Solutions to Exam 2

1. Let U denote an open subset of $\mathbb{R}^n$, and let $F\colon U \to \mathbb{R}^m$ be a vector field on U.

   (a) State precisely what it means for F to be differentiable at $u \in U$.

   Answer: The vector valued function F is said to be differentiable at $u \in U$ iff there exists a liner transformation $T_{_u}\colon \mathbb{R}^n \to \mathbb{R}^m$ such that

   $$F(u+w) = F(u) + T_{_u}(w) + E_{_u}(w),$$

   for $w \in \mathbb{R}^n$ with $\|w\|$ sufficiently small, where

   $$\lim_{\|w\|\to 0} \frac{\|E_{_u}(w)\|}{\|w\|} = 0.$$

   □

   (b) Suppose that a vector field $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear map. Prove that F is differentiable at every $u \in \mathbb{R}^n$, and compute its derivative map,

   $$DF(u)\colon \mathbb{R}^n \to \mathbb{R}^m,$$

   at u, for all $u \in \mathbb{R}^n$.

   Solution: Since F is linear, we have that

   $$F(u+w) = F(u) + F(w) \quad \text{for all } u, w \in \mathbb{R}^n.$$

   Thus, setting $T_{_u}(w) = F(w)$, we see that

   $$E_{_u}(w) = F(u+w) - F(u) - T_{_u}(w) = F(u) + F(w) - F(u) - F(w) = \mathbf{0}$$

   for all $w \in \mathbb{R}^n$. Consequently,

   $$\frac{\|E_{_u}(w)\|}{\|w\|} = 0 \quad \text{for all } w \neq \mathbf{0}.$$

   We therefore see that

   $$F(u+w) = F(u) + T_{_u}(w) + E_{_u}(w)$$

   where $T_u = F$ and

   $$\lim_{\|w\|\to 0} \frac{\|E_{_u}(w)\|}{\|w\|} = 0.$$

In other words, F is differentiable at every $u \in \mathbb{R}^n$, and its derivative map at u is F for all $u \in \mathbb{R}^n$; that is,

$$DF(u)w = F(w) \quad \text{for all } w \in \mathbb{R}^n.$$

$\square$

2. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ denote the scalar field on $\mathbb{R}^2$ defined by

$$f(x,y) = x^{2/3}y^{1/3} \quad \text{ for all } \ (x,y) \in \mathbb{R}^2.$$

(a) Show that the partial derivatives of f at $(0,0)$ exist and compute them.

Solution: For $h \in \mathbb{R}$ with $h \neq 0$, compute

$$\frac{f(0+h,0) - f(0,0)}{h} = \frac{0-0}{h} = 0,$$

so that

$$\lim_{h\to 0} \frac{f(0+h,0) - f(0,0)}{h} = 0.$$

This shows that $\dfrac{\partial f}{\partial x}(0,0)$ exists and

$$\frac{\partial f}{\partial x}(0,0) = 0.$$

Similarly,

$$\frac{\partial f}{\partial y}(0,0) = 0.$$

$\square$

(b) Show that f is not differentiable at $(0,0)$.

Solution: If f was differentiable at $(0,0)$, then we would have that

$$f(h,k) = f(0,0) + Df(0,0)\begin{pmatrix} h \\ k \end{pmatrix} + E(h,k),$$

where E has the property that

$$\lim_{(h,k)\to(0,0)} \frac{|E(h,k)|}{\sqrt{h^2+k^2}} = 0, \tag{1}$$

and

$$Df(0,0) = \begin{bmatrix} \dfrac{\partial f}{\partial x}(0,0) & \dfrac{\partial f}{\partial y}(0,0) \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}.$$

It then follows that

$$f(h,k) = E(h,k),$$

and therefore, for $(h,k) \neq (0,0)$,

$$\frac{|E(h,k)|}{\sqrt{h^2+k^2}} = \frac{|h|^{2/3}|k|^{1/3}}{\sqrt{h^2+k^2}}.$$

In particular, if we take $h = k = t$, where $t \in \mathbb{R}$ and $t \neq 0$, we get that

$$\frac{|E(t,t)|}{\sqrt{t^2+t^2}} = \frac{1}{\sqrt{2}},$$

which contradicts (1). This contradiction shows that f is not differentiable at $(0,0)$. □

3. For fixed vectors v and u in $\mathbb{R}^n$, define $\sigma\colon \mathbb{R} \to \mathbb{R}^n$ by

$$\sigma(t) = u + tv \quad \text{for all } t \in \mathbb{R}.$$

Prove that σ is differentiable at every $t \in \mathbb{R}$ and compute its derivative map, $D\sigma(t)$, for all $t \in \mathbb{R}$

Proof: Fix $t \in \mathbb{R}$. For any $h \in \mathbb{R}$,

$$\sigma(t+h) = u + (t+h)v = u + tv + hv = \sigma(t) + hv.$$

It then follows that

$$\sigma(t+h) = \sigma(t) + D\sigma(t)h + E(h), \quad \text{for all } h \in \mathbb{R},$$

where

$$D\sigma(t)h = hv, \quad \text{for all } h \in \mathbb{R},$$

is a linear map from $\mathbb{R}$ to $\mathbb{R}^n$, and

$$E(h) = \mathbf{0} \quad \text{for all } h \in \mathbb{R}.$$

Consequently,

$$\lim_{|h|\to 0}\frac{\|E(h)\|}{|h|}=0.$$

We have therefore proved that $\sigma(t)=u+tv$ is differentiable at every $t\in\mathbb{R}$, and the derivative map, $D\sigma(t)\colon\mathbb{R}\to\mathbb{R}^n$ is given by

$$D\sigma(t)h=hv,\quad\text{for all } h\in\mathbb{R}.$$

□

4. Let I denote an open interval of real numbers and define $\sigma\colon\ I\to\mathbb{R}^2$ by

$$\sigma(t)=(x(t),y(t))\quad\text{for all } t\in I,$$

where $x(t)$ and $y(t)$ are real valued functions of $t\in I$.

Prove that σ is differentiable at $t\in I$ if and only if both x and y are differentiable at $t\in I$. Furthermore, $D\sigma(t)h=h(x'(t),y'(t))$ for all $h\in\mathbb{R}$.

Proof: Assume first that σ is differentiable at $t\in I$. Then, there exists a linear map $D\sigma(t)\colon\mathbb{R}\to\mathbb{R}^2$ for which

$$\sigma(t+h)=\sigma(t)+D\sigma(t)h+E(h),\quad\text{for } |h| \text{ small enough in } \mathbb{R},\tag{2}$$

and

$$\lim_{|h|\to 0}\frac{\|E(h)\|}{|h|}=0.\tag{3}$$

Since $D\sigma(t)$ is linear, there exists a vector (a_1,a_2) in $\mathbb{R}^2$ such that

$$D\sigma(t)h=h(a_1,a_2)\quad\text{for all }\ h\in\mathbb{R}.\tag{4}$$

We then have from equation (2) that, for $|h|$ sufficiently small,

$$(x(t+h),y(t+h))=(x(t),y(t))+h(a_1,a_2)+E(h).$$

Thus, for $|h|$ small enough and $h\neq 0$,

$$\left(\frac{x(t+h)-x(t)}{h},\frac{y(t+h)-y(t)}{h}\right)=(a_1,a_2)+\frac{1}{h}E(h),$$

or

$$\left(\frac{x(t+h)-x(t)}{h}-a_1,\frac{y(t+h)-y(t)}{h}-a_2\right)=\frac{1}{h}E(h).\tag{5}$$

Taking norms on both sides of equation (5) we obtain that

$$\sqrt{\left(\frac{x(t+h)-x(t)}{h}-a_1\right)^2+\left(\frac{y(t+h)-y(t)}{h}-a_2\right)^2}=\frac{\|E(h)\|}{|h|}.$$

Consequently,

$$\left|\frac{x(t+h)-x(t)}{h}-a_1\right|\leqslant\frac{\|E(h)\|}{|h|}.$$

It then follows from (3) and the Squeeze theorem that

$$\lim_{h\to 0}\left|\frac{x(t+h)-x(t)}{h}-a_1\right|=0,$$

which shows that x is differentiable at t and $x'(t)=a_1$. Similarly, we deduce that y is differentiable at t and $y'(t)=a_2$. Thus, by equation (4),

$$D\sigma(t)h=h(x'(t),y'(t))\quad\text{for all } h\in\mathbb{R}.$$

Conversely, suppose that the functions x and y are differentiable at t and define

$$E(h)=\sigma(t+h)-\sigma(t)-h(x'(t),y'(t))$$

for $h\in\mathbb{R}$ with $|h|$ sufficiently small so that $t+h\in I$. Then,

$$E(h)=(x(t+h)-x(t),y(t+h)-y(t))-h(x'(t),y'(t)).$$

Thus, for $h\neq$ and $|h|$ small enough,

$$\frac{1}{h}E(h)=\left(\frac{x(t+h)-x(t)}{h}-x'(t),\frac{y(t+h)-y(t)}{h}-y'(t)\right). \tag{6}$$

Taking norms on both sides of (6) and squaring, we obtain

$$\frac{\|E(h)\|^2}{|h|^2}=\left(\frac{x(t+h)-x(t)}{h}-x'(t)\right)^2+\left(\frac{y(t+h)-y(t)}{h}-y'(t)\right)^2. \tag{7}$$

Thus, since

$$x'(t)=\lim_{h\to 0}\frac{x(t+h)-x(t)}{h}$$

and

$$y'(t)=\lim_{h\to 0}\frac{y(t+h)-y(t)}{h},$$

it follows from (7) that

$$\lim_{h\to 0} \frac{\|E(h)\|^2}{|h|^2} = 0,$$

from which we get that

$$\lim_{h\to 0} \frac{\|E(h)\|}{|h|} = 0.$$

We therefore conclude that σ is differentiable at t. □

5. Let U denote an open subset of $\mathbb{R}^n$ and Q an open subset of $\mathbb{R}^m$. Consider the maps $F\colon U \to \mathbb{R}^m$ and $G\colon Q \to \mathbb{R}^k$.

   (a) State the Chain Rule in the context of the functions F and G and the open sets given above. Be explicit as to what your assumptions and conclusions are.

   Answer: Assume that $F(U) \subseteq Q$, and that F is differentiable at $u \in U$ and G is differentiable at $F(u)$. Then the composite map

   $$G \circ F\colon U \to \mathbb{R}^k$$

   is differentiable at u and the derivative map of $G \circ F$ at u is

   $$D(G \circ F)(u) = DG(F(u)) \cdot DF(u).$$

   □

   (b) Use the Chain Rule to prove the following: If f is a differentiable scalar field on an open set $U \in \mathbb{R}^n$ and $\sigma\colon I \to \mathbb{R}^n$ is a differentiable path such that $\sigma(I) \subseteq U$, then the function $g\colon I \to \mathbb{R}$ defined by

   $$g(t) = f(\sigma(t)) \quad \text{for all } t \in I,$$

   is differentiable on I. Give a formula for computing $g'(t)$ for all $t \in I$ in terms of the gradient of f and $\sigma'(t)$.

   Solution: Observe that $g = f \circ \sigma$. It then follows by the Chain Rule that g is differentiable on I and

   $$\begin{aligned} g'(t) &= Df(\sigma(t)) \cdot D\sigma(t) \\ &= \nabla f(\sigma(t)) \cdot \sigma'(t). \end{aligned}$$

   □

(c) Use your result from the previous part to prove that, if f is differentiable at $u \in U$, then the limit

$$\lim_{t\to 0} \frac{f(u+t\widehat{v}) - f(u)}{t},$$

where $\widehat{v}$ denotes a unit vector, exists. Give an interpretation to your result.

Solution: Define $\sigma\colon \mathbb{R} \to \mathbb{R}^n$ by $\sigma(t) = u + t\widehat{v}$ for all $t \in \mathbb{R}$, and put $g(t) = f(\sigma(t)) = f(u + t\widehat{v})$. Then, $g = f \circ \sigma$ and σ is differentiable, by the result of Problem 3 in this exam. It then follows by the Chain Rule that g is differentiable on I and, by the result of the previous part,

$$\begin{aligned} g'(t) &= \nabla f(\sigma(t)) \cdot \sigma'(t) \\ &= \nabla f(u + t\widehat{v}) \cdot \widehat{v}, \quad \text{for all } \ t \in \mathbb{R}, \end{aligned}$$

where we have used the result of Problem 3 again. In particular, we get that

$$g'(0) = \nabla f(u) \cdot \widehat{v},$$

which shows that the limit

$$\lim_{t\to 0} \frac{f(u+t\widehat{v}) - f(u)}{t}$$

exists. This limit gives the rate of change of f at u in the direction of $\widehat{v}$. Thus, the rate of change of f at u in the direction of $\widehat{v}$ is the component of the orthogonal projection of the gradient of f at u along the vector $\widehat{v}$.

□

(d) Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x,y) = 3xy - y^2$ for all $(x,y) \in \mathbb{R}^2$. Compute the gradient of f at the point $(2,1)$, and find a direction, $\widehat{v}$, along which f is increasing the fastest at $(2,1)$. Justify your result.

Solution: Compute

$$\nabla f(x,y) = 3y\widehat{i} + (3x - 2y)\widehat{j}$$

for $(x,y) = (2,1)$ to get that

$$\nabla f(2,1) = 3\widehat{i} + 4\widehat{j}.$$

The rate of change of f at $(2,1)$ along the direction of $\widehat{v}$ is $\nabla f(2,1)\cdot\widehat{v}$ by the result of the previous part. By the Cauchy–Schwarz inequality,

$$-\|\nabla f(2,1)\| \leqslant \nabla f(2,1)\cdot\widehat{v} \leqslant \|\nabla f(2,1)\|$$

and the largest value of $\nabla f(2,1)\cdot\widehat{v}$, namely $\|\nabla f(2,1)\|$, is attained when $\widehat{v}$ is a unit vector in the direction of $\nabla(2,1)$. Thus,

$$\widehat{v} = \frac{1}{\|\nabla f(2,1)\|}\nabla f(2,1) = \frac{3}{5}\widehat{i} + \frac{4}{5}\widehat{j}.$$

□

6. Let f be a scalar field in $\mathbb{R}^n$ defined by $f(v) = \|v\|^2$ for all $v \in \mathbb{R}^n$.

(a) Prove that f is differentiable on $\mathbb{R}^n$ and use this fact to prove that the function $g\colon I \to \mathbb{R}$ defined by

$$g(t) = \|\sigma(t)\|^2, \quad \text{for all } t \in I,$$

where the path σ is differentiable on an open interval I, is differentiable and compute $g'(t)$.

Solution: Write v in terms of its coordinates, $(x_1, x_2, \ldots, x_n)$, to get that

$$f(v) = x_1^2 + x_2^2 + \cdots + x_n^2.$$

Thus, the partial derivatives of f are

$$\frac{\partial f}{\partial x_j}(v) = 2x_j, \quad \text{for } j = 1, 2, \ldots, n,$$

which are continuous functions. It then follows that f is C^1 and is therefore differentiable with derivative gradient

$$\nabla f(v) = (2x_1, 2x_2, \ldots, 2x_n) = 2v, \quad \text{for all } v \in \mathbb{R}^n.$$

Next, observe that $g(t) = f(\sigma(t)) = f \circ \sigma(t)$ for all $t \in I$. It then follows by the Chain Rule that g is differentiable on I and

$$g'(t) = \nabla f(\sigma(t)) \cdot \sigma'(t) \quad \text{for all } t \in I,$$

or

$$g'(t) = 2\sigma(t) \cdot \sigma'(t) \quad \text{for all } t \in I.$$

□

(b) Let g be the function defined in part (a) above. Prove that if g has a critical point at $t_o \in I$, then $\sigma(t_o)$ and $\sigma'(t_o)$ are orthogonal (or perpendicular) to each other.

Solution: If g has a critical point at $t_o \in I$, then $g'(t_o) = 0$. We then get, by the result in part (a), that

$$2\sigma(t_o) \cdot \sigma'(t_o) = 0,$$

which shows that $\sigma(t_o)$ and $\sigma'(t_o)$ are orthogonal to each other. □

(c) Find the point (or points) along the path in $\mathbb{R}^2$ given by

$$\sigma(t) = (t, t^2 - 1), \qquad \text{for } t \in \mathbb{R},$$

which are the closest to the origin $(0, 0)$ in $\mathbb{R}^2$.

Solution: $\sigma(t)$ is closest to the origin when $\|\sigma(t)\|$ is the smallest possible. This might occur when $g(t) = \|\sigma(t)\|^2$ has a minimum or, equivalently, when $\sigma(t)$ and $\sigma'(t)$ are orthogonal to each other, by the result of part (b).
Compute $\sigma'(t) = (1, 2t)$ and $\sigma(t) \cdot \sigma'(t) = t + 2t(t^2 - 1)$. Then, $\sigma(t)$ might be closest to the origin when

$$t + 2t(t^2 - 1) = 0,$$

or

$$t(1 + 2t^2 - 2) = 0,$$

which yields the critical points

$$t = 0 \quad \text{or} \quad t = \pm\frac{\sqrt{2}}{2}.$$

Evaluating $g(t) = \|\sigma(t)\|^2 = t^2 + (t^2 - 1)^2$ at these critical points we obtain that

$$g(0) = 1 \quad \text{and} \quad g\left(\pm\frac{\sqrt{2}}{2}\right) = \frac{3}{4}.$$

Hence, the points $\left(\pm\frac{\sqrt{2}}{2}, -\frac{1}{2}\right)$ are the closest to the origin. □