Solutions to Review Problems for Exam 2

1. Define the scalar field $f\colon \mathbb{R}^n \to \mathbb{R}$ by $f(v) = \dfrac{1}{2}\|v\|^2$ for all $v \in \mathbb{R}^n$. Show that f is differentiable on $\mathbb{R}^n$ and compute the linear map $Df(u)\colon \mathbb{R}^n \to \mathbb{R}$ for all $u \in \mathbb{R}^n$. What is the gradient of f at u for all $x \in \mathbb{R}^n$?

   ***Solution**:* Let u and w be any vector in $\mathbb{R}^n$ and consider

$$\begin{aligned} f(u+w) &= \frac{1}{2}\|u+w\|^2 \\ &= \frac{1}{2}(u+w)\cdot(u+w) \\ &= \frac{1}{2}u\cdot u + u\cdot w + \frac{1}{2}w\cdot w \\ &= \frac{1}{2}\|u\|^2 + u\cdot w + \frac{1}{2}\|w\|^2. \end{aligned}$$

   Thus,

$$f(u+w) - f(u) - u\cdot w = \frac{1}{2}\|w\|^2.$$

   Consequently,

$$\frac{|f(u+w) - f(u) - u\cdot w|}{\|w\|} = \frac{1}{2}\|w\|,$$

   from which we get that

$$\lim_{\|w\|\to 0} \frac{|f(u+w) - f(u) - u\cdot w|}{\|w\|} = 0,$$

   and therefore f is differentiable at u with derivative map $Df(u)$ given by

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

   Hence, $\nabla f(u) = u$ for all $u \in \mathbb{R}^n$. □

2. Let $g\colon\ [0,\infty) \to \mathbb{R}$ be a differentiable, real–valued function of a single variable, and let $f(x,y) = g(r)$ where $r = \sqrt{x^2+y^2}$.

(a) Compute $\dfrac{\partial r}{\partial x}$ in terms of x and r, and $\dfrac{\partial r}{\partial y}$ in terms of y and r.

***Solution**:* Take the partial derivative of $r^2 = x^2 + y^2$ on both sides with respect to x to obtain

$$\frac{\partial(r^2)}{\partial x} = 2x.$$

Applying the chain rule on the left–hand side we get

$$2r\frac{\partial r}{\partial x} = 2x,$$

which leads to

$$\frac{\partial r}{\partial x} = \frac{x}{r}.$$

Similarly, $\dfrac{\partial r}{\partial y} = \dfrac{y}{r}$. □

(b) Compute ∇f in terms of $g'(r)$, r and the vector $\mathbf{r} = x\widehat{i} + y\widehat{j}$.

***Solution**:* Take the partial derivative of $f(x,y) = g(r)$ on both sides with respect to x and apply the Chain Rule to obtain

$$\frac{\partial f}{\partial x} = g'(r)\frac{\partial r}{\partial x} = g'(r)\frac{x}{r}.$$

Similarly, $\dfrac{\partial f}{\partial y} = g'(r)\dfrac{y}{r}$.
It then follows that

$$\begin{aligned}
\nabla f &= \frac{\partial f}{\partial x}\widehat{i} + \frac{\partial f}{\partial y}\widehat{j} \\
&= g'(r)\frac{x}{r}\widehat{i} + g'(r)\frac{y}{r}\widehat{j} \\
&= \frac{g'(r)}{r}(x\widehat{i} + y\widehat{j}) \\
&= \frac{g'(r)}{r}\mathbf{r}.
\end{aligned}$$

□

3. Let $f \colon U \to \mathbb{R}$ denote a scalar field defined on an open subset U of $\mathbb{R}^n$, and let $\widehat{u}$ be a unit vector in $\mathbb{R}^n$. If the limit

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

exists, we call it the *directional derivative of f at v in the direction of the unit vector $\widehat{u}$*. We denote it by $D_{\widehat{u}}f(v)$.

(a) Show that if f is differentiable at $v \in U$, then, for any unit vector $\widehat{u}$ in $\mathbb{R}^n$, the directional derivative of f in the direction of $\widehat{u}$ at v exists, and

$$D_{\widehat{u}}f(v) = \nabla f(v) \cdot \widehat{u},$$

where $\nabla f(v)$ is the gradient of f at v.

Proof: Suppose that f is differentiable at $v \in U$. Then,

$$f(v + w) = f(v) + Df(v)w + E(w),$$

where

$$Df(v)w = \nabla f(v) \cdot w,$$

and

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

Thus, for any $t \in \mathbb{R}$,

$$f(v + t\widehat{u}) = f(v) + t\nabla f(v) \cdot \widehat{u} + E(t\widehat{u}),$$

where

$$\lim_{|t| \to 0} \frac{|E(t\widehat{u})|}{|t|} = 0,$$

since $\|t\widehat{u}\| = |t|\|\widehat{u}\| = |t|$.
We then have that, for $t \neq 0$,

$$\frac{f(v + t\widehat{u}) - f(v)}{t} - \nabla f(v) \cdot \widehat{u} = \frac{E(t\widehat{u})}{t},$$

and consequently

$$\left| \frac{f(v + t\widehat{u}) - f(v)}{t} - \nabla f(v) \cdot \widehat{u} \right| = \frac{|E(t\widehat{u})|}{|t|},$$

from which we get that

$$\lim_{t \to 0} \left| \frac{f(v + t\widehat{u}) - f(v)}{t} - \nabla f(v) \cdot \widehat{u} \right| = 0.$$

□

(b) Suppose that $f\colon U \to \mathbb{R}$ is differentiable at $v \in U$. Prove that if $D_{\widehat{u}}f(v) = 0$ for every unit vector $\widehat{u}$ in $\mathbb{R}^n$, then $\nabla f(v)$ must be the zero vector.

Proof: Suppose, by way of contradiction, that $\nabla f(v) \neq \mathbf{0}$, and put

$$\widehat{u} = \frac{1}{\|\nabla f(v)\|}\nabla f(v).$$

Then, $\widehat{u}$ is a unit vector, and therefore, by the assumption,

$$D_{\widehat{u}}f(v) = 0,$$

or

$$\nabla f(v) \cdot \widehat{u} = 0.$$

But this implies that

$$\nabla f(v) \cdot \frac{1}{\|\nabla f(v)\|}\nabla f(v) = 0,$$

where

$$\begin{aligned} \nabla f(v) \cdot \frac{1}{\|\nabla f(v)\|}\nabla f(v) &= \frac{1}{\|\nabla f(v)\|}\nabla f(v) \cdot \nabla f(v) \\ &= \frac{1}{\|\nabla f(v)\|}\|\nabla f(v)\|^2 \\ &= \|\nabla f(v)\|. \end{aligned}$$

It then follows that $\|\nabla f(v)\| = 0$, which contradicts the assumption that $\nabla f(v) \neq \mathbf{0}$. Therefore, $\nabla f(v)$ must be the zero vector. □

(c) Suppose that $f\colon U \to \mathbb{R}$ is differentiable at $v \in U$. Use the Cauchy–Schwarz inequality to show that the largest value of $D_{\widehat{u}}f(v)$ is $\|\nabla f(v)\|$ and it occurs when $\widehat{u}$ is in the direction of $\nabla f(v)$.

Proof. If f is differentiable at x, then $D_{\widehat{u}}f(x) = \nabla f(x) \cdot \widehat{u}$, as was shown in part (a). Thus, by the Cauchy–Schwarz inequality,

$$|D_{\widehat{u}}f(x)| \leqslant \|\nabla f(x)\|\|\widehat{u}\| = \|\nabla f(x)\|,$$

since $\widehat{u}$ is a unit vector. Hence,

$$-\|\nabla f(x)\| \leqslant D_{\widehat{u}}f(x) \leqslant \|\nabla f(x)\|$$

for any unit vector $\widehat{u}$, and so the largest value that $D_{\widehat{u}}f(x)$ can have is $\|\nabla f(x)\|$.

If $\nabla f(x) \neq \mathbf{0}$, then $\widehat{u} = \dfrac{1}{\|\nabla f(x)\|}\nabla f(x)$ is a unit vector, and

$$\begin{aligned} D_{\widehat{u}}f(x) &= \nabla f(x) \cdot \widehat{u} \\ &= \nabla f(x) \cdot \frac{1}{\|\nabla f(x)\|}\nabla f(x) \\ &= \frac{1}{\|\nabla f(x)\|}\nabla f(x) \cdot \nabla f(x) \\ &= \frac{1}{\|\nabla f(x)\|}\|\nabla f(x)\|^2 \\ &= \|\nabla f(x)\|. \end{aligned}$$

Thus, $D_{\widehat{u}}f(x)$ attains its largest value when $\widehat{u}$ is in the direction of $\nabla f(x)$.

□

4. The scalar field $f\colon U \to \mathbb{R}$ is said to have a *local minimum* at $x \in U$ if there exists $r > 0$ such that $B_r(x) \subseteq U$ and

$$f(x) \leqslant f(y) \quad \text{for every} \quad y \in B_r(x).$$

Prove that if f is differentiable at $x \in U$ and f has a local minimum at x, then $\nabla f(x) = \mathbf{0}$, the zero vector in $\mathbb{R}^n$.

Proof. Let $\widehat{u}$ be a unit vector and $t \in \mathbb{R}$ be such that $|t| < r$; then,

$$f(x + t\widehat{u}) \geqslant f(x),$$

from which we get that

$$f(x + t\widehat{u}) - f(x) \geqslant 0.$$

Dividing by $t > 0$ we then have that

$$\frac{f(x + t\widehat{u}) - f(x)}{t} \geqslant 0.$$

Thus, letting $t \to 0^+$, we get that

$$D_{\widehat{u}} f(x) \geqslant 0, \tag{1}$$

since f is differentiable at x. Similarly, dividing by $t < 0$, we have

$$\frac{f(x + t\widehat{u}) - f(x)}{t} \leqslant 0,$$

from which we obtain, letting $t \to 0^-$, that

$$D_{\widehat{u}} f(x) \leqslant 0. \tag{2}$$

Combining (1) and (2) we then have that

$$D_{\widehat{u}} f(x) = 0,$$

where $\widehat{u}$ is an arbitrary unit vector. It then follows from the previous problem that $\nabla f(x) = \mathbf{0}$. $\square$

5. Let I denote an open interval in $\mathbb{R}$. Suppose that $\sigma\colon I \to \mathbb{R}^n$ and $\gamma\colon I \to \mathbb{R}^n$ are paths in $\mathbb{R}^n$. Define a real valued function $f\colon I \to \mathbb{R}$ of a single variable by

$$f(t) = \sigma(t) \cdot \gamma(t) \quad \text{for all } \ t \in I;$$

that is, $f(t)$ is the dot product of the two paths at t.

Show that if σ and γ are both differentiable on I, then so is f, and

$$f'(t) = \sigma'(t) \cdot \gamma(t) + \sigma(t) \cdot \gamma'(t) \quad \text{for all } \ t \in I.$$

***Solution**:* Let $t \in I$ and assume that both σ and γ are differentiable at t. Then,

$$\sigma(t + h) = \sigma(t) + h\sigma'(t) + E_1(h), \quad \text{for } \ |h| \text{ sufficiently small,}$$

where

$$\lim_{h \to 0} \frac{\|E_1(h)\|}{|h|} = 0. \tag{3}$$

Similarly,

$$\gamma(t + h) = \gamma(t) + h\gamma'(t) + E_2(h), \quad \text{for } \ |h| \text{ sufficiently small,}$$

where
$$\lim_{h\to 0}\frac{\|E_2(h)\|}{|h|}=0. \tag{4}$$
It then follows that, for $|h|$ sufficiently small,

$$\begin{aligned}
f(t+h) &= \sigma(t+h)\cdot\gamma(t+h)\\
&= (\sigma(t)+h\sigma'(t)+E_1(h))\cdot(\gamma(t)+h\gamma'(t)+E_2(h))\\
&= \sigma(t)\cdot\gamma(t)+h\sigma(t)\cdot\gamma'(t)+\sigma(t)\cdot E_2(h))+h\sigma'(t)\cdot\gamma(t)\\
&\quad +h^2\sigma'(t)\cdot\gamma'(t)+h\sigma'(t)\cdot E_2(h)+E_1(h)\cdot\gamma(t)\\
&\quad +hE_1(h)\cdot\gamma'(t)+E_1(h)\cdot E_2(h)\\
&= f(t)+h[\sigma(t)\cdot\gamma'(t)+\sigma'(t)\cdot\gamma(t)]+h^2\sigma'(t)\cdot\gamma'(t)\\
&\quad +\sigma(t)\cdot E_2(h))+h\sigma'(t)\cdot E_2(h)+E_1(h)\cdot\gamma(t)\\
&\quad +hE_1(h)\cdot\gamma'(t)+E_1(h)\cdot E_2(h)
\end{aligned}$$

Rearranging terms and dividing by $h\neq 0$ and $|h|$ small enough, we then have that

$$\begin{aligned}
\frac{f(t+h)-f(t)}{h} &= \sigma(t)\cdot\gamma'(t)+\sigma'(t)\cdot\gamma(t)++h\sigma'(t)\cdot\gamma'(t)\\
&\quad +\sigma(t)\cdot\frac{E_2(h)}{h}+\sigma'(t)\cdot E_2(h)+\frac{E_1(h)}{h}\cdot\gamma(t)\\
&\quad +E_1(h)\cdot\gamma'(t)+E_1(h)\cdot\frac{E_2(h)}{h}
\end{aligned}$$

Observe that, as $h\to 0$, all the terms on the right hand side of the previous expression which involve E_1 or E_2 go to 0, by virtue of the Cauchy–Schwarz inequality and (3) and (4). Therefore, we obtain that
$$\lim_{h\to 0}\frac{f(t+h)-f(t)}{h}=\sigma(t)\cdot\gamma'(t)+\sigma'(t)\cdot\gamma(t).$$
Hence, f is differentiable at t, and its derivative at t is

$$f'(t)=\sigma(t)\cdot\gamma'(t)+\sigma'(t)\cdot\gamma(t).$$

Since t is an arbitrary element of I, the result follows. □

6. Let $\sigma\colon I\to\mathbb{R}^n$ denote a differentiable path in $\mathbb{R}^n$. Show that if $\|\sigma(t)\|$ is constant for all $t\in I$, then $\sigma'(t)$ is orthogonal to $\sigma(t)$ for all $t\in I$.

Solution: Let $\|\sigma(t)\| = c$, where c denotes a constant. Then,

$$\|\sigma(t)\|^2 = c^2,$$

or

$$\sigma(t) \cdot \sigma(t) = c^2.$$

Differentiating with respect to t on both sides, and using the result of the previous problem, we obtain that

$$\sigma(t) \cdot \sigma'(t) + \sigma'(t) \cdot \sigma(t) = 0,$$

or, by the symmetry of the dot–product,

$$2\sigma'(t) \cdot \sigma(t) = 0,$$

or

$$\sigma'(t) \cdot \sigma(t) = 0.$$

Hence, $\sigma'(t)$ is orthogonal to $\sigma(t)$ for all $t \in I$. □

7. A particle is following a path in three–dimensional space given by

$$\sigma(t) = (e^t, e^{-t}, 1 - t) \quad \text{for} \quad t \in \mathbb{R}.$$

At time $t_o = 1$, the particle flies off on a tangent.

(a) Where will the particle be at time $t_1 = 2$?

Solution: Find the tangent line to the path at $\sigma(1)$:

$$\overrightarrow{r}(t) = \sigma(1) + (t - 1)\sigma'(1),$$

where

$$\sigma'(t) = (e^t, -e^{-t}, -1) \quad \text{for} \quad t \in \mathbb{R}.$$

Then,

$$\overrightarrow{r}(t) = (e, 1/e, 0) + (t - 1)(e, -1/e, -1).$$

The parametric equations of the tangent line then are

$$\begin{cases} x = e + e(t - 1) \\ y = 1/e - (t - 1)/e \\ z = 1 - t \end{cases}$$

When $t = 2$, the particle will be at the point in $\mathbb{R}^3$ with coordinates

$$(2e, 0, -1).$$

□

(b) Will the particle ever hit the xy–plane? Is so, find the location on the xy plane where the particle hits.

Answer: The particle leaves the path at the point with coordinates $(e, 1/e, 0)$ on the xy–plane. After that, it doesn't come back to it. □

8. Let U denote an open and convex subset of $\mathbb{R}^n$. Suppose that $f\colon U \to \mathbb{R}$ is differentiable at every $x \in U$. Fix x and y in U, and define $g\colon [0,1] \to \mathbb{R}$ by

$$g(t) = f(x + t(y - x)) \quad \text{for} \quad 0 \leqslant t \leqslant 1.$$

(a) Explain why the function g is well defined.

Solution: Since U is convex, $x + t(y - x)$ is in U for $0 \leqslant t \leqslant 1$. Thus, $f(x + t(y - x))$ is defined for $t \in [0, 1]$. □

(b) Show that g is differentiable on $(0, 1)$ and that

$$g'(t) = \nabla f(x + t(y - x)) \cdot (y - x) \quad \text{for} \quad 0 < t < 1.$$

Solution: Apply the Chain Rule to the maps $f\colon U \to \mathbb{R}$ and $\sigma\colon (0, 1) \to \mathbb{R}^n$ given by

$$\sigma(t) = x + t(y - x) \quad \text{for} \quad t \in (0, 1).$$

Since U is convex, it follows that $\sigma(t) \in U$ for all $t \in (0, 1)$. Consequently, $\sigma((0, 1)) \subseteq U$ and therefore $f \circ \sigma\colon (0, 1) \to \mathbb{R}$ is defined. Furthermore, by the Chain Rule, $f \circ \sigma$ is differentiable with

$$D(f \circ \sigma)(t) = Df(\sigma(t))D\sigma(t) = \nabla f(\sigma(t)) \cdot \sigma'(t) \quad \text{for all } t \in (0, 1).$$

Note that $g = f \circ \sigma$ and $\sigma'(t) = y - x$ for all $t \in (0, 1)$. Hence,

$$g'(t) = \nabla f(x + t(y - x)) \cdot (y - x) \quad \text{for} \quad 0 < t < 1,$$

which was to be shown. □

(c) Use the Mean Value Theorem for derivatives to show that there exists a point z is the line segment connecting x to y such that

$$f(y) - f(x) = D_{\widehat{u}} f(z) \|y - x\|,$$

where $\widehat{u}$ is the unit vector in the direction of the vector $y - x$; that is, $\widehat{u} = \dfrac{1}{\|y - x\|}(y - x)$.

(*Hint:* Observe that $g(1) - g(0) = f(y) - f(x)$.)

***Solution**:* Assume that $x \neq y$, for if $x = y$ the equality certainly holds true.
By the Mean Value Theorem, there exists $\tau \in (0,1)$ such that

$$g(1) - g(0) = g'(\tau)(1-0) = g'(\tau).$$

It then follows that

$$f(y) - f(x) = \nabla f(x + \tau(y-x)) \cdot (y-x).$$

Put $z = x + \tau(y-x)$; then, z is a point in the line segment connecting x to y, and

$$\begin{aligned} f(y) - f(x) &= \nabla f(z) \cdot (y-x) \\ &= \nabla f(z) \cdot \frac{y-x}{\|y-x\|}\|y-x\| \\ &= \nabla f(z) \cdot \widehat{u}\; \|y-x\| \\ &= D_{\widehat{u}} f(z)\|y-x\|, \end{aligned}$$

where $\widehat{u} = \dfrac{1}{\|y-x\|}(y-x)$. □

9. Prove that if U is an open and convex subset of $\mathbb{R}^n$, and $f\colon U \to \mathbb{R}$ is differentiable on U with $\nabla f(v) = \mathbf{0}$ for all $v \in U$, then f must be a constant function.

***Solution**:* Fix $x_o \in U$; then, since U is convex, for any $x \in U\backslash\{x_o\}$, the line segment connecting x_o to x is entirely contained in U. Furthermore, by the argument in part (c) of the previous problem, there exists z in the line segment connecting x_o to x such that

$$f(x) - f(x_o) = D_{\widehat{u}} f(z)\|x - x_o\|,$$

where $\widehat{u} = \dfrac{1}{\|x-x_o\|}(x - x_o)$.
Now, $D_{\widehat{u}} f(z) = \nabla f(z) \cdot \widehat{u} = 0$, since $\nabla f(x) = \mathbf{0}$ for all $x \in U$. Therefore,

$$f(x) = f(x_o).$$

Since x was arbitrary, it follows that f maps every element in U to $f(x_o)$; that is, f is a constant function. □

10. Let f be a scalar field defined on (x, y) where $x = r\cos\theta$, $y = r\sin\theta$. Show that
$$\nabla f = \frac{\partial f}{\partial r}\overrightarrow{\mathbf{u}_r} + \frac{1}{r}\frac{\partial f}{\partial \theta}\overrightarrow{\mathbf{u}_\theta},$$
where $\overrightarrow{\mathbf{u}_r} = (\cos\theta, \sin\theta)$ and $\overrightarrow{\mathbf{u}_\theta} = (-\sin\theta, \cos\theta)$.

Hint: First find $\partial f/\partial r$ and $\partial f/\partial\theta$ in terms of $\partial f/\partial x$ and $\partial f/\partial y$ and then solve for $\partial f/\partial x$ and $\partial f/\partial y$ int terms of $\partial f/\partial r$ and $\partial f/\partial\theta$.

***Solution**:* Given $f(x, y)$ where $x = r\cos\theta$ and $y = r\sin\theta$, the Chain Rule implies that
$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r}$$
and
$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta},$$
where
$$\begin{aligned}\frac{\partial x}{\partial r} &= \cos\theta,\\ \frac{\partial y}{\partial r} &= \sin\theta,\\ \frac{\partial x}{\partial \theta} &= -r\sin\theta,\\ \frac{\partial y}{\partial \theta} &= r\cos\theta.\end{aligned}$$
It then follows that
$$\begin{pmatrix}\dfrac{\partial f}{\partial r}\\[2ex] \dfrac{\partial f}{\partial \theta}\end{pmatrix} = \begin{pmatrix}\cos\theta\dfrac{\partial f}{\partial x} + \sin\theta\dfrac{\partial f}{\partial y}\\[2ex] -r\sin\theta\dfrac{\partial f}{\partial x} + r\cos\theta\dfrac{\partial f}{\partial y}\end{pmatrix}$$
or
$$\begin{pmatrix}\dfrac{\partial f}{\partial r}\\[2ex] \dfrac{\partial f}{\partial \theta}\end{pmatrix} = \begin{pmatrix}\cos\theta & \sin\theta\\[2ex] -r\sin\theta & r\cos\theta\end{pmatrix}\begin{pmatrix}\dfrac{\partial f}{\partial x}\\[2ex] \dfrac{\partial f}{\partial y}\end{pmatrix}.$$

Observe that the 2×2 matrix $\begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix}$ is invertible with inverse

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix}^{-1} = \frac{1}{r}\begin{pmatrix} r\cos\theta & -\sin\theta \\ r\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\dfrac{\sin\theta}{r} \\ \sin\theta & \dfrac{\cos\theta}{r} \end{pmatrix}.$$

We then have that

$$\begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\dfrac{\sin\theta}{r} \\ \sin\theta & \dfrac{\cos\theta}{r} \end{pmatrix} \begin{pmatrix} \dfrac{\partial f}{\partial r} \\ \dfrac{\partial f}{\partial \theta} \end{pmatrix}.$$

Which can be written as

$$\begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{pmatrix} = \frac{\partial f}{\partial r}\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} + \frac{1}{r}\frac{\partial f}{\partial \theta}\begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

Transposing the matrices on both sides yields the result. □

11. Let U be an open subset of $\mathbb{R}^n$ and I be an open interval. Suppose that $f\colon U \to \mathbb{R}$ is a differentiable scalar field and $\sigma\colon I \to \mathbb{R}^n$ be a differentiable path whose image lies in U. Suppose also that $\sigma'(t)$ is never the zero vector. Show that if f has a local maximum or a local minimum at some point on the path, then ∇f is perpendicular to the path at that point.

    Suggestion: Consider the real valued function of a single variable $g(t) = f(\sigma(t))$ for all $t \in I$.

    ***Solution**:* If f has a local maximum or minimum at $\sigma(t_o)$, then $g'(t_o) = 0$, where, by the Chain rule,

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

It then follows that

$$\nabla f(\sigma(t_o)) \cdot \sigma'(t_o) = 0,$$

and, consequently, $\nabla f(\sigma(t_o)$ is perpendicular to the tangent to the path at $\sigma(t_o)$. □

12. Let $\sigma\colon [a,b] \to \mathbb{R}^n$ be a differentiable, one–to–one path. Suppose also that $\sigma'(t)$, is never the zero vector. Let $h\colon [c,d] \to [a,b]$ be a one–to–one and onto map such that $h'(t) \neq 0$ for all $t \in [c,d]$. Define

$$\gamma(t) = \sigma(h(t)) \quad \text{for all} \quad t \in [c,d].$$

$\gamma\colon [c,d] \to \mathbb{R}^n$ is a called a *reparametrization* of σ

(a) Show that γ is a differentiable, one–to–one path.

Solution: Since $\gamma = \sigma \circ h$ is the composition of two differentiable maps, it follows from the Chain Rule that γ is differentiable.
To show that γ is one–to–one, suppose that $\gamma(t_1) = \gamma(t_1)$ for t_1 and t_2 in I. It then follows that

$$\sigma(h(t_1)) = \sigma(h(t_2)).$$

Thus, since σ is one–to–one,

$$h(t_1) = h(t_2),$$

from which we get that $t_1 = t_2$ since h is one–to–one. Consequently, γ is one–to–one. □

(b) Compute $\gamma'(t)$ and show that it is never the zero vector.

Solution: By the Chain Rule,

$$\gamma'(t) = h'(t)\sigma'(h(t)) \quad \text{for all} \quad t \in T.$$

Thus, since neither $h'(t)$ nor $\sigma'(t)$ are zero, $\gamma'(t)$ is never the zero vector. □

(c) Show that σ and γ have the same image in $\mathbb{R}^n$.

Solution: We show that

$$\sigma([a,b]) = \gamma([c,d]). \tag{5}$$

Let $x \in \gamma([c,d])$; then, there exists $t \in [c,d]$ such that

$$x = \gamma(t) = \sigma(h(t)).$$

Thus, there exists $h(t) \in [a,b]$ such that $x = \sigma(h(t))$; that is, $x \in \sigma([a,b])$. Hence,

$$\gamma([c,d]) \subseteq \sigma([a,b]). \tag{6}$$

To show the reverse inclusion, let $x \in \sigma([a,b])$. Then, there exists $\tau \in [a,b]$ such that

$$x = \sigma(\tau).$$

Since $h\colon [c,d] \to [a,b]$ is onto, there exists $t \in [c,d]$ such that $\tau = h(t)$. Thus,

$$x = \sigma(h(t)) = \gamma(t),$$

which shows that $x \in \gamma([c,d])$. It then, follows that

$$\sigma([a,b]) \subseteq \gamma([c,d]). \tag{7}$$

Combining the inclusions in (6) and (7) we obtain the set equality in (5). □