i.e., $[\mathbf{D}\big(\mathbf{g}(\mathbf{f}(\mathbf{0}))][\mathbf{Df}(\mathbf{0})] = I$. But by definition, $[\mathbf{Df}(\mathbf{0})]^{-1}$ does not exist so $\mathbf{g}$ cannot exist.

b. False; example 1.9.4 provides a counterexample.

1.8.13 Call $S(A) = A^2 + A$ and $T(A) = A^{-1}$. We have $F = T \circ S$, so

$$\begin{aligned}[\mathbf{D}F(A)]H &= [\mathbf{D}T(A^2+A)][\mathbf{D}S(A)]H \\ &= [\mathbf{D}T(A^2+A)](AH + HA + H) \\ &= -(A^2+A)^{-1}(AH+HA+H)(A^2+A)^{-1}.\end{aligned}$$

It isn't really possible to simplify this much.

1.9.1 Except at $\begin{pmatrix}0\\0\end{pmatrix}$, the partial derivatives of f are given by

$$D_1 f\begin{pmatrix}x\\y\end{pmatrix} = \frac{2x^5 + 4x^3y^2 - 2xy^4}{(x^2+y^2)^2} \quad \text{and} \quad D_2 f\begin{pmatrix}x\\y\end{pmatrix} = \frac{4x^2y^3 - x^4y + 2y^5}{(x^2+y^2)^2}.$$

At the origin, they are both given by

$$D_1 f\begin{pmatrix}0\\0\end{pmatrix} = D_2 f\begin{pmatrix}0\\0\end{pmatrix} = \lim_{h\to 0}\frac{1}{h}\left(\frac{h^4}{h^2}\right) = 0.$$

Thus there are partial derivatives everywhere, and we need to check that they are continuous. The only problem is at the origin. One easy way to show this is to remember that

$$|x| \le \sqrt{x^2+y^2} \quad \text{and} \quad |y| \le \sqrt{x^2+y^2}.$$

Then both partial derivatives satisfy

$$|D_i f\begin{pmatrix}x\\y\end{pmatrix}| \le 8\frac{(x^2+y^2)^{5/2}}{(x^2+y^2)^2} = \sqrt{x^2+y^2}.$$

Thus the limit of both partials at the origin is 0, so the partials are continuous and f is differentiable everywhere.

1.9.2 a. There is not problem except at the origin; everywhere else the function is differentiable, since it is a quotient of two polynomials, and the denominator does not vanish.

At the origin, to compute the directional derivative in the direction $\begin{pmatrix}x\\y\end{pmatrix}$, we must show that the limit

$$\lim_{h\to 0}\frac{1}{h}\left(\frac{3h^3x^2y - h^3y^3}{h^2(x^2+y^2)}\right) = \frac{3x^2y - y^3}{x^2+y^2}$$

exists, which it evidently does. But the limit, which is in fact the function itself, is not a linear function of $\begin{pmatrix}x\\y\end{pmatrix}$, so the function is not differentiable.

b.

c.

1.9.3 a. It means that there is a line matrix $[a,\ b]$ such that

$$\lim_{\vec{\mathbf{h}}\to\mathbf{0}}\frac{\sin\left(\dfrac{h_1^2h_2^2}{h_1^2-h_2^2}\right) - ah_1 - bh_2}{(h_1^2+h_2^2)^{1/2}} = 0.$$

b. Since f vanishes identically on both axes, both partials exist, and are 0 at the origin. In fact, $D_1 f$ vanishes on the x-axis and $D_2 f$ vanishes on the y-axis.

c. We know that if f is differentiable at the origin, then its partial derivatives exist at the origin and are the numbers a, b of part a. Thus for f to be differentiable at the origin, we must have

$$\lim_{\vec{\mathbf{h}} \to \mathbf{0}} \frac{\sin\left(\dfrac{h_1^2 h_2^2}{h_1^2 - h_2^2}\right)}{(h_1^2 + h_2^2)^{1/2}} = 0,$$

and this is indeed the case, since

$$|\sin(h_1^2 h_2^2)| \le |h_1^2 h_2^2| < |h_1|^2 + |h_2|^2 \quad \text{when } |h_1|,\ |h_2| < 1.$$

Solutions for Review Exercises, Chapter 1

1.1 a. not a subspace: $\vec{\mathbf{0}}$ is not on the line. b. not a subspace: $\vec{\mathbf{0}}$ is not on the line. c. a subspace.

1.2 $AB = \begin{bmatrix} 1+ab & a \\ a & 0 \end{bmatrix} \quad BA = \begin{bmatrix} 1 & a \\ a+b & ab \end{bmatrix}$. So $AB = BA$ if and only if

$$ab = 0 \quad \text{and} \quad a+b = a. \quad \text{So } b = 0 \text{ and } a \text{ can be anything.}$$

1.3 If A and B are upper-triangular matrices, then if $i > j$ we know that $a_{i,j} = b_{i,j} = 0$. Using the definition of matrix multiplication

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j},$$

we see that if $i > j$, then in the summation either $a_{i,k} = 0$ or $b_{k,j} = 0$, so $c_{i,j} = \sum_{k=1}^{n} 0 = 0$. If for all $i > j$ $c_{i,j} = 0$, then C is upper-triangular, so if A, B are upper-triangular, AB is also upper-triangular.

1.4 a. Since $z_1 + z_2 = \alpha_1 + \alpha_2 + (\beta_1 + \beta_2)i$,

$$\underbrace{\begin{bmatrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{bmatrix}}_{T_{z_1}} + \underbrace{\begin{bmatrix} \alpha_2 & \beta_2 \\ -\beta_2 & \alpha_2 \end{bmatrix}}_{T_{z_2}} = \underbrace{\begin{bmatrix} \alpha_1 + \alpha_2 & \beta_1 + \beta_2 \\ -\beta_1 - \beta_2 & \alpha_1 + \alpha_2 \end{bmatrix}}_{T_{z_1+z_2}}$$

Since $z_1 z_2 = \alpha_1\alpha_2 - \beta_1\beta_2 + (\alpha_1\beta_2 + \alpha_2\beta_1)i$, $T_{z_1z_2}$ is the matrix on the left, which is the product of the two matrices on the right:

$$\underbrace{\begin{bmatrix} \alpha_1\alpha_2 - \beta_1\beta_2 & \alpha_1\beta_2 + \alpha_2\beta_1 \\ -\alpha_1\beta_2 - \alpha_2\beta_1 & \alpha_1\alpha_2 - \beta_1\beta_2 \end{bmatrix}}_{T_{z_1z_2}} = \begin{bmatrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{bmatrix} \begin{bmatrix} \alpha_2 & \beta_2 \\ -\beta_2 & \alpha_2 \end{bmatrix}.$$

b. $T_z^{-1} = \dfrac{1}{\alpha^2+\beta^2} \underbrace{\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}}_{T_{\bar{z}}}$. This corresponds to $\dfrac{1}{z} = \dfrac{1}{\alpha + i\beta} = \dfrac{\alpha - i\beta}{\alpha^2+\beta^2} = \dfrac{\bar{z}}{\alpha^2+\beta^2}$.

c. A good choice is T_i. This gives $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

1.5 a. Labeling the vertices in the direction of the arrows: $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

b. Labeling from top left clockwise: $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

c. Labeling starts at center, then bottom left clockwise: $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$

1.6 a. The mapping $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \mapsto \begin{bmatrix} x_2 \\ x_4 \end{bmatrix}$ is linear; denoting the mapping by T we have

$$T(a\vec{\mathbf{x}}) = \begin{bmatrix} ax_2 \\ ax_4 \end{bmatrix} = aT(\vec{\mathbf{x}}) \quad \text{and} \quad T(\vec{\mathbf{x}} + \vec{\mathbf{w}}) = \begin{bmatrix} x_2 + w_2 \\ x_4 + w_2 \end{bmatrix} = T(\vec{\mathbf{x}}) + T(\vec{\mathbf{w}}).$$

Its matrix is $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

b. This mapping is not linear. Denote it by T. Then

$$T(a\vec{\mathbf{x}}) = \begin{bmatrix} a^2 x_2 x_4 \\ a(x_1 + x_3) \end{bmatrix} \neq aT(\vec{\mathbf{x}}) = \begin{bmatrix} ax_2 x_4 \\ a(x_1 + x_3) \end{bmatrix}.$$

1.7 a. Yes there is; its matrix is $[T] = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 2 & 1 & -1 \end{bmatrix}$. We computed this matrix by denoting by $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3, \vec{\mathbf{v}}_4$ the four input vectors and determining what combinations of these vectors give the four standard basis vectors in $\mathbb{R}^4$. For example, $\vec{\mathbf{e}}_4 = \vec{\mathbf{v}}_3 - \vec{\mathbf{v}}_2$, so the fourth column of the matrix is $T\vec{\mathbf{e}}_4 = T\vec{\mathbf{v}}_3 - T\vec{\mathbf{v}}_2 = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$. Similarly, $\vec{\mathbf{e}}_1 = \vec{\mathbf{v}}_4 - \vec{\mathbf{v}}_3 + \vec{\mathbf{v}}_2$ so $T(\vec{\mathbf{e}}_1) = T\vec{\mathbf{v}}_4 - T\vec{\mathbf{v}}_3 + T\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. We then confirmed that this matrix does indeed satisfy the four equations of the exercise.

b. No, it is not linear; we have $[T] \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$, not $\begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}$. Another way to see this is to say that if S were linear, then by part (a) we would have

$$S\vec{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad S\vec{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \quad S\vec{\mathbf{e}}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad S\vec{\mathbf{e}}_4 = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix},$$

which by linearity should give $S \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = S(\vec{\mathbf{e}}_1 + \vec{\mathbf{e}}_2 + \vec{\mathbf{e}}_3 + \vec{\mathbf{e}}_4) = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

1.8

$$T = \begin{bmatrix} \cos(30°) & 0 & \sin(30°) \\ 0 & 1 & 0 \\ -\sin(30°) & 0 & \cos(30°) \end{bmatrix}$$

1.9 a. The matrices of S and T are

$$[S] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad [T] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

b. The matrices of the compositions are given by matrix multiplication:

$$[S \circ T] = [S][T] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad [T \circ S] = [T][S] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

c. The matrices $[S \circ T]$ and $[T \circ S]$ are inverses of each other: you can either compute it out, or note that since S and T are reflections, we have $S \circ S = T \circ T = I$, so

$$T \circ (S \circ S) \circ T = T \circ T = I \quad \text{and} \quad S \circ (T \circ T) \circ S = S \circ S = I.$$

d. They are the rotations by $2\pi/3$ and $-2\pi/3$ around the line $x = y = z$, counterclockwise if you look from a point of this line with positive coordinates towards the origin.

1.10 Let θ be the angle between A and $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, where each matrix is viewed as a vector in $\mathbb{R}^4$. Then

$$\cos\theta = \frac{A \cdot A^{-1}}{|A|\,|A^{-1}|} = \operatorname{sgn}(ad-bc)\frac{2ad - b^2 - c^2}{a^2+b^2+c^2+d^2}$$

The matrices are orthogonal if $2ad = b^2 + c^2$.

1.11 Below we denote by $|\overrightarrow{\text{side}}|$ the length of the side.

a. Because the side and the diagonal define a right triangle,

$$\text{angle between } \overrightarrow{\text{side}} \text{ and } \overrightarrow{\text{diagonal}} = \arccos\left(\frac{|\overrightarrow{\text{side}}|}{|\overrightarrow{\text{diagonal}}|}\right)$$

$$\theta_x = \arccos\left(\frac{a}{\sqrt{a^2+b^2+c^2}}\right) \qquad \theta_y = \arccos\left(\frac{b}{\sqrt{a^2+b^2+c^2}}\right)$$

$$\theta_z = \arccos(\frac{c}{\sqrt{a^2+b^2+c^2}})$$

b. Volume (parallelepiped) = abc = area(base) × height, but
height = length(diagonal) × sin(angle of diagonal with face), so

$$\begin{aligned}
\text{angle } &= \arcsin\left(\frac{abc}{\text{area(base)}\sqrt{a^2+b^2+c^2}}\right)\\
\theta_{x-y} &= \arcsin\left(\frac{abc}{ab\sqrt{a^2+b^2+c^2}}\right) = \arcsin\left(\frac{c}{\sqrt{a^2+b^2+c^2}}\right)
\end{aligned}$$

$$\begin{aligned}
\theta_{x-z} &= \arcsin\left(\frac{abc}{ac\sqrt{a^2+b^2+c^2}}\right) = \arcsin\left(\frac{b}{\sqrt{a^2+b^2+c^2}}\right)\theta_{y-z}\\
&= \arcsin\left(\frac{abc}{bc\sqrt{a^2+b^2+c^2}}\right) = \arcsin\left(\frac{a}{\sqrt{a^2+b^2+c^2}}\right)
\end{aligned}$$

1.12 a. We have

$$\vec{\mathbf{a}}\times\vec{\mathbf{b}} = \begin{bmatrix}1\\0\\1\end{bmatrix}\times\begin{bmatrix}2\\-1\\1\end{bmatrix} = \begin{bmatrix}1\\1\\-1\end{bmatrix},\ \vec{\mathbf{b}}\times\vec{\mathbf{c}} = \begin{bmatrix}2\\-1\\1\end{bmatrix}\times\begin{bmatrix}0\\1\\-1\end{bmatrix} = \begin{bmatrix}0\\2\\2\end{bmatrix},$$

$$\vec{\mathbf{c}}\times\vec{\mathbf{a}} = \begin{bmatrix}0\\1\\-1\end{bmatrix}\times\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}1\\-1\\-1\end{bmatrix}$$

so

$$Q_A = \begin{bmatrix}0&2&2\\1&-1&-1\\1&1&-1\end{bmatrix}.$$

b. We have

$$Q_A\,A = \begin{bmatrix}0&2&2\\1&-1&-1\\1&1&-1\end{bmatrix}\begin{bmatrix}1&2&0\\0&-1&1\\1&1&-1\end{bmatrix} = \begin{bmatrix}2&0&0\\0&2&0\\0&0&2\end{bmatrix}.$$

c. For the product $Q_A\,A$ in general, we find

$$\begin{array}{cc}
 & \begin{bmatrix}\vec{\mathbf{a}} & \vec{\mathbf{b}} & \vec{\mathbf{c}}\end{bmatrix}\\
\begin{bmatrix}(\vec{\mathbf{b}}\times\vec{\mathbf{c}})^\top\\(\vec{\mathbf{c}}\times\vec{\mathbf{a}})^\top\\(\vec{\mathbf{a}}\times\vec{\mathbf{b}})^\top\end{bmatrix} & \begin{bmatrix}\vec{\mathbf{a}}\cdot(\vec{\mathbf{b}}\times\vec{\mathbf{c}}) & 0 & 0\\0 & \vec{\mathbf{b}}\cdot(\vec{\mathbf{c}}\times\vec{\mathbf{a}}) & 0\\0 & 0 & \vec{\mathbf{c}}\cdot(\vec{\mathbf{a}}\times\vec{\mathbf{b}})\end{bmatrix}
\end{array}.$$

The zeroes are there because the cross product of two vectors is orthogonal to both, and in each off-diagonal entry we have the dot product of the cross product of two vectors with one of the two. The entries on the diagonal are all equal to the determinant of A. So $Q_A\,A = (\det A)I$.

d. The two problems are identical.

1.13 a. The normalized vectors are:

$$\text{i. } \frac{1}{\sqrt{14}}\begin{bmatrix}2\\1\\3\end{bmatrix},\qquad \text{ii. } \frac{1}{\sqrt{13}}\begin{bmatrix}-2\\3\end{bmatrix},\qquad \text{iii. } \frac{1}{\sqrt{7}}\begin{bmatrix}\sqrt{3}\\0\\2\end{bmatrix}.$$

b. The angle θ satisfies $\cos\theta = \dfrac{2\sqrt{3}+6}{7\sqrt{2}}$, i.e., $\theta = \arccos\dfrac{2\sqrt{3}+6}{7\sqrt{2}}$.

1.14 a. We have $\theta_n = \arccos\dfrac{\sqrt{6}n(n+1)}{2\sqrt{2n^4+3n^3+n^2}}$, since

$$\cos\theta_n = \frac{\vec{\mathbf{v}}\cdot\vec{\mathbf{w}}}{|\vec{\mathbf{v}}||\vec{\mathbf{w}}|} = \frac{\dfrac{n(n+1)}{2}}{\sqrt{n}\sqrt{\dfrac{n(n+1)(2n+1)}{6}}} = \frac{\sqrt{6}n(n+1)}{2\sqrt{2n^4+3n^3+n^2}}$$

b. As $n\to\infty$, $\theta = 30^\circ = \pi/6$. The important terms are those with n^2; the others are negligible. This gives

$$\lim_{n\to\infty}\frac{n^2\sqrt{6}}{2n^2\sqrt{2}} = \frac{\sqrt{3}}{2}.$$

We can justify the statement "the others are negligible" by writing

$$\lim_{n\to\infty}\frac{\sqrt{6}n(n+1)}{2\sqrt{2n^4+3n^3+n^2}} = \lim_{n\to\infty}\frac{\sqrt{6}\,n^2(1+\frac{1}{n})}{2\sqrt{2}\,n^2\sqrt{1+\frac{3}{2n}+\frac{1}{2n^2}}} = \frac{\sqrt{6}\,n^2}{2\sqrt{2}\,n^2} = \frac{\sqrt{6}}{2\sqrt{2}}.$$

Solution 1.15: Notation for sequences can be ambiguous. A sequence $x_1, x_2, \dots$ may be written x_i, but x_i can also refer to some collection of the elements of the sequence. The correct way to describe the sequence in part a is: Let $i\mapsto \mathbf{x}_i$ be a sequence in $\cap_i C_i$, i.e, a map $\mathbb{N}\to\cap_i C_i$.

1.15 a. Let $C_i,\ i\in I$ be some collection of closed subsets of $\mathbb{R}^n$. We will use proposition 1.5.17 to show that their intersection is closed. Indeed, let $\mathbf{x}_1, \mathbf{x}_2, \dots$ be a convergent sequence in $\cap_{i\in I}C_i$, converging in $\mathbb{R}^n$ to some $\mathbf{x}_0$. Then the sequence $\mathbf{x}_i$ belongs to each C_i, and since the C_i are closed, we have $\mathbf{x}_0\in C_i$ for each $i\in I$. Therefore $\mathbf{x}_0\in\cap_{i\in I}C_i$.

b. Again, we will use proposition 1.5.17. Let $C_1,\dots,C_m$ be a finite collection of closed subsets of $\mathbb{R}^m$. Suppose that $\mathbf{x}_1,\mathbf{x}_2,\dots$ is a convergent sequence in the union $\cup_{i=1}^m C_i$, converging in $\mathbb{R}^n$ to some $\mathbf{x}_0$. Then infinitely many of the entries of the sequence must be elements of a single C_k; these form a subsequence, which still converges to $\mathbf{x}_0$ by proposition 1.5.19. Hence $\mathbf{x}_0$ is an element of C_k, hence also an element of $\cup_{i=1}^m C_i$. It follows from proposition 1.5.17 that the union is closed.

c. The union of the closed sets $[0,(n-1)/n]$, $n = 2,3,4,\dots$ is the nonclosed set $[0,1)$.

1.16 Suppose that $\mathbf{x}_1,\mathbf{x}_2,\dots$ is a convergent sequence, converging in $\mathbb{R}^n$ to some $\mathbf{x}_0$; and that each $\mathbf{x}_i\in\overline{U}, i=1,2,\dots$. We need to prove that in that case $\mathbf{x}_0$ also belongs to $\overline{U}$, i.e. that there exists a sequence $\mathbf{y}_i$ in U (as opposed to $\overline{U}$) which converges to $\mathbf{x}_0$. By hypothesis, there exist sequences $\mathbf{z}_{i,j}$ in U such that $\mathbf{z}_{1,i},\ \mathbf{z}_{2,i},\ \mathbf{z}_{3,i},\dots$ converges to $\mathbf{x}_i$ for each i. For each $N=1,2,\dots$, find $i(N)$ such that $|\mathbf{x}_{i(N)}-\mathbf{x}_0| < 1/(2N)$, then find $j(N)$ such that $\mathbf{z}_{j(N),i(N)}-\mathbf{x}_{i(N)}| < 1/(2N)$. Consider the sequence

$$\mathbf{z}_{j(1),i(1)},\ \mathbf{z}_{j(2),i(2)},\ \mathbf{z}_{j(3),i(3)},\dots.$$

It is a sequence in U, and since

$$|\mathbf{x}_0-\mathbf{z}_{j(N),i(N)}| \le |\mathbf{x}_0-\mathbf{x}_{i(N)}| + |\mathbf{x}_{i(N)}-\mathbf{z}_{j(N),i(N)}| < 1/N,$$

this sequence converges to $\mathbf{x}_0$.

1.17 a. The derivative of the function zy^2 at $\mathbf{p} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is $[0\ 2\ 1]$, so the directional derivatives at $\mathbf{p}$ in the directions $\vec{\mathbf{e}}_1$, $\vec{\mathbf{e}}_2$, $\vec{\mathbf{e}}_3$, $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$ are

$$[0\ 2\ 1]\vec{\mathbf{e}}_1 = 0, \qquad [0\ 2\ 1]\vec{\mathbf{e}}_2 = 2, \qquad [0\ 2\ 1]\vec{\mathbf{e}}_3 = 1, \qquad [0\ 2\ 1]\vec{\mathbf{v}}_1 = \sqrt{2}/2$$
$$[0\ 2\ 1]\vec{\mathbf{v}}_2 = 3\sqrt{2}/2.$$

So the function zy^2 increases most slowly in the direction $\vec{\mathbf{e}}_1$.

b. The derivative of the function $2x^2 - y^2$ at $\mathbf{p}$ is $[4\ -2\ 0]$, giving the directional derivatives at $\mathbf{p}$

$$[4\ \ -2\quad 0]\vec{\mathbf{e}}_1 = 4, \qquad [4\ \ -2\quad 0]\vec{\mathbf{e}}_2 = -2, \qquad [4\ \ -2\quad 0]\vec{\mathbf{e}}_3 = 0$$
$$[4\ \ -2\quad 0]\vec{\mathbf{v}}_1 = 2\sqrt{2}, \qquad [4\ \ -2\quad 0]\vec{\mathbf{v}}_2 = -\sqrt{2}.$$

So to make $2x^2 - y^2$ increase as much as possible, also choose direction $\vec{\mathbf{e}}_1$.

3. To make $2x^2 - y^2$ decrease as much as possible, choose direction $\vec{\mathbf{e}}_2$.

1.18 a. $f(x) = \left| \begin{pmatrix} 2 \\ 3 \end{pmatrix} - \begin{pmatrix} x \\ x^2 \end{pmatrix} \right|^2 = x^4 - 5x^2 - 4x + 13.$

b. $Df = 4x^3 - 10x - 4 = 0 \Rightarrow x^3 - \frac{10}{4}x - 1 = 0$, so in the nomenclature of Section 0.7 (Equation A2.10) we have $p = -\frac{10}{4}$, and $q = -1$, and the discriminant $\Delta = 27q^2 + 4p^3 = 27 + 4(-10/4)^3$ is negative. That means that there are three real roots, and they can be found using the formula of Exercise A2.6. This gives the following roots:

$$x_k = \frac{\cos\left(\frac{1}{3}\arccos\left(\sqrt{\frac{54}{125}}\right) + \frac{2\pi k}{3}\right)}{\sqrt{3/10}}, \qquad k = 0, 1, 2.$$

c. The function goes to positive infinity when $x \to \pm\infty$, so we know it must take on an absolute minimum. This minimum must occur at one of the roots of the derivative of the function. Since the function takes on a lower value at x_0 than at either x_1 or x_2, $r = f(x_0)$.

1.19 a. This uses a trick:

$$\frac{x+y}{x^2 - y^2} = \frac{x+y}{(x+y)(x-y)} = \frac{1}{x-y};$$

there is no limit as $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, since when $\begin{pmatrix} x \\ y \end{pmatrix}$ is close to the origin, $x - y$ is also small (perhaps 0), so the quotient is big (or undefined).

b. Again the limit does not exist: on the line $y = kx$, this function is

$$\frac{x^4(1+k^2)^2}{x(1+k)} = x^3\frac{(1+k^2)^2}{1+k}.$$

Choose $\epsilon > 0$, and set $1 + k = \epsilon^4$ and $x = \epsilon$; the function becomes

$$\frac{\epsilon^3}{\epsilon^4}(1+k^2) > \frac{1}{\epsilon}.$$

Thus there are points near the origin where the function is arbitrarily large. But there are also points where the function is arbitrarily close to 0, taking $x = y = \epsilon$.

The statement
$$\lim_{u\to 0} u \ln |u| = 0,$$
can be proved using l'Hôpital's rule applied to $\dfrac{\ln|u|}{1/u}$. We used this already in solution 1.5.21.

c. Since $\lim_{u\to 0} u \ln|u| = 0$, the limit is 0.

d. Let us look at the function on the line $y = \epsilon$, for some $\epsilon > 0$. The function becomes

$$(x^2+\epsilon^2)(\ln|x| + \ln\epsilon) = x^2\ln|x| + \epsilon^2\ln|x| + x^2\ln\epsilon + \epsilon^2\ln\epsilon.$$

When $|x|$ is small, the first, third, and fourth terms are small. But the second is not. If $x = e^{-1/\epsilon^3}$, for instance, the second term is

$$-\frac{\epsilon^2}{\epsilon^3} = -\frac{1}{\epsilon},$$

which will become arbitrarily large as $\epsilon \to 0$. Thus along the curve $x = e^{-1/\epsilon^3}, y = \epsilon$, the function tends to $-\infty$. But along the curve $x = y = \epsilon$, the function tends to 0, so it has no limit.

1.20 Proof of Theorem 1.5.28

a. Choose $\epsilon > 0$, and find $\delta_1, \delta_2 > 0$ such that

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)| < \epsilon/2 \quad \text{when } \mathbf{x}\in U \text{ and } |\mathbf{x}-\mathbf{x}_0| < \delta_1$$
$$|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{x}_0)| < \epsilon/2 \quad \text{when } \mathbf{x}\in U \text{ and } |\mathbf{x}-\mathbf{x}_0| < \delta_2.$$

Set $\delta = \min\{\delta_1, \delta_2\}$ Then

$$\begin{aligned}|h(\mathbf{x})f(\mathbf{x}) - h(\mathbf{x}_0)f(\mathbf{x}_0)| &\le |h(\mathbf{x})f(\mathbf{x}) - h(\mathbf{x}_0)f(\mathbf{x})| + |h(\mathbf{x}_0)f(\mathbf{x}) - h(\mathbf{x}_0)f(\mathbf{x}_0)|\\ &\le |f(\mathbf{x})|\,|h(\mathbf{x}) - h(\mathbf{x}_0)| + |h(\mathbf{x}_0)|\,|f(\mathbf{x}) - f(\mathbf{x}_0)|\\ &\le \big(|f(\mathbf{x}_0)| + \epsilon\big)\epsilon + |h(\mathbf{x}_0)|\epsilon = \big(|f(\mathbf{x}_0)| + |h(\mathbf{x}_0)| + \epsilon\big)\epsilon\end{aligned}$$

when $\mathbf{x}\in U$ and $|\mathbf{x}-\mathbf{x}_0| < \delta$.

b. Choose $\epsilon > 0$, and find $\delta_1, \delta_2 > 0$ such that

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)| < \epsilon \quad \text{when } \mathbf{x}\in U \text{ and } \mathbf{x}-\mathbf{x}_0| < \delta_1$$
$$|h(\mathbf{x}) - h(\mathbf{x}_0)| < \epsilon \quad \text{when } \mathbf{x}\in U \text{ and } \mathbf{x}-\mathbf{x}_0| < \delta_2.$$

Set $\delta = \min\{\delta_1, \delta_2\}$ Then

$$\begin{aligned}|h(\mathbf{x})f(\mathbf{x}) - h(\mathbf{x}_0)f(\mathbf{x}_0)| &\le |h(\mathbf{x})f(\mathbf{x}) - h(\mathbf{x}_0)f(\mathbf{x})| + |h(\mathbf{x}_0)f(\mathbf{x}) - h(\mathbf{x}_0)f(\mathbf{x}_0)|\\ &\le |f(\mathbf{x})|\,|h(\mathbf{x}) - h(\mathbf{x}_0)| + |h(\mathbf{x}_0)|\,|f(\mathbf{x}) - f(\mathbf{x}_0)|\\ &\le (|f(\mathbf{x}_0)| + \epsilon)\epsilon + |h(\mathbf{x}_0|\epsilon = (|f(\mathbf{x}_0)| + |h(\mathbf{x}_0| + \epsilon)\epsilon\end{aligned}$$

when $\mathbf{x}\in U$ and $|\mathbf{x}-\mathbf{x}_0| < \delta$.

c. Using (b), it is enough to show that $1/h$ is continuous at $\mathbf{x}_0$. Choose $\epsilon > 0$, and find $\delta > 0$

$$|h(\mathbf{x}) - h(\mathbf{x}_0)| < \epsilon \text{ and } |h(\mathbf{x})| > |h(\mathbf{x}_0)|/2 \quad \text{when } \mathbf{x}\in U \text{ and } |\mathbf{x}-\mathbf{x}_0| < \delta.$$

Then

$$\left|\frac{1}{h(\mathbf{x})} - \frac{1}{h(\mathbf{x}_0)}\right| = \left|\frac{h(\mathbf{x}_0) - h(\mathbf{x})}{h(\mathbf{x})h(\mathbf{x}_0)}\right| \le \frac{2\epsilon}{|h(x_0)|^2}.$$

d. We can write $\mathbf{f} \cdot \mathbf{g} = \sum f_i g_i$. Each of the summands is continuous by (b), so their sum is continuous by (a).

e. Choose $\epsilon > 0$, and find $\delta > 0$ such that $|h(\mathbf{x})| \le \epsilon$ when $\mathbf{x} \in U$ and $|\mathbf{x} - \mathbf{x}_0| < \delta$. By hypothesis, there exists M such that $|\mathbf{f}(\mathbf{x})| \le M$ for all $\mathbf{x} \in U$. Thus

$$|h(\mathbf{x})f(\mathbf{x})| \le \epsilon M$$

when $\mathbf{x} \in U$ and $|\mathbf{x} - \mathbf{x}_0| < \delta$.

Proof of Theorem 1.5.29. [not yet written]

1.21 Let π_n be π to n places (e.g., $\pi_2 = 3.14$). Then $\pi - \pi_n < 10^{-n}$. We can do the same with e. So:

$$\left| \mathbf{a}_n - \begin{bmatrix} \pi \\ e \end{bmatrix} \right| < \sqrt{2 \cdot 10^{2(-n)}} = (\sqrt{2})10^{-n} < 10^{-n+1}.$$

This means that the best we can say in general is that to get $|\mathbf{a}_n - \begin{bmatrix} \pi \\ e \end{bmatrix}| < 10^{-m}$ we need $n = m + 1$. When $n = 3$, one less is enough, because $(.59\dots)^2 + (.28\dots)^2 < 1$. But when $m = 4$, we really need $m = 5$, since $(.92\dots)^2 + (.81\dots)^2 > 1$

1.22 If $\theta = 2k\pi$ with $k \in \mathbb{Z}$, then the sequence is constant and converges. Otherwise the sequence does not converge ($m\theta$ does not converge modulo 2π). The sequence always has a convergent subsequence: if θ is a rational multiple of π then there is a constant subsequence. Otherwise $\forall M \in N, \epsilon > 0\, \exists m > M$ such that $|m\theta| < \epsilon$ modulo 2π ($\frac{\theta}{2\pi}$ is irrational).

1.23 Let $a_n z^n$ be the term of highest degree. Let $c > 0$ be the real number such that $a_n c^n = p(c) - a_n c^n$. Then for any $R > c$ we know that $p(z) \neq 0$ for any $|z| \ge R$. (The term in z^{10} is nonzero and the other terms together cannot match it.) We know by the fundamental theorem of algebra that $p(z)$ must have at least one root (and we know by corollary 1.6.14 that it has exactly ten). Therefore we know that $p(z)$ has a root for $|z| < R$ if $R > c$.

1.24 The derivative of f is the $n \times n$ matrix whose i, j^{th} entry is

$$D_j f_i = D_j((x_1^2 + x_2^2 + \dots + x_n^2)x_i).$$

This is $2x_i x_j$ if $i \neq j$ and $(x_1^2 + x_2^2 + \dots + x_n^2) + 2x_i^2$ if $i = j$.

1.25 We have $\sqrt{x^2} = |x|$, so $\lim_{h \to 0} \frac{1}{h}(f(h) - f(0)) = \lim_{h \to 0} \frac{|h|}{h}$ which does not exist (since it is ± 1 depending on whether h is positive or negative).

For $\sqrt[3]{x^2}$, we have $\lim_{h \to 0} \frac{1}{h}(f(h) - f(0)) = \lim_{h \to 0} \frac{h^{2/3}}{h} = \lim_{h \to 0} \frac{1}{h^{1/3}}$, which tends to $\pm\infty$.

$\sqrt{x^4} = x^2$ is differentiable, of course.

1.26 a. The derivative of

$$C : \text{Mat}\,(2, 2) \to \text{Mat}\,(2, 2), \quad A \mapsto A^3$$

is $[DC(A)]H = A^2H + AHA + HA^2$. This is seen as follows:

$$(A+H)^3 = A^3 + A^2H + AHA + HA^2 + AH^2 + HAH + H^2A + H^3,$$

and the linear terms in H are $A^2H + AHA + HA^2$; this should be the derivative (see Remark 1.7.6). We can confirm this is true by writing

$$\begin{aligned}\frac{|(A+H)^3 - A^3 - (A^2H + AHA + HA^2)|}{|H|} &= \frac{|AH^2 + HAH + H^2A + H^3|}{|H|}\\ &\le \frac{3|A||H|^2 + |H|^2|H|}{|H|} = 3|A||H| + |H|^2.\end{aligned}$$

Clearly this goes to 0 as $H \to 0$. (Note that taking absolute values of the numerator is justified because if the length of a matrix goes to 0, the matrix necessarily goes to 0.) What theorem justifies the inequality above?[1]

b. Let us denote the mapping $A \mapsto A^k$ by P (for "power"). The linear terms in H of $P(A+H) - P(A)$ are

$$\sum_{i=0}^{k-1} A^iHA^{k-1-i} = HA^{k-1} + AHA^{k-2} + \cdots A^{k-2}HA + A^{k-1}H. \qquad (1)$$

(You can find this by trying some values of k, simplifying the computations by immediately disregarding quadratic and higher terms of H. For example, $(A+H)^3$ has only four relevant terms, $A^3 + AHA^2 + HA^3 + A^3H$; when multiplying this by $(A+H)$ to compute $(A+H)^4 - A^4$, we don't bother to multiply H times any term that has an H.)

We confirm that equation (1) is indeed the derivative by the following argument: the difference

$$(A+H)^k - A^k - \sum_{i=0}^{k-1} A^iHA^{k-1-i}$$

consists of sum of (a lot of) terms X_m, each of which is a product of at least $j \ge 2$ factors of H, and $k-j$ factors of A. By proposition 1.4.11, we have $|X_m| \le |H|^j|A|^{k-j}$. Therefore

$$\frac{|X_m|}{|H|} \le |H|^{j-1}|A|^{k-j}, \quad \text{and} \quad \lim_{H\to 0}\frac{|X_m|}{|H|} = 0.$$

Finally, we have

$$\lim_{H\to 0}\frac{1}{|H|}\left|(A+H)^k - A^k - \sum_{i=0}^{k-1} A^iHA^{k-1-i}\right| \le \lim_{H\to 0}\sum_m \frac{|X_m|}{|H|} = 0.$$

1.27 a. Both partials exist at $\begin{pmatrix}0\\0\end{pmatrix}$ and are 0, but the function is not differentiable. Indeed, if it were, the derivative would necessarily be the Jacobian matrix, i.e., the 0 matrix, and we would have

$$\lim_{\vec{\mathbf{h}}\to\mathbf{0}}\frac{f(\vec{\mathbf{h}}) - f(\mathbf{0}) - [0,0]\vec{\mathbf{h}}}{|\vec{\mathbf{h}}|} = 0.$$

[1] Theorem 1.4.9, the triangle inequality.

But writing the definition out leads to

$$\lim_{\vec{\mathbf{h}}\to\mathbf{0}} \frac{h_1^2 h_2}{(h_1^2+h_2^2)\sqrt{h_1^2+h_2^2}} = 0,$$

which isn't true: for instance, if you set $h_1 = h_2 = t$, the expression above becomes

$$\frac{t^3}{2\sqrt{2}\,|t|^3}, \quad \text{which does not become small as } t \to 0.$$

Solution 1.27,part b: If f were differentiable, then by the chain rule the composition would be differentiable, and it isn't.

b. This function is not differentiable. If you set $g(t) = \begin{pmatrix} t \\ t \end{pmatrix}$, then $(f \circ g)(t) = 2|t|$ is not differentiable at $t = 0$, but g is differentiable at $t = 0$, so f is not differentiable at the origin, which is $g(0)$.

c. Here are two proofs of part c:

First proof This isn't even continuous at the origin, although both partials exist there and are 0. But if you set $x = t, y = t$, then

For x small, $\sin x \approx x$; see equation 3.4.6 for the Taylor polynomial of $\sin(x)$.

$$\frac{\sin(xy)}{x^2+y^2} = \frac{\sin t^2}{2t^2} \to \frac{1}{2}, \quad \text{as } t \to 0.$$

Second proof This function is not differentiable; although both partial derivatives exist at the origin, the function itself is not continuous at the origin. For example, along the diagonal,

$$\lim_{t\to 0} \frac{\sin t^2}{2t^2} = \frac{1}{2},$$

but the limit along the antidiagonal is $-1/2$:

$$\lim_{t\to 0} \frac{\sin(-t^2)}{2t^2} = -\frac{1}{2}.$$

1.28 a. The area is given by the length of the cross product of the two vectors:

$$\left| \begin{bmatrix} u \\ 0 \\ u^2 \end{bmatrix} \times \begin{bmatrix} 0 \\ v^2 \\ v \end{bmatrix} \right| = \left| \begin{bmatrix} -u^2v^2 \\ -uv \\ uv^2 \end{bmatrix} \right| = \sqrt{u^4v^4 + u^2v^2 + u^2v^4}.$$

b. The quantity under the square root is strictly positive except on the axes, so A is differentiable except on the axes, and we find

$$\left[\mathbf{D}A\begin{pmatrix} u \\ v \end{pmatrix}\right] = \frac{1}{\sqrt{u^4v^4 + u^2v^4 + u^2v^2}}[2u^3v^4+uv^2+uv^4,\ 2u^4v^3+u^2v+2u^2v^3].$$

Thus $\left[\mathbf{D}A\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right] = [4/\sqrt{3},\ -5/\sqrt{3}]$. Evaluated on $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ this gives $-6/\sqrt{3}$.

c. Among vectors with length 1, the direction in which the area increases fastest is $\begin{bmatrix} 4/\sqrt{41} \\ -5/\sqrt{41} \end{bmatrix}$. To compute this, we use $\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = |\vec{\mathbf{v}}| \cdot |\vec{\mathbf{w}}| \cos\theta$, where θ is the angle between the two vectors; this quantity is greatest when $\cos\theta = 1$, i.e., when the two vectors point in the same direction. Therefore we wanted a vector with length 1 that points in the same direction as $\begin{bmatrix} 4/\sqrt{3} \\ -5/\sqrt{3} \end{bmatrix}$, which is any multiple of $\begin{bmatrix} 4 \\ -5 \end{bmatrix}$. To normalize $\begin{bmatrix} 4 \\ -5 \end{bmatrix}$, we divide by its length.

d. At $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, A is still $\sqrt{3}$, and $\left[\mathbf{D}A\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right] = [4/\sqrt{3},\ 5/\sqrt{3}]$. Evaluated on $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ this gives $14/\sqrt{3}$.

e. Among vectors with length 1, the direction in which the area increases fastest is $\begin{bmatrix} 4/\sqrt{41} \\ 5/\sqrt{41} \end{bmatrix}$.

f. The mapping A is not differentiable at points where $u \neq 0$ and $v = 0$ and at points where $u = 0$ and $v \neq 0$. For instance, if $u = 1$, we have

$$A\begin{pmatrix} 1 \\ v \end{pmatrix} = |v|\sqrt{1 + 2v^2}.$$

1.29 a. By the chain rule this is

$$[\mathbf{D}f(t)] = \left[-\frac{1}{t + \sin t}, \frac{1}{t^2 + \sin(t^2)}\right]\begin{bmatrix} 1 \\ 2t \end{bmatrix} = \frac{2t}{t^2 + \sin(t^2)} - \frac{1}{t + \sin t}.$$

b. We defined f for $t > 1$, but we will analyze it for all values of t. The function f is increasing for all $t > 0$ and decreasing for all $t < 0$. First let us see that f increases for $t > 0$, i.e., that the derivative is strictly positive for all $t > 0$. To see this, put the derivative on a common denominator:

$$[\mathbf{D}f(t)] = \frac{2t^2 + 2t\sin t - t^2 - \sin(t^2)}{(t^2 + \sin(t^2))(t + \sin t)}. \tag{1}$$

For $x > 0$, we have $x > \sin x$ (try graphing the functions x and $\sin x$), so for $t > 0$, the denominator is strictly positive. We need to show that the numerator is also strictly positive, i.e.,

$$t^2 + 2t\sin t - \sin(t^2) > 0. \tag{2}$$

For $-\pi < t \leq \pi$, this is true: for $-\pi < t \leq \pi$, we know that t and $\sin t$ are both positive or both negative, so $2t\sin t > 0$, and we have $t^2 > \sin(t^2)$ for $t \neq 0$. (Do not worry about $t^2 < t$ for small t; if you set $x = t^2$, the formula $x > \sin x$ still applies for $x > 0$.)

For $t > \pi$, equation (2) is also true: in that case,

$$t^2 + 2t\sin t - \sin(t^2) \geq t(t + 2\sin t) - 1 \geq t(\pi - 2) - 1 > \pi - 1.$$

To show that the function is decreasing for $t < 0$, we must show that the derivative is negative. For $t < 0$, the numerator is still strictly positive, by the argument above. But the denominator is negative: $t^2 + \sin(t^2)$ is positive, but $t + \sin t$ is negative.

The function f tends to $-\infty$ as t tends to 0. To see this, note that if s is small, $\sin s$ is very close to s (see the margin note for solution 1.27). So

$$\begin{aligned}\int_t^{t^2} \frac{ds}{s + \sin s} &\approx \int_t^{t^2} \frac{ds}{2s} \\ &= \frac{1}{2}\Big(\ln|t^2| - \ln|t|\Big) \\ &= \frac{1}{2}(2\ln|t| - \ln|t| \\ &= \frac{\ln|t|}{2}.\end{aligned}$$

1.30 a. Let $f(A) = A^3$ and $g(A) = A^{-1}$. Then:

$$\begin{aligned}[\mathbf{D}f \circ g(A)]H &= [\mathbf{D}f(g(A))][\mathbf{D}g(A)]H = [\mathbf{D}f(g(A))]\ \overbrace{(-A^{-1}HA^{-1})}^{\text{from example 1.8.6}} \\ &= (A^{-1})^2(-A^{-1}HA^{-1}) + (A^{-1})(-A^{-1}HA^{-1})(A^{-1}) + (-A^{-1}HA^{-1})(A^{-1})^2 \\ &= -A^{-3}HA^{-1} - A^{-2}HA^{-2} - A^{-1}HA^{-3}\end{aligned}$$

b. If $f(A) = A^n$ and g is A^{-1} as above, then

$$[\mathbf{D}f(g(A))]H = \sum_{i=0}^{n-1} A^{-i}HA^{-(n-1-i)} = \sum_{i=0}^{n-1} A^{-i}HA^{i+1-n},$$

so:

$$[\mathbf{D}(f \circ g)(A)]H = \sum_{i=0}^{n-1} A^{-i}(-A^{-1}HA^{-1})A^{i+1-n} = -\sum_{i=0}^{n-1} A^{-i-1}HA^{i-n}$$

1.31 Set $f(A) = A^{-1}$ and $g(A) = AA^\top + A^\top A$. Then $F = f \circ g$, and we wish to compute

$$\begin{aligned}[\mathbf{D}F(A)]H &= [\mathbf{D}f \circ g(A)]H = [\mathbf{D}f(g(A))][\mathbf{D}g(A)]H \\ &= [\mathbf{D}f(AA^\top + A^\top A)]\ \underbrace{[\mathbf{D}g(A)]H}_{\substack{\text{new increment} \\ \text{for } \mathbf{D}f}}\ .\end{aligned} \tag{1}$$

Solution 1.31: Remember (see example 1.7.18) that we cannot treat equation (1) as matrix multiplication. To treat the derivatives of f and g as matrices, we would have to identify Mat (n, n) with $\mathbb{R}^{n^2}$; the derivatives would be $n^2 \times n^2$ matrices. Instead we think of the derivatives as linear transformations.

The linear terms in H of

$$g(A + H) - g(A) = (A + H)(A + H)^\top + (A + H)^\top(A + H) - AA^\top - A^\top A$$

are $AH^\top + HA^\top + A^\top H + H^\top A$; this is $[\mathbf{D}g(A)]H$, which is the new increment for $\mathbf{D}f$.

We know from proposition 1.7.19 that $[\mathbf{D}f(A)]H = -A^{-1}HA^{-1}$, which we will rewrite as

$$[\mathbf{D}f(B)]K = -B^{-1}KB^{-1} \tag{2}$$

to avoid confusion. We substitute $AH^\top + HA^\top + A^\top H + H^\top A$ for the increment K in equation (2) and $g(A) = AA^\top + A^\top A$ for B. This gives

$$\begin{aligned}[\mathbf{D}F(A)]H &= [\mathbf{D}f(AA^\top + A^\top A)][\mathbf{D}g(A)]H \\ &= \underbrace{(-AA^\top + A^\top A)^{-1}}_{-B^{-1}}\underbrace{(AH^\top + HA^\top + A^\top H + H^\top A)}_{K}\underbrace{(AA^\top + A^\top A)^{-1}}_{B^{-1}}.\end{aligned}$$

There is no obvious way to simplify this expression.

1.32 a. The only time the partials might not exist is when $x, y = 0$ (the numerator and the denominator are differentiable). If $x = 0$ or $y = 0$ then $xy = 0$ so $f(x, y)$ is constant and so the partials exist and are 0.

b. The only time f might not be differentiable is at the origin (the denominator goes to 0). If $x = y \neq 0$ then $f(x, y) = \frac{1}{2}$ so f is not differentiable (or indeed continuous) at the origin.

1.33 a. Except at the origin, all partials exist by theorem 1.7.9. At the origin, all partials exist and are 0, since the function vanishes identically on all three coordinate axes.

b. By theorem 1.8.1, the function is differentiable everywhere except at the origin. At the origin, the function is not differentiable. In fact, it isn't even continuous: the limit

$$\lim_{h\to 0} f\begin{pmatrix} t \\ t \\ t \end{pmatrix} = \lim_{h\to 0} \frac{h^3}{3h^4} = \lim_{h\to 0} \frac{1}{3h} \quad \text{does not exist.}$$

1.34

1.35

$$A^2 = \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)^2 = \begin{bmatrix} a^2+bc & b(a+d) \\ c(a+d) & bc+d^2 \end{bmatrix}$$

So if we want $A^2 = 0$ we need: $a^2 = d^2 = -bc$ $(a = \pm d)$ and either $b = c = 0$ or $a = -d$. If $a = d = 0$ then one of b or $c = 0$. If $a = d \neq 0$ then $b = c = 0$ so $a = d = 0$ so $A = 0$. If $a = -d \neq 0$ then we have $c = \frac{-a^2}{b}$.

If we want $A^2 = I$ we need $a^2 + bc = d^2 + bc = 1$ $(a = \pm d)$ and either $b = c = 0$ or $a = -d$. If $a = d = 0$ then $b = \frac{1}{c}$. If $b = c = 0$ then $a, d = \pm 1$. If one of $b, c \neq 0$ then we have $a = -d$, $c = \frac{1-a^2}{b}$.

If we want $A^2 = -I$ then we have almost the same. If $a = d = 0$ then $b = \frac{-1}{c}$. If $b = c = 0$ then $a, d = \pm i$. If one of $b, c \neq 0$ then we have $a = -d$, $c = \frac{-1-a^2}{b}$.

1.36

1.37 a. This is a case where it is much easier to think of first rotating the telescope so that it is in the (x, z)-plane, then changing the elevation, then rotating back. This leads to the following product of matrices:

$$\begin{bmatrix} \cos\theta_0 & -\sin\theta_0 & 0 \\ \sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{bmatrix} \begin{bmatrix} \cos\theta_0 & \sin\theta_0 & 0 \\ -\sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} \cos^2\theta_0\cos\varphi - \sin^2\theta_0 & \cos\theta_0\sin\theta_0(\cos\varphi - 1) & -\sin\varphi\cos\theta_0 \\ cos\theta_0\sin\theta_0(\cos\varphi - 1) & \cos\varphi\sin^2\theta_0 + \cos^2\theta_0 & -\sin\theta_0\sin\varphi \\ \sin\varphi\cos\theta_0 & \sin\varphi\sin\theta_0 & \cos\varphi \end{bmatrix}$$

You may wonder about the signs of the $\sin\omega$ terms. Once the telescope is level, it is pointing in the direction of the x-axis. You, the astronomer rotating the telescope, are at the negative x end of the telescope. If you rotate it counterclockwise, as seen by you, the matrix is as we say. On the other hand, we are not absolutely sure that the problem is unambiguous as stated.

b. It is best to think of first rotating the telescope into the (x,z)-plane, then rotating it until it is horizontal (or vertical), then rotating it on its own axis, and then rotating it back (in two steps). This leads to the following product of matrices:

$$\begin{bmatrix} \cos\theta_0 & -\sin\theta_0 & 0 \\ \sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\varphi_0 & 0 & -\sin\varphi_0 \\ 0 & 1 & 0 \\ \sin\varphi_0 & 0 & \cos\varphi_0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{bmatrix}$$
$$\begin{bmatrix} \cos\varphi_0 & 0 & \sin\varphi_0 \\ 0 & 1 & 0 \\ -\sin\varphi_0 & 0 & \cos\varphi_0 \end{bmatrix} \begin{bmatrix} \cos\theta_0 & \sin\theta_0 & 0 \\ -\sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$