PMATH 351 Real Analysis, Solutions to the Exercises for Chapter 2

1: Determine which of the following functions $d:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ are metrics on $\mathbb{R}$.

(a) $d(x,y)=(x-y)^2$

Solution: This is not a metric on $\mathbb{R}$ since it does not satisfy the triangle inequality. For example, if $x=0$, $y=1$ and $z=2$ then $d(x,y)+d(y,z)=1+1=2<4=d(x,z)$.

(b) $d(x,y)=\sqrt{|x-y|}$

Solution: This is a metric on $\mathbb{R}$. It is clearly positive definite and symmetric, and for $x,y,z\in\mathbb{R}$ we have

$$\begin{aligned} d(x,z) &= \sqrt{|x-z|} \\ &\leq \sqrt{|x-y|+|y-z|}\ , \text{ by the triangle inequality in } \mathbb{R} \\ &\leq \sqrt{|x-y|}+\sqrt{|y-z|}\ , \text{ by the triangle inequality in } \mathbb{R}^2 \\ &= d(x,y)+d(y,z)\,. \end{aligned}$$

(c) $d(x,y)=|x^2-y^2|$

Solution: This is not a metric on $\mathbb{R}$ since it is not positive definite. For example, if $x=1$ and $y=-1$ then we have $d(x,y)=0$ but $x\neq y$.

(d) $d(x,y)=\dfrac{|x-y|}{1+|x-y|}$

Solution: This is a metric. More generally, if d_1 is any metric on a set X, and if $F:[0,\infty)\to[0,\infty)$ is any function which satisfies

(1) $F(x)\geq 0$ for all $x\geq 0$, with $F(x)=0 \iff x=0$,
(2) $F(x)\leq F(y)$ for all $x\leq y\in\mathbb{R}$, and
(3) $F(x+y)\leq F(x)+F(y)$ for all $x,y\geq 0$,

then the map $d_2(x,y)=F\big(d_1(x,y)\big)$ is also a metric. Indeed property (1) ensures that d_2 is positive-definite, we need no requirement on F to ensure that d_2 is symmetric, and properties (2) and (3) ensure that d_2 satisfies the triangle inequality, since for $x,y,z\in\mathbb{R}$ we have

$$\begin{aligned} d_2(x,z) &= F\big(d_1(x,z)\big)\leq F\big(d_1(x,y)+d_1(y,z)\big)\ , \text{ using propery (2)} \\ &\leq F\big(d_1(x,y)\big)+F\big(d_1(y,z)\big)\ , \text{ using property (3)} \\ &= d_2(x,y)+d_2(y,z)\,. \end{aligned}$$

We note that if $F(0)=0$, and $F'(t)>0$ for all $t\geq 0$, then F satisfies properties (1) and (2), and if in addition $F''(t)<0$ for all $t\geq 0$, then F also satisfies property (3). Indeed, say $0\leq x\leq y\leq x+y$. Using the Mean Value Theorem, choose a with $0\leq a\leq x$ so that $F(x)-F(0)=F'(a)(x-0)$, that is $F(x)=F'(a)\,x$, and choose b with $y\leq b\leq x+y$ so that $F(x+y)-F(y)=F'(b)(x+y-y)$, that is $F(x+y)-F(y)=F'(b)\,x$. Since $F''(t)<0$ for all $t\geq 0$, $F'(t)$ is decreasing, so

$$a\leq b\Longrightarrow F'(b)\geq F'(a)\Longrightarrow F'(b)\,x\geq F'(a)\,x\Longrightarrow F(x+y)-F(y)\geq F(x)\,.$$

When d_1 is the standard metric on $\mathbb{R}$ and $F(t)=\dfrac{t}{1+t}$, we have $F(0)=0$, $F'(t)=\dfrac{1}{(1+t)^2}>0$ for all $t\geq 0$ and $F''(t)=\dfrac{-2}{(1+t)^3}<0$ for all $t\geq 0$, and we have $d=d_2$.

2: (a) Let $S = \{(x,y) \in \mathbb{R}^2 \big| y > x^2\}$. Prove, from the definition of an open set, that S is open in $\mathbb{R}^2$.

Solution: Let $(a,b) \in S$ so we have $b > a^2$ and hence $\sqrt{b} > |a|$. Let $r = \min\left(\frac{b-a^2}{2}, \frac{\sqrt{b}-|a|}{2}\right)$. We claim that $B\big((a,b),r\big) \subseteq S$. Let $(x,y) \in B\big((a,b),r\big)$. Note that

$$|x-a| \leq \sqrt{(x-a)^2+(y-b)^2} = d\big((a,b),(x,y)\big) < r \leq \tfrac{\sqrt{b}-|a|}{2}$$

and similarly

$$|y-b| < r \leq \tfrac{b-a^2}{2}.$$

It follows that $|x|-|a| \leq |x-a| < \frac{\sqrt{b}-|a|}{2}$ so that $|x| \leq \frac{\sqrt{b}+|a|}{2}$ and that $b-y \leq |y-b| < \frac{b-a^2}{2}$ so that $y > \frac{b+a^2}{2}$. Note that $0 \leq \big(\sqrt{b}-|a|\big)^2 = b+a^2-2|a|\sqrt{b}$ so we have $2|a|\sqrt{b} \leq b+a^2$. It follows that

$$x^2 < \left(\tfrac{\sqrt{b}+|a|}{2}\right)^2 = \tfrac{b+a^2+2|a|\sqrt{b}}{4} \leq \tfrac{b+a^2}{2} < y.$$

Since $y > x^2$ we have $(x,y) \in S$. This shows that $B\big((a,b),r\big) \subseteq S$, as claimed, and so S is open.

(b) Define $f : \mathbb{R} \to \mathbb{R}^2$ by $f(t) = \left(\frac{2t}{t^2+1}, \frac{t^2-1}{t^2+1}\right)$. Show that Range$(f)$ is not closed in $\mathbb{R}^2$.

Solution: To solve this problem, it helps to draw a picture of Range$(f) \subseteq \mathbb{R}^2$. By plotting points, you will see that Range(f) looks like the unit circle centred at $(0,0)$ with the point $(0,1)$ removed and, if you wish, you can show that this is indeed the case. Let $S = \text{Range}(f)$ and let $a = (0,1)$. Let $x(t) = \frac{2t}{t^2+1}$ and $y(t) = \frac{t^2-1}{t^2+1}$ so that $f(t) = \big(x(t), y(t)\big)$. We claim that $a \in \overline{S}$ but $a \notin S$. It is clear that $a \notin S$ because to get $f(t) = a$ we need $x(t) = 0$ and $y(t) = 1$, but to get $x(t) = \frac{2t}{t^2+1} = 0$ we must choose $t = 0$, but when $t = 0$ we have $y(t) = \frac{t^2-1}{t^2+1} = -1 \neq 1$. To show that $a \in \overline{S}$, we shall show that for all $r > 0$ we have $B(a,r) \cap S \neq \emptyset$. Let $r > 0$. Since $\lim\limits_{t\to\infty} x(t) = 0$ and $\lim\limits_{t\to\infty} y(t) = 1$ we can choose $t \in \mathbb{R}$ so that $\big|x(t)-0\big| < \frac{r}{2}$ and $\big|y(t)-1\big| < \frac{r}{2}$. Then we have

$$\big|f(t)-a\big| = \big|(x(t),y(t))-(0,1)\big| = \big|\big(x(t)\,,\,y(t)-1\big)\big| \leq |x(t)| + |y(t)-1| < \tfrac{r}{2} + \tfrac{r}{2} = r$$

and so $f(t) \in B(a,r) \cap S$. This shows that for all $r > 0$ we have $B(a,r) \cap S \neq \emptyset$, and so $a \in \overline{S}$. Since $a \in \overline{S}$ but $a \notin S$ we see that $S \neq \overline{S}$ and so S is not closed.

3: Determine which of the following statements are true for every metric space (X,d) and every $A \subseteq X$.

(a) $\overline{B(a,r)} = \overline{B}(a,r)$ for every $a \in X$ and every $r > 0$.

Solution: This is false. For example, let $X = \mathbb{Z}$ with the standard metric. Then $B(0,1) = \{0\} = \overline{B(0,1)}$, but $\overline{B}(0,1) = \{-1,0,1\}$.

(b) $\left(\overline{A}\right)^c = \left(A^c\right)^\circ$.

Solution: This is true. Indeed

$$\begin{aligned}\left(\overline{A}\right)^c &= \left(\bigcap\{K \subseteq X \big| K \text{ closed}\,, A \subseteq K\}\right)^c \\ &= \bigcup\{K^c \subseteq X \big| K \text{ closed}\,, A \subseteq K\} \\ &= \bigcup\{U \subseteq X \big| U^c \text{ closed}\,, A \subseteq U^c\} \\ &= \bigcup\{U \subseteq X \big| U \text{ open}\,, U \subseteq A^c\} = \left(A^c\right)^\circ.\end{aligned}$$

(c) If $A = A^\circ$ then $A = \left(\overline{A}\right)^\circ$.

Solution: This is false. For example, let $X = \mathbb{R}^n$ with the standard metric, and let $A = B^*(0,1)$. Then $\overline{A} = \overline{B}(0,1)$ and $\left(\overline{A}\right)^\circ = B(0,1)$.

(d) If $A = \overline{A}$ then $\partial(\partial A) = \partial A$.

Solution: This is true, as we now prove. Note first that, for any set A, we have $\partial A = \overline{A} \setminus A^\circ = \overline{A} \cap (A^\circ)^c$, which is closed, since it is the intersection of two closed sets. Now suppose that A is closed. We claim that $(\partial A)^\circ = \emptyset$. Indeed, if we had $a \in (\partial A)^\circ$, then we could choose $r > 0$ so that $B(a,r) \subseteq \partial A$, but then we would have $B(a,r) \subseteq \partial A \subseteq \overline{A} = A$ so that $a \in A^\circ$, and this is not possible since $a \in A^\circ \Longrightarrow a \notin \partial A \Longrightarrow a \notin (\partial A)^\circ$. Since ∂A is closed with $(\partial A)^\circ = \emptyset$, we have $\partial(\partial A) = \overline{\partial A} \setminus (\partial A)^\circ = \partial A \setminus \emptyset = \partial A$.

4: (a) Let (X,d) be a metric space with X uncountable. Show that for every $a \in X$ there exists $r > 0$ such that $B(a,r)$ is uncountable.

Solution: Suppose, for a contradiction, that there exists $a \in X$ such that for all $r > 0$, $B(a,r)$ is at most countable. Choose such a point a. Since for every $x \in X$ there exists $n \in \mathbb{N}$ with $n > d(a,x)$ so that $x \in B(a,n)$, we have $X = \bigcup_{n=1}^{\infty} B(a,n)$, which is a union of at most countable sets, and hence X is at most countable.

(b) Let (X,d) be a metric space with the property that for every $a \in X$ there exists $r > 0$ such that $B(a,r)$ is countable. Determine whether X must be countable.

Solution: It is not necessarily the case that X is countable. For example, let $X = \mathbb{R} \times \left(\mathbb{Q} \cap (0,1)\right)$ with the metric given by

$$d\big((x,p),(y,q)\big) = d_0(x,y) + d_1(p,q)$$

where d_0 is the discrete metric on $\mathbb{R}$ and d_1 is the standard metric on $\mathbb{Q} \cap (0,1)$. Notice that when $x = y$ we have $d\big((x,p),(y,q)\big) = d_1(p,q) < 1$ and when $x \neq y$ we have $d\big((x,p),(y,q)\big) = 1 + d(p,q) \geq 1$. Thus for $(x,p) \in X$, we have

$$B\big((x,p),1\big) = \{x\} \times \left(\mathbb{Q} \cap (0,1)\right).$$

Note that $\left|\mathbb{Q} \cap (0,1)\right| = \aleph_0$ $\big($proof: since $\mathbb{Q} \cap (0,1) \subseteq \mathbb{Q}$ we have $\left|\mathbb{Q} \cap (0,1)\right| \leq |\mathbb{Q}| = \aleph_0$, and since the map $F : \mathbb{N} \to \mathbb{Q} \cap (0,1)$ given by $F(n) = \frac{1}{2^n}$ is 1:1 we have $\aleph_0 = |\mathbb{N}| \leq \left|\mathbb{Q} \cap (0,1)\right|\big)$ and so

$$\left|B\big((x,p),1\big)\right| = \left|\{x\} \times \left(\mathbb{Q} \cap (0,1)\right)\right| = 1 \cdot \aleph_0 = \aleph_0\,,$$

but $|X| = 2^{\aleph_0}$ $\big($proof: $|X| = \left|\mathbb{R} \times \left(\mathbb{Q} \cap (0,1)\right)\right| = 2^{\aleph_0} \cdot \aleph_0$, and we have $2^{\aleph_0} \cdot \aleph_0 \leq 2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0 + \aleph_0} = 2^{\aleph_0}$ and $2^{\aleph_0} = 2^{\aleph_0} \cdot 1 \leq 2^{\aleph_0} \cdot \aleph_0\big)$.

5: (a) Show that there is no inner product on $\mathbb{R}^2$ which induces the 1-norm $\| \ \|_1$.

Solution: If there were such an inner product, say $\langle \ , \ \rangle$, then by the polarization identity we would have

$$\langle (1,0),(0,1)\rangle = \tfrac{1}{2}\big(\|(1,1)\|_1{}^2 - \|(1,0)\|_1{}^2 - \|(0,1)\|_1{}^2\big) = \tfrac{1}{2}(4-1-1) = 1 \text{ , and}$$
$$\langle -(1,0),(0,1)\rangle = \tfrac{1}{2}\big(\|(-1,1)\|_1{}^2 - \|(-1,0)\|_1{}^2 - \|(0,1)\|_1{}^2\big) = \tfrac{1}{2}(4-1-1) = 1\,,$$

but this is not possible since by linearity, we must have $\langle -(1,0),(0,1)\rangle = -\langle (1,0),(0,1)\rangle$.

(b) Let $T = \big\{U \subseteq \mathbb{R} \,\big|\, U = \emptyset \text{ or } \mathbb{R}\setminus U \text{ is finite}\big\}$. Show that T is a topology on $\mathbb{R}$ which is not induced by any metric on $\mathbb{R}$ (T is called the *cofinite topology* on $\mathbb{R}$).

Solution: First we show that T is a topology on $\mathbb{R}$. Clearly, we have $\emptyset \in T$ and $\mathbb{R} \in T$ (since $\mathbb{R}^c = \emptyset$, which is finite). Suppose that $U_\alpha \in T$ for each $\alpha \in A$. If every $U_\alpha = \emptyset$ then $\bigcup\limits_{\alpha\in A} U_\alpha = \emptyset$, so $\bigcup\limits_{\alpha\in A} U_\alpha \in T$. If $U_\beta \neq \emptyset$ for some $\beta \in A$, then since $U_\beta \subseteq \bigcup\limits_{\alpha\in A} U_\alpha$, we have $\big(\bigcup\limits_{\alpha\in A} U_\alpha\big)^c \subseteq {U_\beta}^c$, which is finite, so $\bigcup\limits_{\alpha\in A} U_\alpha \in T$. Suppose that $U_k \in T$ for each $k = 0,1,\cdots,n$. If some $U_k = \emptyset$ then $\bigcap\limits_{k=0}^{n} U_k = \emptyset$, so $\bigcap\limits_{k=0}^{n} U_k \in T$. If no $U_k = \emptyset$ then each ${U_k}^c$ is finite, so $\big(\bigcap\limits_{k=0}^{n} U_k\big)^c = \bigcup\limits_{k=0}^{n} {U_k}^c$, which is a finite union of finite sets, and hence finite, so $\bigcap\limits_{k=0}^{n} U_k \in T$.

Next we show that T cannot be induced by any metric. Let d be any metric on $\mathbb{R}$. Let $r = \frac{1}{2}\,d(0,1)$. Note that $B(0,r)\cap B(1,r) = \emptyset$ since if we had $x \in B(0,r)\cap B(1,r)$ then we would have $d(0,x) < r$ and $d(1,x) < r$ and so $2r = d(0,1) \leq d(0,x) + d(x,1) < r + r = 2r$, which is not possible. Thus in the topology which is induced by any metric, there exist two disjoint non-empty sets. On the other hand, in the cofinite topology T on $\mathbb{R}$, given any two non-empty sets $U_1, U_2 \in T$, as shown in the previous paragraph we have that $(U_1 \cap U_2)^c$ is finite, and so $U_1 \cap U_2 \neq \emptyset$.

6: Let A denote the set of all real-valued sequences $(a_n)_{n\geq 1}$ for which $|a_n| \leq \frac{1}{2^n}$ for all $n \in \mathbb{Z}^+$.

(a) Show that $A^\circ = \emptyset$ in (ℓ_1, d_1).

Solution: Let $a = (a_n)_{n\geq 1} \in A$. Let $r > 0$. Choose $N \in \mathbb{Z}^+$ so that $\frac{1}{2^N} < \frac{r}{2}$. Define $b = (b_n)_{n\geq 1}$ by

$$b_n = \begin{cases} a_n\,, & n \neq N \\ \frac{r}{2}\,, & n = N\,. \end{cases}$$

Then $b \in \ell_1$ since $\sum\limits_{n=1}^{\infty} |b_n| = \sum\limits_{n=1}^{\infty} |a_n| - a_N + \frac{r}{2} < \infty$, and $b \notin A$ since $b_N = \frac{r}{2} > \frac{1}{2^N}$, and $b \in B_1(a,r)$ since

$$\|b-a\|_1 = \sum_{n=1}^{\infty} |b_n - a_n| = \big|\tfrac{r}{2} - a_N\big| \leq \tfrac{r}{2} + |a_N| \leq \tfrac{r}{2} + \tfrac{1}{2^N} < \tfrac{r}{2} + \tfrac{r}{2} = r\,.$$

(b) Show that $\overline{A} = A$ in (ℓ_1, d_1).

Solution: We must show that A is closed, or equivalently that A^c is open. Let $a = (a_n)_{n\geq 1} \in A^c$. Choose $N \in \mathbb{Z}^+$ so that $|a_N| > \frac{1}{2^N}$. Let $r = |a_N| - \frac{1}{2^N}$. We claim that $B_1(a,r) \subseteq A^c$. Let $x = (x_n)_{n\geq 1} \in B_1(a,r)$. Then

$$|a_N| - |x_N| \leq |a_N - x_N| \leq \sum_{n=1}^{\infty} |a_n - x_n| = \|a - x\|_1 < r = |a_N| - \tfrac{1}{2^N}\,.$$

It follows that $|x_N| > \frac{1}{2^N}$, and so $x = (x_n)_{n\geq 1} \notin A$, that is $x \in A^c$ as required.

7: (a) Show that ℓ_1 is neither open nor closed in the metric space (ℓ_∞, d_∞).

Solution: By Part 2 of Theorem 2.49, to show that ℓ_1 is not closed in (ℓ_∞, d_∞), it suffices to find a limit point of ℓ_1 which does not lie in ℓ_1. Let $a = \left(\frac{1}{n}\right)_{n\geq 1} = \left(1, \frac{1}{2}, \frac{1}{3}, \cdots\right)$. Note that $a \in \ell_\infty$ with $\|a\|_\infty = a_1 = 1$ but $a \notin \ell_1$ since $\|a\|_1 = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$. We claim that a is a limit point of ℓ_1. Let $r > 0$. Choose $m \in \mathbb{Z}^+$ such that $\frac{1}{m} < r$. Let $x = (x_n)_{n\geq 1}$ with $x_n = \frac{1}{n}$ for $n \leq m$ and $x_n = 0$ for $n > m$, that is let $x = \left(1, \frac{1}{2}, \cdots, \frac{1}{m}, 0, 0, \cdots\right)$. Then $x \neq a$, and $x \in \ell_1$ with $\|x\|_1 = \sum_{n=1}^{m} \frac{1}{n} < \infty$, and $\|x - a\|_\infty = \left\|\left(0, \cdots, 0, \frac{1}{m+1}, \frac{1}{m+2}, \cdots\right)\right\|\infty = \frac{1}{m+1} < r$, and so we have $x \in B^*(a, r) \cap \ell_1$. Thus a is a limit point of ℓ_1, as claimed, hence ℓ_1 is not closed in (ℓ_∞, d_∞).

To see that ℓ_1 is not open, and indeed to see that ${\ell_1}^\circ = \emptyset$ in (ℓ_∞, d_∞), note that given $a = (a_n) \in \ell_1$ and $0 < r \in \mathbb{R}$, we can choose $b = (b_n)$ to be the sequence given by $b_n = a_n + \frac{r}{2}$. Then we have $b \in B_\infty(a, r)$ but $b \notin \ell_1$.

(b) Determine whether every set $U \subseteq \ell_1$ which is open in (ℓ_1, d_2) is also open in (ℓ_1, d_1).

Solution: We show that this is indeed the case. Let $a = (a_n)_{n\geq 1}$ and $x = (x_n)_{n\geq 1}$ be in ℓ_2. For all $N \in \mathbb{Z}^+$, the Triangle Inequality gives

$$\sqrt{\sum_{n=1}^{N} (a_n - x_n)^2} \leq \sum_{n=1}^{N} |a_n - x_n| \,.$$

Taking the limit as $N \to \infty$ we obtain

$$d_2(a, x) = \sqrt{\sum_{n=1}^{\infty} (a_n - x_n)^2} \leq \sum_{n=1}^{\infty} |a_n - x_n| = d_1(a, x) \,.$$

It follows that for all $0 < r \in \mathbb{R}$ we have $B_1(a, r) \subseteq B_2(a, r)$ since

$$x \in B_1(a, r) \Longrightarrow d_1(a, x) < r \Longrightarrow d_2(a, x) \leq d_1(a, x) < r \Longrightarrow x \in B_2(a, r) \,.$$

Now suppose that $U \subseteq \ell_1$ is open in (ℓ_1, d_2). Let $a \in U$. Choose $0 < r \in \mathbb{R}$ so that $B_2(a, r) \subseteq U$. Then $B_1(a, r) \subseteq B_2(a, r) \subseteq U$. Thus U is also open in (ℓ_1, d_1).

(c) Determine whether every set $U \subseteq \ell_1$ which is open in (ℓ_1, d_1) is also open in (ℓ_1, d_2).

Solution: We shall show that for $a \in \ell_1$ and $0 < r \in \mathbb{R}$, the ball $B_1(a, r)$, which is open in (l_1, d_1), is not open in (l_1, d_2). Let $a = (a_n)_{n\geq 1} \in \ell_1$ and let $0 < r \in \mathbb{R}$. We claim that $a \notin B_1(a, r)^\circ$ in (ℓ_1, d_2). We must show that for every $0 < s \in \mathbb{R}$, $B_2(a, s) \not\subseteq B_1(a, r)$. Let $s > 0$. Recall that $\sum \frac{1}{n^2}$ converges but $\sum \frac{1}{n}$ diverges. Let

$$p = \sqrt{\sum_{n=1}^{\infty} \frac{1}{n^2}}$$

$\left(\text{in fact, } p = \frac{\pi}{\sqrt{6}}\right)$. Then we have $\sum_{n=1}^{\infty} \frac{s^2/p^2}{n^2} = s^2$ and $\sum_{n=1}^{\infty} \frac{s/p}{n} = \infty$. Choose $N \in \mathbb{Z}^+$ so that $\sum_{n=1}^{N} \frac{s/p}{n} > r$. Let $b = (b_n)_{n\geq 1}$ be the sequence given by $b_n = \frac{s/p}{n}$ for $1 \leq n \leq N$ and $b_n = 0$ for $n > N$. Note that $b \in \mathbb{R}^\infty \subseteq \ell_1$ so $a + b \in \ell_1$. We have $a + b \in B_2(a, s)$ because

$$d_2(a, a + b) = \sqrt{\sum_{n=1}^{N} \frac{s^2/p^2}{n^2}} < \sqrt{\sum_{n=1}^{\infty} \frac{s^2/p^2}{n^2}} = s \,,$$

but $a + b \notin B_1(a, r)$ because

$$d_1(a, a + b) = \sum_{n=1}^{N} \frac{s/p}{n} > r \,.$$

Thus $B_2(a, r) \not\subseteq B_1(a, r)$, as required. We remark that a minor modification of the above argument can be used to show that $B_1(a, r)^\circ = \emptyset$ in (ℓ_1, d_2).

8: (a) Verify that we can define a metric on the space $M_{k\times\ell}(\mathbb{R})$ of real $k\times\ell$ matrices by $d(A,B)=\text{rank}(B-A)$.

Solution: For $A,B\in M_{k\times\ell}(\mathbb{R})$, we define $d(A,B)=\text{rank}(A-B)$. It is clear that d is positive definite, that is $d(A,B)\geq 0$ for all $A,B\in M_{k\times\ell}(\mathbb{R})$ with $d(A,B)=0$ if and only if $A=B$, because only the zero matrix has rank zero. It is also clear that d is symmetric, that is $d(A,B)=d(B,A)$, since for any matrix X we have $\text{rank}(X)=\text{rank}(-X)$. We need to verify that d satisfies the triangle inequality. Let $A,B,C\in M_{k\times\ell}(\mathbb{R})$. Let $X=A-B$, $Y=B-C$ and $Z=C-A$. Note that $X+Y=Z$. Let $u_1,\cdots,u_\ell$ be the columns of X, let $v_1,\cdots,v_\ell$ be the columns of Y and let $w_1,\cdots,w_\ell$ be the columns of Z. Since $X+Y=Z$ we have $w_i=u_i+v_i$ for all indices i, and so $\text{Span}\{w_1,\cdots,w_\ell\}\subseteq\text{Span}\{u_1,\cdots,u_\ell,v_1,\cdots,v_\ell\}$. Let $U=\text{Col}(A)=\text{Span}\{u_1,\cdots,u_\ell\}$, $V=\text{Col}(B)=\text{Span}\{v_1,\cdots,v_\ell\}$ and $W=\text{Col}(Z)=\text{Span}\{w_1,\cdots,w_\ell\}$. Since $W=\text{Span}\{w_1,\cdots,w_\ell\}\subseteq\text{Span}\{u_1,\cdots,u_\ell,v_1,\cdots,v_\ell\}=U+V$ we have

$$\text{rank}(Z)=\dim W\leq\dim(U+V)=\dim U+\dim V-\dim(U\cap V)\leq\dim U+\dim V=\text{rank}(X)+\text{rank}(Y),$$

and so

$$d(A,C)=\text{rank}(A-C)=\text{rank}(Z)\leq\text{rank}(X)+\text{rank}(Y)=\text{rank}(A-B)+\text{rank}(B-C)=d(A,B)+d(B,C).$$

(b) Verify that we can define a metric on the unit sphere $\mathbb{S}^2=\{u\in\mathbb{R}^3\,\big|\,\|u\|=1\}$ by $d(u,v)=\cos^{-1}(u\bullet v)$ where $u\bullet v$ is the standard inner product (the dot product) in $\mathbb{R}^3$. Hint: you may wish to use properties of the cross product in $\mathbb{R}^3$.

Solution: For $u,v\in\mathbb{S}^2$ we define $d(u,v)=\cos^{-1}(u\bullet v)$. It is clear that d is symmetric. Note that d is positive definite because for all $u,v\in\mathbb{S}^2$ we have $d(u,v)=\cos^{-1}(u\bullet v)\in[0,\pi]$ and we have

$$\|u-v\|^2=\|u\|^2+\|v\|^2-2(u\bullet v)=2-2(u\bullet v)$$

so that

$$d(u,v)=0\iff\cos^{-1}(u\bullet v)=0\iff u\bullet v=1\iff\|u-v\|^2=0\iff u=v.$$

To show that d satisfies the Triangle Inequality, recall (or verify) that for vectors $u,v,w,x\in\mathbb{R}^3$ we have

$$(u\times v)\bullet(w\times x)=(u\bullet w)(v\bullet x)-(v\bullet w)(u\bullet x).\qquad(1)$$

For $0\neq u,v\in\mathbb{R}^3$, recall that the angle $\theta=\theta(u,v)$ between u and v is given by $\theta=\cos^{-1}\dfrac{u\bullet v}{\|u\|\,\|v\|}$. Using Property (1) of the cross product, we have

$$\begin{aligned}\sin^2\theta=1-\cos^2\theta=1-\frac{(u\bullet v)^2}{\|u\|^2\|v\|^2}&=\frac{\|u\|^2\|v\|^2-(u\bullet v)^2}{\|u\|^2\|v\|^2}\\ =\frac{(u\bullet u)(v\bullet v)-(u\bullet v)(u\bullet v)}{\|u\|^2\|v\|^2}&=\frac{(u\times v)\bullet(u\times v)}{\|u\|^2\|v\|^2}=\frac{\|u\times v\|^2}{\|u\|^2\|v\|^2}.\end{aligned}$$

Since $0\leq\theta\leq\pi$ so that $\sin\theta\geq 0$ it follows that

$$\sin\theta=\frac{\|u\times v\|}{\|u\|\,\|v\|}.$$

When $u,v\in\mathbb{S}^2$, so $|u|=|v|=1$, note that $d(u,v)=\cos^{-1}(u\bullet v)=\theta(u,v)$. When $u,v,w\in\mathbb{S}^2$, using the Cauchy Schwarz Inequality and Property (1), we have

$$\begin{aligned}\cos\big(\theta(u,v)+\theta(v,w)\big)&=\cos\theta(u,v)\cos\theta(v,w)-\sin\theta(u,v)\sin\theta(v,w)\\&=(u\bullet v)(v\bullet w)-\|u\times v\|\,\|v\times w\|\\&\leq(u\bullet v)(v\bullet w)-(u\times v)\bullet(v\times w)\\&=(u\bullet v)(v\bullet w)-\big((u\bullet v)(v\bullet w)-(u\bullet w)(v\bullet v)\big)\\&=u\bullet w=\cos\theta(u,w).\end{aligned}$$

Since $\cos\theta$ is decreasing with θ, it follows that $\theta(u,w)\leq\theta(u,v)+\theta(v,w)$, that is $d(u,w)\leq d(u,v)+d(v,w)$.