I. Differentiability of functions from $\mathbb{R}^n$ to $\mathbb{R}$

1. We saw that the definition of derivative of one variable functions as $\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$ (instantaneous rate of change) does not lend itself to a generalization to higher dimensions. Firstly, the failure is because there are more than one such rates that can be asked. Secondly, even with all partial derivatives existing, the function need not even be continuous. So we take a different look at the derivative of a one variable function as follows. We call a real number α to be the derivative of f at a point x if $\lim_{h\to 0}\frac{|f(x+h)-f(x)-\alpha\cdot h|}{|h|}=0$.

   (a) α gives us a degree-one polynomial function of h, defined by $g(h)=f(x)+\alpha\cdot h$ for $h\in\mathbb{R}$. This function g, for small h, best approximates the value of the function f at a point $x+h$ near x in the above limit-sense. α can be thought of as a function acting on the error h to give a real number $\alpha\cdot h$.

   (b) α is also seen as the slope of the tangent line to the graph of the function f at the point x in the domain, *i.e.* at the point $(x,f(x))$ on the graph.

   Using this as a cue, we define the derivative of a real-valued function of several variables.

2. *Definition:* Let $\mathbf{x}\in A\subset\mathbb{R}^n$ be a point such that A contains an ϵ-ball around $\mathbf{x}$ for some $\epsilon>0$. We say that a function $f:A\subset\mathbb{R}^n\to\mathbb{R}$ is <u>differentiable</u> at the point $\mathbf{x}$ if there is a vector $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)\in\mathbb{R}^n$ such that $\lim_{\mathbf{h}\to 0}\frac{|f(\mathbf{x}+\mathbf{h})-f(\mathbf{x})-\alpha\cdot\mathbf{h}|}{\|\mathbf{h}\|}=0$. When the function is differentiable, this vector $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ is called the <u>derivative</u> of f at the point $\mathbf{x}$. The derivative of f at the point $\mathbf{x}$ is denoted by $Df(\mathbf{x})$.

3. Analogous to its one-variable counterpart, we have the following:

   (a) This vector α gives us a homogeneous, degree-one polynomial in n variables, defined by $g(h)=f(\mathbf{x})+\alpha\cdot\mathbf{h}$ for $\mathbf{h}\in\mathbb{R}^n$. This function g, (although defined for any $\mathbf{h}\in\mathbb{R}^n$) for $\mathbf{h}$ with small $\|\mathbf{h}\|$, best approximates the function f at a point $\mathbf{x}+\mathbf{h}$ near $\mathbf{x}$ in the above limit-sense. α can be thought of as a function acting on the error $\mathbf{h}$ to give a real number $\alpha\cdot\mathbf{h}$. To emphasize this point, $Df(\mathbf{x}):\mathbb{R}^2\to\mathbb{R}$ is the function $Df(\mathbf{x})(\mathbf{h})=\alpha\cdot\mathbf{h}$.

   (b) Here too, like functions of one-variable, $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)$, can be thought of as the "n-dimensional slope" to the graph of f in $\mathbb{R}^{n+1}$. For example when $n=2$, α indicates the plane in $\mathbb{R}^3$ given by $z=\alpha_1 x+\alpha_2 y+constant$.

   (c) If f is differentiable at all points in its domain, then it is simply called a differentiable function.

4. Example: For $(x,y)\in\mathbb{R}^2$, let $f(x,y)=x^2-y^2$. For this function f at $(1,3)$ the vector $\alpha=(2,-6)$ works as the derivative.

5. *Proposition:* The derivative of f (as defined above), when it exists, is unique. (proof is easy).

6. *Theorem:* Suppose that $f:A\subset\mathbb{R}^n\to\mathbb{R}$ is differentiable at a point $(x_1,x_2,\ldots,x_n)\in A$. Then the vector $\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\ldots,\frac{\partial f}{\partial x_n}\right)$ gives the derivative of f at the point $(x_1,x_2,\ldots,x_n)$.

   Proof: Since the derivative of f exists, there is a vector $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)$, as in the definition of the derivative, at the point $(x_1,x_2,\ldots,x_n)$. If the error vector is chosen along the first co-ordinate direction as $\mathbf{h}=(h,0,\ldots,0)$, then the definition of derivative gives that

   $$\lim_{h\to 0}\frac{|f(x_1+h,x_2,\ldots,x_n)-f(x_1,x_2,\ldots,x_n)-\alpha_1\cdot h|}{|h|}=0$$

   But then, from the definition of the first partial derivative of f, $\frac{\partial f}{\partial x_1}(x_1,x_2,\ldots,x_n)$ also plays the role of α_1. So, by above proposition, $\alpha_1=\frac{\partial f}{\partial x_1}(x_1,x_2,\ldots,x_n)$.

7. Given a function $f:A\subset\mathbb{R}^2\to\mathbb{R}$, we saw that its graph $z=f(x,y)$ can be pictured in $\mathbb{R}^3$. Much like the tangent line to the graph of a function in $\mathbb{R}^2$, we can associate a tangent plane to the graph of $f(x,y)$ in $\mathbb{R}^3$ at a point $(x_0,y_0)\in A$, which corresponds to the point $(x_0,y_0,f(x_0,y_0))$ on the graph.

8. Suppose that the function $f:A\subset\mathbb{R}^2\to\mathbb{R}$ is differentiable at a point $(x_0,y_0)\in A$. Define the tangent plane to the graph of f in $\mathbb{R}^3$ at a point $(x_0,y_0,f(x_0,y_0))$ where $(x_0,y_0)\in A$ as the plane given by $z=\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)+f(x_0,y_0)$.

9. This immediately tells us that the normal to the tangent plane is the vector $\left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1\right)$.

10. Example: To find the equation of the tangent plane to the graph of the function given by $f(x, y) = x^2 - y^2$, where $(x, y) \in \mathbb{R}^2$, at the point $(1, 3)$, note that $(2, -6)$ served as the derivative. So the equation of the tangent plane is $z = 2(x-1) - 6(y-3) - 8$, $i.e. z = 2x - 6y + 8$.

II. Differentiability of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

1. Recall that $Df(\mathbf{x})(\mathbf{h}) = \alpha \cdot \mathbf{h}$, which can also be written as the matrix multiplication

$$\left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right] \cdot \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}$$

2. It is this view point of Df that allows us to generalize the definition of derivative to an m-component function of n-variables.

3. *Definition:* Let $\mathbf{x} \in A \subset \mathbb{R}^n$ be a point such that A contains an ϵ-ball around $\mathbf{x}$ for some $\epsilon > 0$. We say that a function $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ is <u>differentiable</u> at the point $\mathbf{x}$ if there is an $m \times n$ matrix B such that $\lim_{\mathbf{h} \to 0} \frac{\|f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - B \cdot \mathbf{h}\|}{\|\mathbf{h}\|} = 0$. When the function f is differentiable, this matrix B is called the <u>derivative</u> of f at the point $\mathbf{x}$. The derivative of f at the point $\mathbf{x}$ is denoted by $Df(\mathbf{x})$.

4. *Theorem:* When the derivative of $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ exists at a point $\mathbf{x} \in A$, the derivative B in the definition is the matrix of partial derivatives of f as follows. The $(i, j)^{th}$ entry of the matrix B call it b_{ij}, is given by $b_{ij} = \frac{\partial f_i}{\partial x_j}$, where f_i denotes the i^{th} component function of f and $\frac{\partial f_i}{\partial x_j}$ is its partial derivative w.r.t. the j^{th} variable.

5. Example 1: For a function $f : A \subset \mathbb{R} \to \mathbb{R}$, the size of the derivative matrix is 1×1, *i.e.* it is a single number, which is signified by the slope of the tangent line to the graph of the function or by the instantaneous rate of change of the function.

6. Example 2: Consider $f(x, y) = (x^2 + 3y, e^{x/y}, 4\cos xy)$, where $(x, y) \in \mathbb{R}^2$, $y \neq 0$. Let us call the component functions of f as f_1, f_2, f_3 respectively. The matrix

$$\begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & 3 \\ \frac{e^{x/y}}{y} & -\frac{xe^{x/y}}{y^2} \\ -4y\sin xy & -4x\sin xy \end{bmatrix}$$

evaluated at a given point (a, b) in the domain of f, gives the derivative function as a matrix at that point (a, b).

7. *Proposition:* If the derivative of $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ exists at a point $\mathbf{x} \in A$, then the function f is continuous at $\mathbf{x}$. Proof:

8. *Proposition:* If all the partial derivatives of a function exist at a point $\mathbf{x}$ and are continuous in a neighborhood of $\mathbf{x}$, then f is differentiable at $\mathbf{x}$. Proof: omitted.

9. Derivative Rules: Suppose that $A \subset \mathbb{R}^n$ and that $f : A \to \mathbb{R}^m$ and $g : A \to \mathbb{R}^m$ are both differentiable at a point $\mathbf{a} \in A$. Let $c \in \mathbb{R}$. Then the following hold:

   (a) The function $h : A \to \mathbb{R}^m$ defined as $h(\mathbf{x}) = c \cdot f(\mathbf{x})$ is differentiable at $\mathbf{a}$ and $Dh(\mathbf{a}) = cDf(\mathbf{a})$.

   (b) The function $h : A \to \mathbb{R}^m$ defined as $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ is differentiable at $\mathbf{a}$ and $Dh(\mathbf{a}) = Df(\mathbf{a}) + Dg(\mathbf{a})$.