a set such that, for any two points in A, the line segment between them is also in A:

   $$\mathbf{x}, \mathbf{y} \in A \implies \{(1-t)\mathbf{x} + t\mathbf{y} \mid 0 \leq t \leq 1\} \subset A$$

   Theorem (MVT)**.** *If $f : A \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable on a convex set A, for any pair of points $\mathbf{x}, \mathbf{y} \in A$, we have*

   $$f(\mathbf{x}) - f(\mathbf{y}) = \nabla f(\mathbf{z}) \cdot (\mathbf{x} - \mathbf{y}) = \mathbf{D}f(\mathbf{z})(\mathbf{x} - \mathbf{y})$$

   for some $\mathbf{z} \in \{(1-t)\mathbf{x} + t\mathbf{y} \mid 0 \leq t \leq 1\} \subset A$.

   (b) (10pts) Why is the analogous statement (using the $\mathbf{D}f$ form of the right hand side) false in general for a vector valued function $f : \mathbb{R}^n \to \mathbb{R}^m$ when $m > 1$?

2. (5pts) What is the derivative of a constant function $f : \mathbb{R}^n \to \mathbb{R}^m$, (so $(x_1, \ldots, x_n) \mapsto (f_1, \ldots, f_m)$, where each $f_i \in \mathbb{R}$ is independent of $\mathbf{x}$)? Prove your answer using the definition of the derivative; i.e. that the derivative is the unique linear function $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ such that

   $$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{T}(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0$$

3. (5pts) What is the derivative of a linear function $f : \mathbb{R}^n \to \mathbb{R}^m$, (so $f(a\mathbf{x} + b\mathbf{y}) = af(\mathbf{x}) + bf(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $a, b \in \mathbb{R}$)? Prove your answer using the definition of the derivative.

4. (10pts) A function $f : A \subset \mathbb{R}^n \to B \subset \mathbb{R}^n$ is said to be **invertible** if there exists a function (called the inverse) $f^{-1} : B \subset \mathbb{R}^n \to A \subset \mathbb{R}^n$ such that

   $$f^{-1} \circ f = \mathrm{Id} : A \to A \quad \text{and} \quad f \circ f^{-1} = \mathrm{Id} : B \to B,$$

   where Id is the **identity function** which maps each point to itself,

   $$\mathrm{Id} : \mathbb{R}^n \to \mathbb{R}^n, \quad \mathbf{x} \mapsto \mathbf{x}.$$

   (Note that the dimension of the domain and range must be the same, and note that $f^{-1}(\mathbf{x})$ does not mean $1/f(\mathbf{x})$ unless $n = 1$, since division does not make sense for $n > 1$.)

   Show that if $f : A \subset \mathbb{R}^n \to B \subset \mathbb{R}^n$ is invertible, and if f is differentiable at $\mathbf{x}_0 \in A$, then f^{-1} is differentiable at $\mathbf{y}_0 = f(\mathbf{x}_0)$ with derivative

   $$\mathbf{D}\left(f^{-1}\right)(\mathbf{y}_0) = \left(\mathbf{D}f(\mathbf{x}_0)\right)^{-1}.$$

   (Hint: use the chain rule and your result from problem 3).

5. (10pts) In class (and in Ch. 2.6 in the book), we discussed the product rule

   $$\mathbf{D}(fg)(\mathbf{x}_0) = g(\mathbf{x}_0)\mathbf{D}f(\mathbf{x}_0) + f(\mathbf{x}_0)\mathbf{D}g(\mathbf{x}_0)$$

   for scalar functions $f, g : \mathbb{R}^n \to \mathbb{R}$, where the product $fg : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{x} \mapsto f(\mathbf{x})g(\mathbf{x})$ is ordinary multiplication in $\mathbb{R}$.

For vector valued functions $f, g : \mathbb{R}^n \to \mathbb{R}^m$, we can form the dot product function (note that it is scalar valued!)

$$f \cdot g : \mathbb{R}^n \to \mathbb{R}, \quad \mathbf{x} \mapsto f(\mathbf{x}) \cdot g(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x}) g_i(\mathbf{x})$$

where the product is the dot product in $\mathbb{R}^m$. Formulate and prove a product rule for the dot product. You may use the scalar version of the product rule in your proof, and any other properties of derivatives discussed in class. (Hint: in formulating your result in a concise manner, you may be interested in the **transpose** operation on matrices $A \mapsto A^{\mathrm{T}}$. If A is an $r \times s$ matrix, A^{T} is a $s \times r$ matrix with $(A^{\mathrm{T}})_{ij} = A_{ji}$, that is, the rows and columns are swapped with one another.)