Fourth Assignment Answers

▷ **Exercise** 1. Let X be a metric space. Recall that a subset of X is called *open* if whenever it contains a point x it also contains all points "sufficiently close" to x, and that a subset F of X is called closed if whenever a sequence in F converges to a point p of X it follows that $p \in F$. Show that a set is open if and only if its complement is closed.

Answer Suppose $U \subset X$ is open. Let $p_1, p_2, \ldots$ be a sequence in U^c, the complement of U, converging to $p \in X$. We must show that p is also in U^c, i.e., that $p \notin U$. By the definition of convergence, there are element p_n of the sequence arbitrarily close to p. But if p were in U then all points sufficiently close to p would have to be in U and **not** in U^c, so p cannot be in U, proving that U^c is closed.

Now suppose F is closed, and let p be in F^c. Since no sequence of points in F can converge to p, there must be some distance $\epsilon > 0$ such that any point within ϵ of p is **not** in F and so is in F^c, proving that F^c is open.

▷ **Exercise** 2. Let X and Y be metric spaces, $f : X \to Y$ continuous, and A an open (resp. closed) subset of Y. Show that $f^{-1}(A)$ is open (resp. closed) in X. Deduce from this that the unit sphere, $S(V)$, in an inner-product space V is a closed and hence compact subset of V.

Answer Suppose that $A \subset Y$ is closed, and let the sequence $x_1, x_2, \ldots$ in $f^{-1}(A)$ converge to some $x \in X$. Since f is continuous, the sequence $f(x_1), f(x_2), \ldots$ in A converges to $f(x)$, and since A is closed, $f(x) \in A$, i.e., $x \in f^{-1}(A)$, so $f^{-1}(A)$ is also closed.

Now suppose that A is open. Since a subset of a metric space is closed if and only if its complement is open, by what we have just proved it will suffice to show that $f^{-1}(A^c) = (f^{-1}(A))^c$. But $x \in f^{-1}(A^c)$ just means that $f(x) \in A^c$, i.e., that $f(x) \notin A$, while $x \in (f^{-1}(A))^c$ means that it is **not** the case that $f(x) \in A$, which is the same thing.

Now $S(V) = \{v \in V \mid ||v|| = 1\} = G^{-1}(1)$, where $G(v) = ||v||$. Since the set $\{1\}$ is closed in $\mathbf{R}$ and $G : V \to \mathbf{R}$ is continuous, $S(V)$ is closed in V. Since the norms of elements in $S(V)$ are bounded, $S(V)$ is a closed and bounded subset of V and hence compact.

▷ **Exercise** 3. Recall that a subset K of a metric space X is called *compact* if every sequence in K has a subsequence that converges to some point of K (the Bolzano-Weierstrass property), Show that if $f : X \to \mathbf{R}$ is continuous real-valued function on X then f must be bounded on any compact subset K of X and in fact there is a point p of K where f assumes its maximum value. (Hint # 1: If it were not bounded above on K, then there would be a sequence $k_n \in K$ such that $f(k_n) > n$. Hint #2: Choose a sequence $k_n \in K$ so that $f(k_n)$ converges to the least upper bound of the values of f on K.)

Answer If f were not bounded on K there would be a sequence $x_1, x_2, \ldots$ in K with $f(x_n) > n$, and since K is compact, this sequence would have a subsequence $x_{i_1}, x_{i_2}, \ldots$ converging to some $k \in K$. But since $f(x_{i_n})$ is unbounded, it could not converge to the real number $f(k)$, so f would not be continuous at k, a contradiction. Hence f must be bounded on K. Then if B is the least upper bound of the values of f on K, we can choose a sequence $k_1, k_2, \ldots$ in K such that $f(k_n) \geq B - \frac{1}{n}$. Since K is compact, there is a subsequence $k_{i_1}, k_{i_2}, \ldots$ converging to some $k \in K$. Then $f(k) = \lim_{n\to\infty} f(k_{i_n}) = B$, so f assumes its maximum value on K at the point k.

▷ **Exercise** 4. Prove the so-called Chain Rule.

Chain Rule. *Let U, V, W be inner-product spaces, Ω an open set of U and O and open set of V. Suppose that $G : \Omega \to V$ is differentiable at $\omega \in \Omega$ and that $F : O \to W$ is differentiable at $p = G(\omega) \in O$. Then $F \circ G$ is differentiable at ω and $D(F \circ G)_\omega = DF_p \circ DG_\omega$.*

Answer For v and u sufficiently small, we have by definition of DG_ω and DF_p that :

$$G(\omega + v) = G(\omega) + DG_\omega(v) + ||v||\rho^G_\omega(v),$$

$$F(p + u) = F(p) + DF_p(u) + ||u||\rho^F_p(u).$$

On the other hand,

$$F \circ G(\omega + v) = F(G(\omega + v))$$
$$= F(G(\omega) + DG_\omega(v) + ||v||\rho^G_\omega(v)) = F(p + u),$$

where $p = G(\omega)$, and $u = DG_\omega(v) + ||v||\rho^G_\omega(v)$, so

$$F \circ G(\omega + v) = F(G(\omega)) + DF_p\left(DG_\omega(v) + ||v||\rho^G_\omega(v)\right) + ||u||\rho^F_p(u)$$

$$= F(G(\omega)) + DF_p(DG_\omega(v)) + ||v||DF_p(\rho^G_\omega(v)) + ||u||\rho^F_p(u),$$

since DF_p is a linear operator. Furthermore,

$$F(p + u) = F(G(\omega)) + DF_p(DG_\omega(v)) + ||v||\left[DF_p(\rho^G_\omega(v)) + \frac{||u||}{||v||}\rho^F_p(u)\right]$$

$$= F(G(\omega)) + DF_p \circ DG_\omega(v) + \rho(v),$$

where

$$\rho(v) = DF_p(\rho^G_\omega(v)) + \frac{||u||}{||v||}\rho^F_p(u).$$

Now we check that $\rho(v)$ goes to 0 as $||v|| \to 0$, which is easy: as $||v|| \to 0$, $\rho^G_\omega(v) \to 0$ and $\rho^F_p(u) \to 0$ by definition, and $\frac{||u||}{||v||} \to DG_\omega(\frac{v}{||v||})$, which is a constant; so $DF_p(\rho^G_\omega(v)) \to 0$ and $\frac{||u||}{||v||}\rho^F_p(u) \to 0$. This proves that $D(F \circ G)_\omega$ exists and equals $DF_p \circ DG_\omega$.

▷ **Exercise** 5. Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be a function from $\mathbf{R}^n$ to $\mathbf{R}^m$. We use the usual conventions, so that if $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$ then its image point $f(x)$ has the m components $(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$. Show that if f is differentiable at a point p of $\mathbf{R}^n$ then the matrix of the differential, Df_p, with respect to the standard bases of $\mathbf{R}^n$ and $\mathbf{R}^m$, is just the so-called "Jacobian matrix" at p, namely the matrix of partial derivatives $\frac{\partial f_i(p)}{\partial x_j}$.

Answer Let $e_1, e_2, \ldots, e_n$ be the standard basis of $\mathbf{R}^n$, and $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ the standard basis of $\mathbf{R}^m$. What we must show is that $Df_p(e_j) = \sum_{i=1}^m \frac{\partial f_i(p)}{\partial x_j}\epsilon_i = \left(\frac{\partial f_1(p)}{\partial x_j}, \ldots, \frac{\partial f_m(p)}{\partial x_j}\right)$. Now recall that, a consequence of the chain-rule was that $Df_p(e_j)$ is the directional derivative of f at p in direction e_j, i.e., $\left(\frac{d}{dt}\right)_{t=0} f(p + te_j) = \left(\frac{d}{dt}\right)_{t=0} (f_1(p + te_j), \ldots, f_m(p + te_j)) = \left(\left(\frac{d}{dt}\right)_{t=0} f_1(p + te_j), \ldots, \left(\frac{d}{dt}\right)_{t=0} f_m(p + te_j)\right) = \left(\frac{\partial f_1(p)}{\partial x_j}, \ldots, \frac{\partial f_m(p)}{\partial x_j}\right)$.

▷ **Exercise** 6. Let $\sigma : (a, b) \to V$ and $\gamma : (a, b) \to V$ be differentiable curves in an inner-product space V. Show that $\frac{d}{dt}\langle\sigma(t), \gamma(t)\rangle = \langle\sigma'(t), \gamma(t)\rangle + \langle\sigma(t), \gamma'(t)\rangle$, and in particular $\frac{d}{dt}\|\sigma(t)\|^2 = 2\langle\sigma(t), \sigma'(t)\rangle$. Deduce that if $\sigma : (a, b) \to V$ has its image in $S(V)$, i.e., if $\|\sigma(t)\|$ is identically one, then $\sigma(t)$ and $\sigma'(t)$ are orthogonal for all t.

Answer First, we note that the inner product is a continuous function, and that $\langle\sigma(t), \gamma(t)\rangle$ is a function from $\mathbf{R}$ to $\mathbf{R}$, so the standard definition of derivative applies. Thus:

$$\frac{d}{dt}\langle\sigma(t), \gamma(t)\rangle = \lim_{h\to 0} \frac{\langle\sigma(t+h), \gamma(t+h)\rangle - \langle\sigma(t), \gamma(t)\rangle}{h}$$

$$\begin{aligned}
&= \lim_{h\to 0} \frac{\langle\sigma(t+h), \gamma(t+h)\rangle - \langle\sigma(t), \gamma(t+h)\rangle + \langle\sigma(t), \gamma(t+h)\rangle - \langle\sigma(t), \gamma(t)\rangle}{h} \\
&= \lim_{h\to 0} \left(\left\langle \frac{\sigma(t+h) - \sigma(t)}{h}, \gamma(t+h)\right\rangle + \left\langle \sigma(t), \frac{\gamma(t+h) - \gamma(t)}{h}\right\rangle\right) \\
&= \langle\sigma'(t), \gamma(t)\rangle + \langle\sigma(t), \gamma'(t)\rangle .
\end{aligned}$$

In particular, $\frac{d}{dt}\|\sigma(t)\|^2 = 2\langle\sigma(t), \sigma'(t)\rangle$ follows from the symmetry of the inner product. If the image of σ lies on $S(V)$, then $\|\sigma(t)\|^2$ is constant, so $\langle\sigma(t), \sigma'(t)\rangle = 0$, which means that $\sigma(t)$ and $\sigma'(t)$ are orthogonal.