1. (a) The inverse has the form

$$U^{-1} = \begin{bmatrix} 1/4 & a & b & c \\ 0 & -1/3 & d & e \\ 0 & 0 & 1 & f \\ 0 & 0 & 0 & 1/2 \end{bmatrix}$$

for some entries that we find by using the matrix equation $UU^{-1} = \mathrm{I}_4$. Multiplying, we have

$$\begin{aligned} UU^{-1} &= \begin{bmatrix} 4 & 2 & -4 & 8 \\ 0 & -3 & 6 & 12 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1/4 & a & b & c \\ 0 & -1/3 & d & e \\ 0 & 0 & 1 & f \\ 0 & 0 & 0 & 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 4a-2/3 & 4b+2d-4 & 4c+2e-4f+4 \\ 0 & 1 & -3d+6 & -3e+6f+6 \\ 0 & 0 & 1 & f+2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \end{aligned}$$

so equating this with I_4 with vanishing everywhere above the diagonal gives successive equations as we work our way into the upper right corner layer by layer from the diagonal: the layer just above the diagonal gives

$$4a - 2/3 = 0, \quad -3d + 6 = 0, \quad f + 2 = 0,$$

which solves to give $a = 1/6$, $d = 2$, $f = -2$. The next layer up gives

$$0 = 4b + 2d - 4 = 4b + 4 - 4 = 4b, \quad 0 = -3e + 6f + 6 = -3e - 12 + 6 = -3e - 6,$$

so $b = 0$ and $e = -2$. Finally, the upper right corner gives

$$0 = 4c + 2e - 4f + 4 = 4c - 4 - 4(-2) + 4 = 4c + 8,$$

so $c = -2$.

Putting it all together, we have

$$U^{-1} = \begin{bmatrix} 1/4 & 1/6 & 0 & -2 \\ 0 & -1/3 & 2 & -2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1/2 \end{bmatrix}.$$

To check this works, we multiply matrices

$$\begin{aligned} UU^{-1} &= \begin{bmatrix} 4 & 2 & -4 & 8 \\ 0 & -3 & 6 & 12 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1/4 & 1/6 & 0 & -2 \\ 0 & -1/3 & 2 & -2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2/3-2/3 & 4-4 & -8-4+8+4 \\ 0 & 1 & -6+6 & 6-12+6 \\ 0 & 0 & 1 & -2+2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \mathrm{I}_4. \end{aligned}$$

(b) The general solution is $\mathbf{x} = U^{-1}\mathbf{b}$, so using the answer in (a) we get

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/6 & 0 & -2 \\ 0 & -1/3 & 2 & -2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} b_1/4 + b_2/6 - 2b_4 \\ -b_2/3 + 2b_3 - 2b_4 \\ b_3 - 2b_4 \\ b_4/2 \end{bmatrix}.$$

The entries on the right give a general formula for each x_i when given $\mathbf{b}$. Taking $(b_1, b_2, b_3, b_4) = (8, -6, 2, 6)$, the general formula for the solution becomes

$$\begin{bmatrix} 2-1-12 \\ 2+4-12 \\ 2-12 \\ 3 \end{bmatrix} = \begin{bmatrix} -11 \\ -6 \\ -10 \\ 3 \end{bmatrix},$$

recovering the answer in Exercise 22.1(a). ◊

2. (a) The matrix product is

$$\begin{bmatrix} 3 & 2 & -2 \\ 15 & 10+2 & -10+2 \\ 9 & 6-4 & -6-4+4 \end{bmatrix} = \begin{bmatrix} 3 & 2 & -2 \\ 15 & 12 & -8 \\ 9 & 2 & -6 \end{bmatrix}.$$

(b) The inverses L^{-1} and U^{-1} have the form

$$L^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ a & 1/2 & 0 \\ b & c & -1 \end{bmatrix}, \quad U^{-1} = \begin{bmatrix} 1/3 & a' & b' \\ 0 & 1 & c' \\ 0 & 0 & -1/4 \end{bmatrix}$$

with entries determined by the vanishing of the off-diagonal entries in

$$LL^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 2 & 0 \\ 3 & -4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ a & 1/2 & 0 \\ b & c & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 5+2a & 1 & 0 \\ 3-4a-b & -2-c & 1 \end{bmatrix}$$

and

$$UU^{-1} = \begin{bmatrix} 3 & 2 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} 1/3 & a' & b' \\ 0 & 1 & c' \\ 0 & 0 & -1/4 \end{bmatrix} = \begin{bmatrix} 1 & 3a'+2 & 3b'+2c'+1/2 \\ 0 & 1 & c'-1/4 \\ 0 & 0 & 1 \end{bmatrix}.$$

This gives the systems of equations

$$5+2a = 0, \ , -2-c = 0, \ 3-4a-b = 0$$

and

$$3a'+2 = 0, \ c'-1/4 = 0, \ 3b'+2c'+1/2 = 0,$$

which we solve in succession. For the first system, we initially solve the first two equations to get $a = -5/2$ and $c = -2$, so the final equation can be solved for b to give $b = 3-4a = 3-(-10) = 13$. For the second system, we initially solve the first two equations to get $a' = -2/3$ and $c' = 1/4$, so the final equation can be solved for b' to give $b' = -(2/3)c' - 1/6 = -1/6 - 1/6 = -1/3$.

Putting it all together, we have

$$L^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -5/2 & 1/2 & 0 \\ 13 & -2 & -1 \end{bmatrix}, \quad U^{-1} = \begin{bmatrix} 1/3 & -2/3 & -1/3 \\ 0 & 1 & 1/4 \\ 0 & 0 & -1/4 \end{bmatrix}.$$

To check that this work, we directly compute the matrix products

$$LL^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 2 & 0 \\ 3 & -4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -5/2 & 1/2 & 0 \\ 13 & -2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 5-5 & 1 & 0 \\ 3+10-13 & -2+2 & 1 \end{bmatrix} = \mathrm{I}_3$$

and

$$UU^{-1} = \begin{bmatrix} 3 & 2 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} 1/3 & -2/3 & -1/3 \\ 0 & 1 & 1/4 \\ 0 & 0 & -1/4 \end{bmatrix} = \begin{bmatrix} 1 & -2+2 & -1+1/2+1/2 \\ 0 & 1 & 1/4-1/4 \\ 0 & 0 & 1 \end{bmatrix} = \mathrm{I}_3.$$

(c) We compute the product $U^{-1}L^{-1}$ is equal to

$$\begin{bmatrix} 1/3 & -2/3 & -1/3 \\ 0 & 1 & 1/4 \\ 0 & 0 & -1/4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -5/2 & 1/2 & 0 \\ 13 & -2 & -1 \end{bmatrix} = \begin{bmatrix} 1/3+5/3-13/3 & -1/3+2/3 & 1/3 \\ -5/2+13/4 & 1/2-1/2 & -1/4 \\ -13/4 & 1/2 & 1/4 \end{bmatrix}$$
$$= \begin{bmatrix} -7/3 & 1/3 & 1/3 \\ 3/4 & 0 & -1/4 \\ -13/4 & 1/2 & 1/4 \end{bmatrix}.$$

To check this is inverse to A, we multiply it against A on the right: this is the product

$$\begin{bmatrix} -7/3 & 1/3 & 1/3 \\ 3/4 & 0 & -1/4 \\ -13/4 & 1/2 & 1/4 \end{bmatrix} \begin{bmatrix} 3 & 2 & -2 \\ 15 & 12 & -8 \\ 9 & 2 & -6 \end{bmatrix}$$

that is equal to

$$\begin{bmatrix} -7+5+3 & -14/3+4+2/3 & 14/3-8/3-2 \\ 9/4-9/4 & 3/2-1/2 & -3/2+3/2 \\ -39/4+15/2+9/4 & -13/2+6+1/2 & 13/2-4-3/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad \diamond$$

3. (a) The matrix product is

$$\begin{bmatrix} 1 & 4/3+2/3 & -5/3-1/3 \\ 1 & 4/3-1/3 & -5/3+1/6-1/2 \\ 1 & 4/3-1/3 & -5/3+1/6+1/2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & -2 \\ 1 & 1 & -1 \end{bmatrix}.$$

(b) The matrix R^{-1} has the form

$$R^{-1} = \begin{bmatrix} 1/\sqrt{3} & a & b \\ 0 & \sqrt{3/2} & c \\ 0 & 0 & \sqrt{2} \end{bmatrix},$$

determined by the vanishing of the off-diagonal entries of

$$RR^{-1} = \begin{bmatrix} \sqrt{3} & 4/\sqrt{3} & -5/\sqrt{3} \\ 0 & \sqrt{2/3} & -1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & a & b \\ 0 & \sqrt{3/2} & c \\ 0 & 0 & \sqrt{2} \end{bmatrix}$$
$$= \begin{bmatrix} 1 & \sqrt{3}\,a+2\sqrt{2} & \sqrt{3}\,b+(4/\sqrt{3})c-5\sqrt{2/3} \\ 0 & 1 & \sqrt{2/3}\,c-1/\sqrt{3} \\ 0 & 0 & 1 \end{bmatrix}.$$

This amounts to the three equations

$$\sqrt{3}\,a+2\sqrt{2}=0, \quad \sqrt{2/3}\,c-1/\sqrt{3}=0, \quad \sqrt{3}\,b+(4/\sqrt{3})c-5\sqrt{2/3}=0.$$

The first two are solved to give $a = -2\sqrt{2/3}$ and $c = 1/\sqrt{2}$, so the final equation says

$$\sqrt{3}\,b + 2\sqrt{2/3} - 5\sqrt{2/3} = 0.$$

This is the same as $\sqrt{3}\,b = 3\sqrt{2/3} = \sqrt{6}$, so $b = \sqrt{2}$.

Putting it all together, we obtain

$$R^{-1} = \begin{bmatrix} 1/\sqrt{3} & -2\sqrt{2/3} & \sqrt{2} \\ 0 & \sqrt{3/2} & 1/\sqrt{2} \\ 0 & 0 & \sqrt{2} \end{bmatrix}.$$

To check that this works, we compute RR^{-1}: this is the matrix product

$$\begin{bmatrix} \sqrt{3} & 4/\sqrt{3} & -5/\sqrt{3} \\ 0 & \sqrt{2/3} & -1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & -2\sqrt{2/3} & \sqrt{2} \\ 0 & \sqrt{3/2} & 1/\sqrt{2} \\ 0 & 0 & \sqrt{2} \end{bmatrix}$$

that works out to be

$$\begin{bmatrix} 1 & -2\sqrt{2} + 2\sqrt{2} & \sqrt{6} + 2\sqrt{2/3} - 5\sqrt{2/3} \\ 0 & 1 & 1/\sqrt{3} - 1/\sqrt{3} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(where the vanishing of the upper right entry is because $2\sqrt{2/3} - 5\sqrt{2/3} = -3\sqrt{2/3} = -\sqrt{6}$).

(c) The matrix $R^{-1}Q^{\top}$ is the product

$$\begin{bmatrix} 1/\sqrt{3} & -2\sqrt{2/3} & \sqrt{2} \\ 0 & \sqrt{3/2} & 1/\sqrt{2} \\ 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ \sqrt{2/3} & -1/\sqrt{6} & -1/\sqrt{6} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

which works out to be

$$\begin{bmatrix} 1/3 - 4/3 & 1/3 + 2/3 - 1 & 1/3 + 2/3 + 1 \\ 1 & -1/2 - 1/2 & -1/2 + 1/2 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 2 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

To check this is inverse to A, we multiply it on the left by A to get the product

$$\begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & -2 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 2 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -1+2 & -2+2 & 2-2 \\ -1+1 & -1+2 & 2-2 \\ -1+1 & -1+1 & 2-1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad \diamond$$

4. (a) Since $Q'R' = QR$, if we multiply both sides by Q^{-1} on the left and R'^{-1} on the right then we get $Q^{-1}Q'R'R'^{-1} = Q^{-1}QRR'^{-1}$. The left side has the factor $R'R'^{-1} = \mathrm{I}_n$ and the right side has the factor $Q^{-1}Q = \mathrm{I}_n$, so making those substitutions gives $Q^{-1}Q'\mathrm{I}_n = \mathrm{I}_n RR'^{-1}$, or in other words $Q^{-1}Q' = RR'^{-1}$ as desired.

The inverse of an orthogonal matrix is known to be orthogonal (it is just the transpose, and we know that an orthogonal matrix has row vectors that are also mutually orthogonal unit vectors), and the inverse of an upper triangular matrix with nonzero diagonal entries is upper triangular with diagonal entries given by the reciprocals of the original diagonal entries (so those are positive if the original diagonal entries are positive). Hence, Q^{-1} is orthogonal and R'^{-1} is upper triangular with positive diagonal entries. We know that products of orthogonal matrices are orthogonal, so $Q^{-1}Q'$ is orthogonal. Likewise, RR'^{-1} is a product of upper triangular matrices, so it is also upper triangular due to how one multiplies such matrices (look at the 3×3 case to see how the pattern goes), and its diagonal entries are the products of the corresponding diagonal entries of R and R'^{-1} (again look at products of 3×3 upper triangular matrices to see how the diagonal works out in a product of two such matrices). So RR'^{-1} has its diagonal entries that are products of two positive numbers each and hence are positive.

(b) If $Q^{-1}Q' = \mathrm{I}_n$ then multiplying both sides by Q on the left gives $Q(Q^{-1}Q') = Q\mathrm{I}_n = Q$. But $Q(Q^{-1}Q') = (QQ^{-1})Q' = \mathrm{I}_nQ' = Q'$, so we have $Q' = Q$ as desired.

Likewise, if $RR'^{-1} = \mathrm{I}_n$ then multiplying both sides by R' on the right gives $(RR'^{-1})R' = \mathrm{I}_nR' = R'$. But $(RR'^{-1})R' = R(R'^{-1}R') = R\mathrm{I}_n = R$, so we have $R = R'$ as desired.

(c) Once we establish the general assertion about such matrices M, let's see how we can conclude by using (a) and (b). By (a), we have the matrix $Q^{-1}Q' = RR'^{-1}$ that is both orthogonal and upper triangular with positive entries. Hence, by the general fact about such matrices M that we're going to establish below, this common matrix must be I_n. Then the hypothesis of (b) holds, so by (b) we get $Q = Q'$ and $R = R'$ as desired.

It remains to show that if an orthogonal $n \times n$ matrix M is also upper triangular with positive diagonal entries then it equals I_n. We will work one step at a time, showing that the jth column of M is $\mathbf{e}_j$ for $j = 1, 2, 3, \ldots, n$ in turn. Since M is upper triangular with positive diagonal entry, its jth column $\mathbf{m}_j$ has vanishing entries below the jth position, with positive jth entry. In other words:

$$\mathbf{m}_j = \begin{bmatrix} m_{1j} \\ \vdots \\ m_{jj} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = m_{1j}\mathbf{e}_1 + \cdots + m_{jj}\mathbf{e}_j$$

with $m_{jj} > 0$.

For $j = 1$ we have $\mathbf{m}_1 = m_{11}\mathbf{e}_1$ with $m_{11} > 0$. But $\mathbf{m}_1$ is a unit vector by orthogonality of M, so $|m_{11}| = 1$, yet $m_{11} > 0$, so $m_{11} = |m_{11}| = 1$. Hence, $\mathbf{m}_1 = 1\mathbf{e}_1 = \mathbf{e}_1$ as desired. Next, for $j = 2$ we have $\mathbf{m}_2 = m_{12}\mathbf{e}_1 + m_{22}\mathbf{e}_2$ with $m_{22} > 0$. But $\mathbf{m}_2$ is orthogonal to $\mathbf{m}_1 = \mathbf{e}_1$ by orthogonality of M, so $0 = \mathbf{m}_1 \cdot \mathbf{m}_2 = \mathbf{e}_1 \cdot \mathbf{m}_2 = m_{12}$. Hence, $\mathbf{m}_2 = m_{22}\mathbf{e}_2$. But $\mathbf{m}_2$ is a unit vector (by orthogonality of M) and $m_{22} > 0$, so $m_{22} = 1$ just as we showed $m_{11} = 1$; this yields that $\mathbf{m}_2 = 1\mathbf{e}_2 = \mathbf{e}_2$, as desired.

The pattern persists. Suppose for some $1 \le j \le n-1$ that we have established $\mathbf{m}_r = \mathbf{e}_r$ for all $r \le j$; we want to deduce that $\mathbf{m}_{j+1} = \mathbf{e}_{j+1}$ and then we can continue all the way to the final column and finish. By the orthogonality of distinct columns of the orthogonal matrix M, we have for $r \le j$ that

$$0 = \mathbf{m}_r \cdot \mathbf{m}_{j+1} = \mathbf{e}_r \cdot \mathbf{m}_{j+1} = m_{r,j+1}.$$

This say that all entries in $\mathbf{m}_{j+1}$ above its diagonal entry vanish, so we are left with $\mathbf{m}_{j+1} = m_{j+1,j+1}\mathbf{e}_{j+1}$ where $m_{j+1,j+1} > 0$. Now we argue as we did for m_{11} and m_{22}: this $m_{j+1,j+1}\mathbf{e}_{j+1} = \mathbf{m}_{j+1}$ is a unit vector, we have $|m_{j+1,j+1}| = 1$. But $m_{j+1,j+1} > 0$, so $m_{j+1,j+1} = |m_{j+1,j+1}| = 1$ and hence $\mathbf{m}_{j+1} = 1\mathbf{e}_{j+1} = \mathbf{e}_{j+1}$ as desired. The pattern carries us through all of the columns, so M and I_n agree in every column and hence are equal as matrices. ◊

5. We can write

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

So

$$A\begin{bmatrix}0\\1\end{bmatrix} = A\left(\begin{bmatrix}1\\2\end{bmatrix} - \begin{bmatrix}1\\1\end{bmatrix}\right) = 3\begin{bmatrix}1\\2\end{bmatrix} - 2\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1\\4\end{bmatrix}.$$

◊

6. (a) We know

$$A\mathbf{v} = \lambda\mathbf{v}$$

and

$$A\mathbf{w} = \mu\mathbf{w}.$$

So together with the hint we get

$$(\lambda\mathbf{v})\cdot\mathbf{w} = \mathbf{v}\cdot(\mu\mathbf{w}).$$

(b) We simplify further to get

$$\lambda(\mathbf{v}\cdot\mathbf{w}) = \mu(\mathbf{v}\cdot\mathbf{w}),$$

so $(\lambda-\mu)(\mathbf{v}\cdot\mathbf{w}) = 0$. Since λ and μ are different, so $\lambda \neq \mu$, we must have $\mathbf{v}\cdot\mathbf{w} = 0$. ◊

7. (a) From the definition, if we swap the roles of $\mathbf{v}$ and $\mathbf{w}$ then the effect is to swap the order of subtraction in each vector entry of the cross product. Hence, $\mathbf{w}\times\mathbf{v} = -(\mathbf{v}\times\mathbf{w})$. Likewise, in the definition of $\mathbf{v}\times\mathbf{v}$ each vector entry is a difference of the form $v_iv_j - v_jv_i = 0$ with $i \neq j$, so it is the zero vector in $\mathbf{R}^3$.

(b) For the calculations,

$$\begin{aligned}
\mathbf{v}\times\mathbf{w} &= \det\begin{bmatrix}-1 & 3\\ 2 & 3\end{bmatrix}\mathbf{e}_1 - \det\begin{bmatrix}2 & 3\\ 1 & 3\end{bmatrix}\mathbf{e}_2 + \det\begin{bmatrix}2 & -1\\ 1 & 2\end{bmatrix}\mathbf{e}_3\\
&= (-3-6)\mathbf{e}_1 - (6-3)\mathbf{e}_2 + (4-(-1))\mathbf{e}_3\\
&= -9\mathbf{e}_1 - 3\mathbf{e}_2 + 5\mathbf{e}_3\\
&= \begin{bmatrix}-9\\-3\\5\end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
\mathbf{w}\times\mathbf{u} &= \det\begin{bmatrix}2 & 3\\ 3 & -2\end{bmatrix}\mathbf{e}_1 - \det\begin{bmatrix}1 & 3\\ 4 & -2\end{bmatrix}\mathbf{e}_2 + \det\begin{bmatrix}1 & 2\\ 4 & 3\end{bmatrix}\mathbf{e}_3\\
&= (-4-9)\mathbf{e}_1 - (-2-12)\mathbf{e}_2 + (3-8)\mathbf{e}_3\\
&= -13\mathbf{e}_1 + 14\mathbf{e}_2 - 5\mathbf{e}_3\\
&= \begin{bmatrix}-13\\14\\-5\end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
(\mathbf{v}\times\mathbf{w})\times\mathbf{u} &= \det\begin{bmatrix}-3 & 5\\ 3 & -2\end{bmatrix}\mathbf{e}_1 - \det\begin{bmatrix}-9 & 5\\ 4 & -2\end{bmatrix}\mathbf{e}_2 + \det\begin{bmatrix}-9 & -3\\ 4 & 3\end{bmatrix}\mathbf{e}_3\\
&= (6-15)\mathbf{e}_1 - (18-20)\mathbf{e}_2 + (-27-(-12))\mathbf{e}_3\\
&= -9\mathbf{e}_1 + 2\mathbf{e}_2 - 15\mathbf{e}_3\\
&= \begin{bmatrix}-9\\2\\-15\end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
\mathbf{v} \times (\mathbf{w} \times \mathbf{u}) &= \det \begin{bmatrix} -1 & 3 \\ 14 & -5 \end{bmatrix} \mathbf{e}_1 - \det \begin{bmatrix} 2 & 3 \\ -13 & -5 \end{bmatrix} \mathbf{e}_2 + \det \begin{bmatrix} 2 & -1 \\ -13 & 14 \end{bmatrix} \mathbf{e}_3 \\
&= (5 - 42)\mathbf{e}_1 - (-10 - (-39))\mathbf{e}_2 + (28 - 13)\mathbf{e}_3 \\
&= -37\mathbf{e}_1 - 29\mathbf{e}_2 + 15\mathbf{e}_3 \\
&= \begin{bmatrix} -37 \\ -29 \\ 15 \end{bmatrix}.
\end{aligned}$$

(c) From the definition,

$$(c\mathbf{v}) \times \mathbf{w} = \begin{bmatrix} (cv_2)w_3 - (cv_3)w_2 \\ (cv_3)w_1 - (cv_1)w_3 \\ (cv_1)w_2 - (cv_2)w_1 \end{bmatrix} = \begin{bmatrix} c(v_2w_3 - v_3w_2) \\ c(v_3w_1 - v_1w_3) \\ c(v_1w_2 - v_2w_1) \end{bmatrix} = c \begin{bmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{bmatrix} = c(\mathbf{v} \times \mathbf{w}).$$

Finally,

$$\begin{aligned}
(\mathbf{v} + \mathbf{v}') \times \mathbf{w} &= \begin{bmatrix} (v_2 + v_2')w_3 - (v_3 + v_3')w_2 \\ (v_3 + v_3')w_1 - (v_1 + v_1')w_3 \\ (v_1 + v_1')w_2 - (v_2 + v_2')w_1 \end{bmatrix} \\
&= \begin{bmatrix} v_2w_3 + v_2'w_3 - v_3w_2 - v_3'w_2 \\ v_3w_1 + v_3'w_1 - v_1w_3 - v_1'w_3 \\ v_1w_2 + v_1'w_2 - v_2w_1 - v_2'w_1 \end{bmatrix} \\
&= \begin{bmatrix} (v_2w_3 - v_3w_2) + (v_2'w_3 - v_3'w_2) \\ (v_3w_1 - v_1w_3) + (v_3'w_1 - v_1'w_3) \\ (v_1w_2 - v_2w_1) + (v_1'w_2 - v_2'w_1) \end{bmatrix} \\
&= \begin{bmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{bmatrix} + \begin{bmatrix} v_2'w_3 - v_3'w_2 \\ v_3'w_1 - v_1'w_3 \\ v_1'w_2 - v_2'w_1 \end{bmatrix} \\
&= \mathbf{v} \times \mathbf{w} + \mathbf{v}' \times \mathbf{w}.
\end{aligned}$$

(d) The length of $\mathbf{v} \times \mathbf{w} = \begin{bmatrix} -9 \\ -3 \\ 5 \end{bmatrix}$ is $\sqrt{81 + 9 + 25} = \sqrt{115}$ and

$$\|\mathbf{v}\| \, \|\mathbf{w}\| \sin(\theta) = \sqrt{14}\sqrt{14} \sin(\theta) = 14 \sin(\theta).$$

To compute $\sin(\theta)$, the dot product formula "$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \, \|\mathbf{w}\| \cos(\theta)$" says

$$9 = \sqrt{14}\sqrt{14} \cos(\theta),$$

so $\cos(\theta) = 9/14$. Hence, $\sin(\theta) = \sqrt{1 - (\cos\theta)^2} = \sqrt{1 - 81/196} = \sqrt{115/196} = \sqrt{115}/14$. Hence, finally, $\|\mathbf{v}\| \, \|\mathbf{w}\| \sin(\theta) = 14 \sin(\theta) = 14(\sqrt{115}/14) = \sqrt{115}$, which matches the calculation of $\|\mathbf{v} \times \mathbf{w}\|$. ♢

8. (a) We want to check that, for each $k \geq 2$, $(A^k)^\top = A^k$. But

$$(A^k)^\top = (AA \cdots A)^\top = A^\top A^\top \cdots A^\top = AA \cdots A = A^k,$$

as desired.

(b) By the Spectral Theorem (Theorem 24.1.4), $\mathbf{R}^n$ has a orthonormal basis of eigenvectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$ for A with respective eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbf{R}$ (not necessarily distinct). We have $A = WDW^{-1}$ where W is the $n \times n$ orthogonal matrix with ith column $\mathbf{w}_i$ and D is the diagonal matrix with ith diagonal entry λ_i. Assume A is invertible, so $A\mathbf{x} = \mathbf{0}$ only for $\mathbf{x} = \mathbf{0}$, so $A\mathbf{v} \neq \mathbf{0}$ for any eigenvector $\mathbf{v}$ of A. In particular, $A\mathbf{w}_i = \lambda_i \mathbf{w}_i$ is nonzero for each i, so every λ_i is nonzero.

If every λ_i is nonzero then D is invertible, with inverse given by the diagonal matrix whose ith entry is $1/\lambda_i$. Then $WD^{-1}W^{-1}$ works as an inverse to A:

$$\begin{aligned}(WD^{-1}W^{-1})(W^{-1}DW^{-1}) = WD^{-1}(WW^{-1})DW^{-1} &= WD^{-1}\mathrm{I}_n DW^{-1} \\ &= WD^{-1}DW^{-1} \\ &= W\mathrm{I}_n W^{-1} \\ &= WW^{-1} \\ &= \mathrm{I}_n\end{aligned}$$

and similarly for the product in the other direction (since $DD^{-1} = \mathrm{I}_n$). ♢

9. (a) Using our rules of matrix algebra, $(M^\top M)^\top = M^\top (M^\top)^\top = M^\top M$, so $M^\top M$ is indeed symmetric.

(b) We can write $q(\mathbf{x}) = \mathbf{x}^\top (M^\top M\mathbf{x})$ as a matrix product. Using the associativity of matrix multiplication, we see that this is equal to $(\mathbf{x}^\top M^\top)(M\mathbf{x}) = (M\mathbf{x}) \cdot (M\mathbf{x}) = \|M\mathbf{x}\|^2$. Since the squared-length of a vector (e.g. $M\mathbf{x}$) is always non-negative, we have $q(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbf{R}^n$. That is, q is positive-semidefinite.

(c) Since we can write $q(\mathbf{x}) = \|M\mathbf{x}\|^2$, we know that $q(\mathbf{x}) = 0$ when the length of the vector $M\mathbf{x}$ is zero. But a vector's length is zero precisely when it is the zero vector. That is, $q(\mathbf{x}) = 0$ precisely when $\mathbf{x} \in N(M)$. The condition that q be positive-definite (given that it is already positive-semidefinite) is the same as the condition that $q(\mathbf{x}) = 0$ only when $\mathbf{x} = 0$. Thus, this is equivalent to the condition that $N(M) = \{\mathbf{0}\}$. ♢

10. (a) We compute

$$A\mathbf{v}_1 = A\begin{bmatrix}1\\-2\\2\end{bmatrix} = \begin{bmatrix}-5+28+4\\-14-8-32\\2+32+20\end{bmatrix} = \begin{bmatrix}27\\-54\\54\end{bmatrix} = 27\begin{bmatrix}1\\-2\\2\end{bmatrix} = 27\mathbf{v}_1,$$

$$A\mathbf{v}_2 = A\begin{bmatrix}2\\2\\1\end{bmatrix} = \begin{bmatrix}-10-28+2\\-28+8-16\\4-32+10\end{bmatrix} = \begin{bmatrix}-36\\-36\\-18\end{bmatrix} = -18\begin{bmatrix}2\\2\\1\end{bmatrix} = -18\mathbf{v}_2,$$

and

$$A\mathbf{v}_3 = A\begin{bmatrix}-2\\1\\2\end{bmatrix} = \begin{bmatrix}10-14+4\\28+4-32\\-4-16+20\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix} = 0\begin{bmatrix}-2\\1\\2\end{bmatrix} = 0\mathbf{v}_3.$$

Thus, the respective eigenvalues are $\lambda_1 = 27$, $\lambda_2 = -18$, $\lambda_3 = 0$.

(b) Since $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis of $\mathbf{R}^3$, for any $\mathbf{x} \in \mathbf{R}^3$ we have

$$\mathbf{x} = \mathbf{Proj}_{\mathbf{v}_1}(\mathbf{x}) + \mathbf{Proj}_{\mathbf{v}_2}(\mathbf{x}) + \mathbf{Proj}_{\mathbf{v}_3}(\mathbf{x})$$

with

$$\mathbf{Proj}_{\mathbf{v}_i}(\mathbf{x}) = \left(\frac{\mathbf{x} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}\right)\mathbf{v}_i.$$

Conveniently, in this case all $\mathbf{v}_i \cdot \mathbf{v}_i$'s are equal to 9, and of course $\mathbf{e}_j \cdot \mathbf{v}_i$ is the jth entry of $\mathbf{v}_i$, so the coefficients of $\mathbf{v}_i$ for each $\mathbf{Proj}_{\mathbf{v}_i}(\mathbf{e}_j)$ is read off from staring at the entries for $\mathbf{v}_i$. This yields:

$$\begin{aligned}
\mathbf{e}_1 &= \frac{1}{9}\mathbf{v}_1 + \frac{2}{9}\mathbf{v}_2 + \frac{-2}{9}\mathbf{v}_3, \\
\mathbf{e}_2 &= \frac{-2}{9}\mathbf{v}_1 + \frac{2}{9}\mathbf{v}_2 + \frac{1}{9}\mathbf{v}_3, \\
\mathbf{e}_3 &= \frac{2}{9}\mathbf{v}_1 + \frac{1}{9}\mathbf{v}_2 + \frac{2}{9}\mathbf{v}_3.
\end{aligned}$$

(c) We need to compute $A^{10}\mathbf{e}_i$ for each $i = 1, 2, 3$. Using the answer to (b) and the relations $A^{10}\mathbf{v}_i = \lambda_i^{10}\mathbf{v}_i$ with the λ_i's all known, we have

$$A^{10}\mathbf{e}_1 = \frac{27^{10}}{9}\mathbf{v}_1 + \frac{2 \cdot (-18)^{10}}{9}\mathbf{v}_2 + 0\mathbf{v}_3 = 3^{28}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} + 2^{11}3^{18}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix},$$

$$A^{10}\mathbf{e}_2 = \frac{-2 \cdot 27^{10}}{9}\mathbf{v}_1 + \frac{2 \cdot (-18)^{10}}{9}\mathbf{v}_2 + 0\mathbf{v}_3 = -2 \cdot 3^{28}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} + 2^{11}3^{18}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix},$$

and

$$A^{10}\mathbf{e}_3 = \frac{2 \cdot 27^{10}}{9}\mathbf{v}_1 + \frac{(-18)^{10}}{9}\mathbf{v}_2 + 0\mathbf{v}_3 = 2 \cdot 3^{28}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} + 2^{10}3^{18}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

Putting everything together, we have

$$A^{10} = \begin{bmatrix} 3^{28} + 2^{12}3^{18} & -2 \cdot 3^{28} + 2^{12}3^{18} & 2 \cdot 3^{28} + 2^{11}3^{28} \\ -2 \cdot 3^{28} + 2^{12}3^{18} & 2^2 3^{28} + 2^{12}3^{18} & -2^2 3^{28} + 2^{11}3^{18} \\ 2 \cdot 3^{28} + 2^{11}3^{18} & -2^2 3^{28} + 2^{11}3^{18} & 2^2 3^{28} + 2^{10}3^{18} \end{bmatrix}.$$

As expected, this is symmetric (as we know it had to be, due to symmetry of A and the behavior of transpose on matrix products).

(d) For the second approach, we will work out the formula $A^{10} = QD^{10}Q^\top$. This says

$$\begin{aligned}
A^{10} &= \begin{bmatrix} 1/3 & 2/3 & -2/3 \\ -2/3 & 2/3 & 1/3 \\ 2/3 & 1/3 & 2/3 \end{bmatrix}\begin{bmatrix} 27^{10} & 0 & 0 \\ 0 & (-18)^{10} & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \end{bmatrix} \\
&= \begin{bmatrix} 1/3 & 2/3 & -2/3 \\ -2/3 & 2/3 & 1/3 \\ 2/3 & 1/3 & 2/3 \end{bmatrix}\begin{bmatrix} 3^{30} & 0 & 0 \\ 0 & 2^{10}3^{20} & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 & -2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}\begin{bmatrix} 3^{28} & 0 & 0 \\ 0 & 2^{10}3^{18} & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & -2 & 2 \\ 2 & 2 & 1 \\ -2 & 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 & -2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}\begin{bmatrix} 3^{28} & -2 \cdot 3^{28} & 2 \cdot 3^{28} \\ 2^{11}3^{18} & 2^{11}3^{18} & 2^{10}3^{18} \\ 0 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 3^{28} + 2^{12}3^{18} & -2 \cdot 3^{28} + 2^{12}3^{18} & 2 \cdot 3^{28} + 2^{11}3^{28} \\ -2 \cdot 3^{28} + 2^{12}3^{18} & 2^2 3^{28} + 2^{12}3^{18} & -2^2 3^{28} + 2^{11}3^{18} \\ 2 \cdot 3^{28} + 2^{11}3^{18} & -2^2 3^{28} + 2^{11}3^{18} & 2^2 3^{28} + 2^{10}3^{18} \end{bmatrix}.
\end{aligned}$$

Indeed this agrees with the answer to (c). $\Diamond$

11. (a) Since $\begin{bmatrix}1\\1\\1\end{bmatrix}$ is a scalar multiple of $\begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}$, we verify instead that $A\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}$:

$$M\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}a+b+c\\b+d+e\\c+e+f\end{bmatrix}.$$

But we assumed that $a+b+c$, $b+d+e$, and $c+e+f$ are all equal to 1, so the right side is indeed equal to $\begin{bmatrix}1\\1\\1\end{bmatrix}$.

(b) If the hypotheses of Proposition 24.4.2 are satisfied with the dominant eigenvalue being 1 and $\mathbf{w} = \begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}$, we have the approximation

$$M^k \approx \frac{1^k}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}\mathbf{w}^\top = \frac{1}{3/9}\begin{bmatrix}1/9 & 1/9 & 1/9\\1/9 & 1/9 & 1/9\\1/9 & 1/9 & 1/9\end{bmatrix} = \begin{bmatrix}1/3 & 1/3 & 1/3\\1/3 & 1/3 & 1/3\\1/3 & 1/3 & 1/3\end{bmatrix}$$

for large k. If the starting populations are given by the vector $\begin{bmatrix}P_A\\P_B\\P_C\end{bmatrix}$, then the populations after k years satisfy the following approximation:

$$M^k\begin{bmatrix}P_A\\P_B\\P_C\end{bmatrix} \approx \begin{bmatrix}1/3 & 1/3 & 1/3\\1/3 & 1/3 & 1/3\\1/3 & 1/3 & 1/3\end{bmatrix}\begin{bmatrix}P_A\\P_B\\P_C\end{bmatrix} = \frac{P_A+P_B+P_C}{3}\begin{bmatrix}1\\1\\1\end{bmatrix},$$

as we claimed.

(c) We compute $M\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\1\end{bmatrix}$, so $\begin{bmatrix}0\\0\\1\end{bmatrix}$ is an eigenvector with eigenvalue 1. Similarly, $M\begin{bmatrix}1\\-1\\0\end{bmatrix} = \begin{bmatrix}-1\\1\\0\end{bmatrix}$, so $\begin{bmatrix}1\\-1\\0\end{bmatrix}$ is an eigenvector with eigenvalue -1. Since 1 is a repeated eigenvalue (and -1 is an eigenvalue of the same magnitude), the hypotheses of Proposition 24.4.2 are not satisfied. (Regarding the "interpretation," a possible scenario is there are three sub-populations of birds, one of which remains forever on island C, while the other two swap between islands A and B.) $\Diamond$

12. (a) Plugging $f(x,t) = A\sin(x-ct) + B\sin(x+ct)$ into the left side of the wave equation yields

$$\frac{\partial^2 f}{\partial t^2} - c^2\frac{\partial^2 f}{\partial x^2} = A(\sin(x-ct))_{tt} + B(\sin(x+ct))_{tt} - c^2(A(\sin(x-ct))_{xx} + B(\sin(x+ct))_{xx}),$$

so we need to compute these second partials.

We compute that $(\sin(x-ct))_x = \cos(x-ct)$, so $(\sin(x-ct))_{xx} = -\sin(x-ct)$, and similarly $(\sin(x+ct))_{xx} = -\sin(x+ct)$. On the other hand $(\sin(x-ct))_t = -c\cos(x-ct)$, so $(\sin(x-ct))_{tt} = -c^2\sin(x-ct)$, and similarly $(\sin(x+ct))_{tt} = -c^2\sin(x+ct)$ (the sign arising from the derivative of $\cos$ being $-\sin$).

Plugging these into our expression for the left side of the wave equation yields

$$A(-c^2\sin(x-ct)) + B(-c^2\sin(x+ct)) - c^2(-A\sin(x-ct) - B\sin(x+ct)).$$

Combining terms for a common trigonometric function, this can be written as

$$(-Ac^2 + c^2A)\sin(x-ct) + (-Bc^2 + c^2B)\sin(x+ct)$$

that in turn vanishes (by cancellation in the coefficients).

(b) Similarly to (a), we compute that if we plug $f(x,t) = Ah(x-ct) + Bh(x+ct)$ into the left side of the wave equation then we get

$$\frac{\partial^2 f}{\partial t^2} - c^2\frac{\partial^2 f}{\partial x^2} = A(h(x-ct))_{tt} + B(h(x+ct))_{tt} - c^2(A(h(x-ct))_{xx} + B(h(x+ct))_{xx}),$$

so we need to compute these second partials.

Similarly to (a) we obtain $(h(x\pm ct))_{xx} = h''(x\pm ct)$ and $(h(x\pm ct))_{tt} = c^2h''(x\pm ct)$. (For $h = \sin$ we have $h'' = -\cos$.) Hence, our expression for the left side of the wave equation becomes

$$A(c^2h''(x-ct)) + B(c^2h''(x+ct)) - c^2(Ah''(x-ct) + Bh''(x+ct)),$$

and combining terms yields

$$(Ac^2 - c^2A)h''(x-ct) + (Bc^2 - c^2B)h''(x+ct),$$

which vanishes for the same cancellation reason as in (a).

(c) By the identity $\cos(x\pm y) = \cos(x)\cos(y) \mp \sin(x)\sin(y)$, we have

$$\cos(x\pm ct) = \cos x\cos(ct) \mp \sin x\sin(ct).$$

Hence, $\frac{1}{2}\cos(x-ct) - \frac{1}{2}\cos(x+ct)$ is equal to

$$\frac{1}{2}(\cos(x)\cos(ct) + \sin(x)\sin(ct)) - \frac{1}{2}(\cos(x)\cos(ct) - \sin(x)\sin(ct)),$$

and in this expression the cos terms cancel out whereas the sin terms combine to yield $\sin(x)\sin(ct)$ as desired. $\diamondsuit$

13. (a) Using the single-variable Chain Rule in each variable separately, we have

$$f_x = (x^2+y)^{-1}(2x) = \frac{2x}{x^2+y},\quad f_y = (x^2+y)^{-1} = \frac{1}{x^2+y}.$$

From these expressions we obtain the second partial derivatives via the quotient rule:

$$f_{xx} = \frac{(x^2+y)2 - 2x(2x)}{(x^2+y)^2} = \frac{-2x^2+2y}{(x^2+y)^2},\quad f_{xy} = \frac{-2x}{(x^2+y)^2},\quad f_{yy} = \frac{-1}{(x^2+y)^2}.$$

Putting it all together, we have

$$(\nabla f)(x,y) = \begin{bmatrix} 2x/(x^2+y) \\ 1/(x^2+y) \end{bmatrix},\quad (\mathrm{H}f)(x,y) = \begin{bmatrix} 2(-x^2+y)/(x^2+y)^2 & -2x/(x^2+y)^2 \\ -2x/(x^2+y)^2 & -1/(x^2+y)^2 \end{bmatrix}.$$

(b) Since $f(1,0) = \ln(1) = 0$, $(\nabla f)(1,0) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, and $(\mathrm{H}f)(1,0) = \begin{bmatrix} -2 & -2 \\ -2 & -1 \end{bmatrix}$, the quadratic approximation at $(1,0)$ is

$$\begin{aligned} f(1+h,k) &\approx f(1,0) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} h \\ k \end{bmatrix} + \frac{1}{2}\begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} -2 & -2 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \\ &= 2h + k + \frac{1}{2}(-2h^2 - 4hk - k^2) = 2h + k - h^2 - 2hk - (1/2)k^2. \end{aligned}$$

(c) Plugging $(h,k) = (0.1, 0.2)$ into the approximation in (b) yields

$$f(1.1, 0.2) \approx 0.2 + 0.2 - 0.01 - 0.04 - (1/2)(0.04) = 0.4 - 0.05 - 0.02 = 0.4 - 0.07 = 0.33.$$

If we had omitted the Hessian term then the approximation would have been 0.4. On a calculator, the "exact" answer is $\ln((1.1)^2 + 0.2) = \ln(1.41) = 0.3435897\ldots$. Including the Hessian term improves the accuracy. $\diamondsuit$

14. (a) First we compute the Hessian of f symbolically. We have

$$f_x = 6xy - 2y + \frac{1}{\sqrt{x+2y}}, \quad f_y = 3x^2 - 2x + \frac{2}{\sqrt{x+2y}},$$

and then

$$f_{xx} = 6y - \frac{1}{2(x+2y)^{3/2}}, \quad f_{xy} = 6x - 2 - \frac{1}{(x+2y)^{3/2}}, \quad f_{yy} = -\frac{2}{(x+2y)^{3/2}},$$

so

$$(\mathrm{H}f)(x,y) = \begin{bmatrix} 6y - 1/(2(x+2y)^{3/2}) & 6x - 2 - 1/((x+2y)^{3/2}) \\ 6x - 2 - 1/((x+2y)^{3/2}) & -2/(x+2y)^{3/2} \end{bmatrix}.$$

(b) At the point $(-1,1)$ the value of $x + 2y$ is 1, so $\sqrt{x+2y} = 1$ there too. Hence,

$$(\mathrm{H}f)(-1,1) = \begin{bmatrix} 6 - 1/2 & -6 - 2 - 1 \\ -6 - 2 - 1 & -2 \end{bmatrix} = \begin{bmatrix} 11/2 & -9 \\ -9 & -2 \end{bmatrix}.$$

At the point $(1,0)$ the value of $x + 2y$ is 1, so $\sqrt{x+2y} = 1$ there too. Hence,

$$(\mathrm{H}f)(1,0) = \begin{bmatrix} -1/2 & 6 - 2 - 1 \\ 6 - 2 - 1 & -2 \end{bmatrix} = \begin{bmatrix} -1/2 & 3 \\ 3 & -2 \end{bmatrix}.$$

(c) The gradient of f at $(-1,1)$ is the vector

$$(\nabla f)(-1,1) = \begin{bmatrix} f_x(-1,1) \\ f_y(-1,1) \end{bmatrix} = \begin{bmatrix} -6 - 2 + 1 \\ 3 + 2 + 2 \end{bmatrix} = \begin{bmatrix} -7 \\ 7 \end{bmatrix}.$$

Hence, for small h, k we have the quadratic approximation

$$\begin{aligned} f(-1+h, 1+k) &\approx f(-1,1) + \begin{bmatrix} -7 \\ 7 \end{bmatrix} \cdot \begin{bmatrix} h \\ k \end{bmatrix} + \frac{1}{2}\begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} 11/2 & -9 \\ -9 & -2 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \\ &= (3+2+2) - 7h + 7k + \frac{1}{2}\begin{bmatrix} 11h/2 - 9k & -9h - 2k \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \\ &= 7 - 7h + 7k + \frac{1}{2}((11/2)h^2 - 9kh - 9hk - 2k^2) \\ &= 7 - 7h + 7k + (11/4)h^2 - 9hk - k^2. \end{aligned}$$

The gradient of f at $(1,0)$ is the vector

$$(\nabla f)(1,0) = \begin{bmatrix} f_x(1,0) \\ f_y(1,0) \end{bmatrix} = \begin{bmatrix} 1 \\ 3-2+2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Hence, for small h, k we have the quadratic approximation

$$\begin{aligned} f(1+h,k) &\approx f(1,0) + \begin{bmatrix} 1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} h \\ k \end{bmatrix} + \frac{1}{2}\begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} -1/2 & 3 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \\ &= 2 + h + 3k + \frac{1}{2}\begin{bmatrix} -h/2+3k & 3h-2k \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \\ &= 2 + h + 3k + \frac{1}{2}(-h^2/2 + 3kh + 3hk - 2k^2) \\ &= 2 + h + 3k - h^2/4 + 3kh - k^2. \end{aligned}$$

$\Diamond$

15. (a) The associated symmetric 2×2 matrix is $A = \begin{bmatrix} -1 & 6 \\ 6 & -1 \end{bmatrix}$. This has trace -2 and determinant $1 - 36 = -35$, so the characteristic polynomial of A is $\lambda^2 + 2\lambda - 35 = (\lambda - 5)(\lambda + 7)$. Hence, the eigenvalues are $\lambda_1 = 5$ and $\lambda_2 = -7$. To find corresponding eigenvectors we compute

$$A - \lambda_1 \mathrm{I}_2 = A - 5\mathrm{I}_2 = \begin{bmatrix} -6 & 6 \\ 6 & -6 \end{bmatrix}, \quad A - \lambda_2 \mathrm{I}_2 = A + 7\mathrm{I}_2 = \begin{bmatrix} 6 & 6 \\ 6 & 6 \end{bmatrix}$$

whose respective null spaces are spanned by $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Since the eigenvalues have opposite signs, $q(x,y)$ is indefinite. Hence, the level sets $q(x,y) = c$ $(c \neq 0)$ are hyperbolas, with asymptotes "closer" to the line spanned by $\mathbf{w}_1$ since $|\lambda_1| < |\lambda_2|$ (i.e., $5 < 7$). The hyperbolas crossing the line spanned by $\mathbf{w}_1$ are $q(x,y) = c$ with $c > 0$ since $q(\mathbf{w}_1) = q(1,1) = 10 > 0$, and the hyperbolas crossing the line spanned by $\mathbf{w}_2$ are $q(x,y) = c$ with $c < 0$ since $q(\mathbf{w}_2) = q(1,-1) = -14 < 0$. The precise picture is as follows (but only the qualitative aspects are expected):

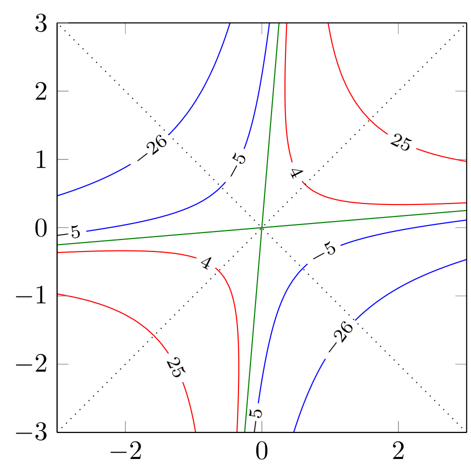

FIGURE 1. Contour plot for $q(x,y) = -x^2 + 12xy - y^2$ with dotted eigenlines and green asymptotes, and $q(x,y) = c$ in red for $c > 0$ and in blue for $c < 0$.

(b) The associated symmetric 2×2 matrix is $A = \begin{bmatrix} 12 & -3 \\ -3 & 4 \end{bmatrix}$. This has trace 16 and determinant $48 - 9 = 39$, so the characteristic polynomial of A is $\lambda^2 - 16\lambda + 39 = (\lambda - 3)(\lambda - 13)$. Hence, the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 13$. To find corresponding eigenvectors we compute

$$A - \lambda_1 \mathrm{I}_2 = A - 3\mathrm{I}_2 = \begin{bmatrix} 9 & -3 \\ -3 & 1 \end{bmatrix}, \quad A - \lambda_2 \mathrm{I}_2 = A - 13\mathrm{I}_2 = \begin{bmatrix} -1 & -3 \\ -3 & -9 \end{bmatrix}$$

whose respective null spaces are spanned by $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

Since the eigenvalues are both positive, $q(x, y)$ is positive-definite. Hence, the level sets $q(x, y) = c$ (with $c > 0$) are ellipses. These ellipses are longer along the line spanned by $\mathbf{w}_1$ since $|\lambda_1| < |\lambda_2|$ (i.e., $3 < 13$). The precise picture is as follows (but only the qualitative aspects are expected):

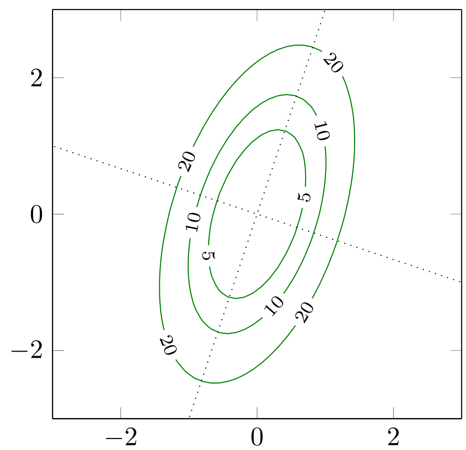

FIGURE 2. Contour plot for $q(x, y) = 12x^2 - 6xy + 4y^2$ with dotted eigenlines; the ellipses are longer along the direction of the 3-eigenline

◊

16. (a) We computed in Exercise I.2.3 that $(\mathrm{H}f)(x, y) = \begin{bmatrix} 16 + 24x & 6 + 6y \\ 6 + 6y & 6x \end{bmatrix}$. Thus $(\mathrm{H}f)(0, 0) = \begin{bmatrix} 16 & 6 \\ 6 & 0 \end{bmatrix}$, $(\mathrm{H}f)(0, -2) = \begin{bmatrix} 16 & -6 \\ -6 & 0 \end{bmatrix}$, $(\mathrm{H}f)(-3/2, -1) = \begin{bmatrix} -20 & 0 \\ 0 & -9 \end{bmatrix}$, and $(\mathrm{H}f)(1/6, -1) = \begin{bmatrix} 20 & 0 \\ 0 & 1 \end{bmatrix}$.

Since a diagonal matrix has its diagonal entries as its eigenvalues, we see that $(-3/2, -1)$ is a local maximum (both eigenvalues are negative) and $(1/6, -1)$ is a local minimum (both eigenvalues are positive). For $(0, 0)$, note that the determinant is $(16)(0) - (6)(6) = -36$, so the eigenvalues are nonzero with opposite signs and so this critical point is a saddle point. Finally, at $(0, -2)$, the determinant is $(16)(0) - (-6)(-6) = -36$ again, so once more the critical point is a saddle point.

(b) The Hessians at $(-3/2, -1)$ and $(1/6, -1)$ are diagonal, so eigenvectors for each are $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Being definite for each, the level sets are approximately ellipses with axes along the coordinate directions. (More specifically, these are respectively approximated by $-20{t_1}^2 - 9{t_2}^2 = c$ with $c < 0$, shifted to be centered at $(-3/2, -1)$, and $20{t_1}^2 + {t_2}^2 = c$ with $c > 0$, shifted to be centered at $(1/6, -1)$; such explicitness doesn't matter for your answer.)

In both cases the ellipse is somewhat longer vertically than horizontally since in both cases the eigenvalue corresponding to the vertical eigenvector is smaller in absolute value ($|-9| < |-20|$ and $|1| < |20|$).

A very detailed picture of the contour plot near both critical points is given in Figure 3 below (the green vertical line through $x = 0$ expresses that x is a factor of $f(x, y)$). You're not expected to come up with anything as precise as this, just to recognize from the Hessians as above that the contour plots consist of approximate nested ellipses around the critical points with approximate axis directions determined by the eigenvalues and eigenvector directions of the Hessian as discussed above and that the longer axis along the direction of the line through $\mathbf{w}_2$ in each case. The "zooming in" around $(1/6, -1)$, which we provide as a bonus, illustrates that the scale on which the quadratic approximation kicks in could be very tiny!

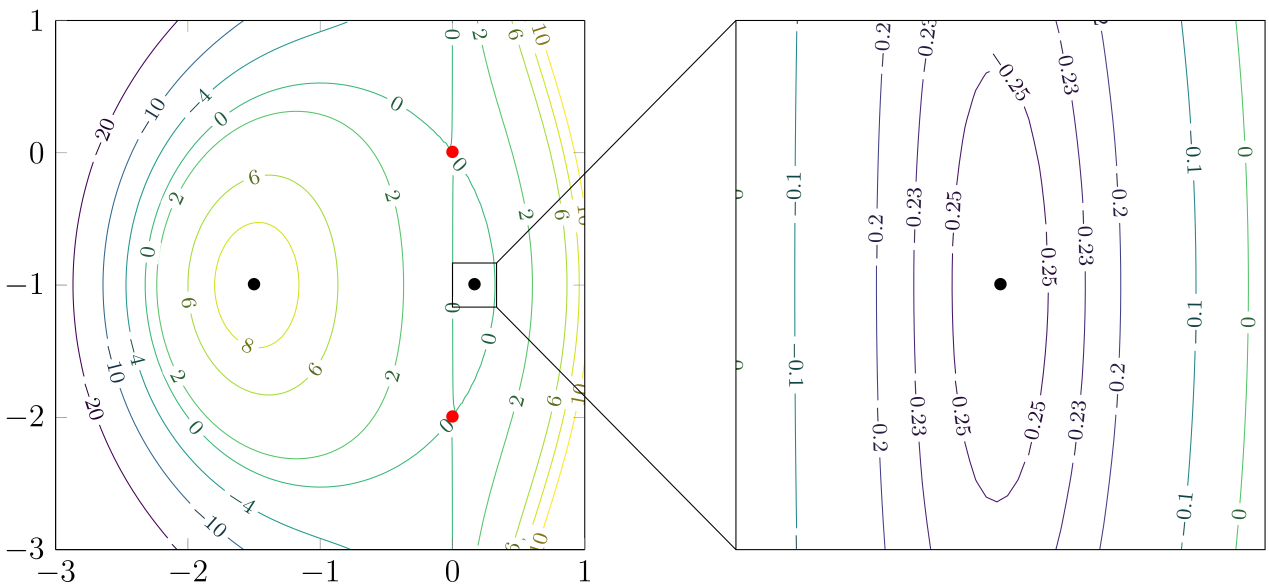

FIGURE 3. Contour plot of $f(x, y) = 8x^2 + 6xy + 4x^3 + 3xy^2$, with approximate nested ellipses around the critical points $(-3/2, -1)$ and $(1/6, -1)$ (zoomed in)

$\diamondsuit$

17. (a) The gradient is

$$\nabla f = \begin{bmatrix} 6x^2 + y^2 + 10x \\ 2xy + 2y \end{bmatrix} = \begin{bmatrix} 6x^2 + y^2 + 10x \\ 2y(x+1) \end{bmatrix}.$$

To find the critical points, we need to solve simultaneously

$$6x^2 + y^2 + 10x = 0, \qquad 2y(x+1) = 0.$$

Solving the second equation, we have either $x = -1$ or $y = 0$. Plugging these conditions into the first equation, if $x = -1$ then the first equation says $y^2 - 4 = 0$, so $y = \pm 2$. On the other hand, if $y = 0$ then the first equation says $6x^2 + 10x = 0$, so $x = 0$ or $x = -5/3$. So the critical points are $(-1, 2)$, $(-1, -2)$, $(0, 0)$, and $(-5/3, 0)$.

(b) We compute the Hessian of f symbolically to be

$$(\mathrm{H}f)(x, y) = \begin{bmatrix} 12x + 10 & 2y \\ 2y & 2x + 2 \end{bmatrix}.$$

Thus,

$$(\mathrm{H}f)(-1,2) = \begin{bmatrix} -2 & 4 \\ 4 & 0 \end{bmatrix}$$

Since the determinant is $(-2)(0) - (4)(4) = -16 < 0$, $(-1,2)$ must be a saddle point. The Hessian

$$(\mathrm{H}f)(-1,-2) = \begin{bmatrix} -2 & -4 \\ -4 & 0 \end{bmatrix}$$

has determinant $(-2)(0) - (-4)(-4) = -16 < 0$, so $(-1,2)$ is also a saddle point.

The Hessian

$$(\mathrm{H}f)(0,0) = \begin{bmatrix} 10 & 0 \\ 0 & 2 \end{bmatrix}$$

is diagonal, so its diagonal entries are its eigenvalues. These are both positive, so $(0,0)$ is a local minimum. The Hessian

$$(\mathrm{H}f)(-5/3,0) = \begin{bmatrix} -10 & 0 \\ 0 & -4/3 \end{bmatrix}$$

is also diagonal, with negative eigenvalues, so $(-5/3,0)$ is a local maximum.

(c) The Hessians at $(0,0)$ and $(-5/3,0)$ are diagonal, so eigenvectors for each are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Being definite for each, the level sets are approximately ellipses with axes along the coordinate directions. (More specifically, these are respectively approximated by $10{t_1}^2 + 2{t_2}^2 = c$ with $c > 0$ that is already centered at $(0,0)$ and $-10{t_1}^2 - 4/3{t_2}^2 = c$ with $c < 0$ shifted to be centered at $(-5/3,0)$; you are not expected to provide such explicitness.)

In both cases the ellipse is longer vertically than horizontally since the standard basis vectors are eigenvectors in these cases (due to the Hessian at each critical point being diagonal) with the eigenvalue for the vertical eigenvector having smaller absolute value in each case ($2 < 10$ and $4/3 < 10$).

A very detailed picture of the contour plot near these two critical points is given in Figure 4 below. You're not expected to come up with anything as precise and detailed as this, just to recognize from the Hessians that the contour plot consists of approximate nested ellipses around each critical point $(0,0)$ and $(-5/3,0)$ with horizontal and vertical axis directions (due to those being the directions of the eigenvectors of the diagonal Hessians) and with the ellipses being longer in the vertical direction.

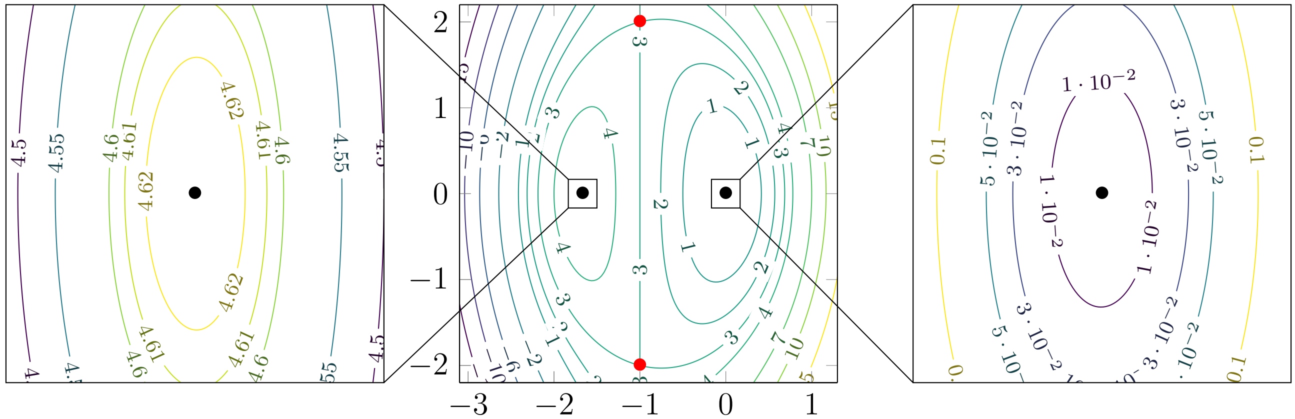

FIGURE 4. Contour plot of $f(x,y) = 2x^3 + xy^2 + 5x^2 + y^2$, with approximate nested ellipses around the critical points $(-5/3, 0)$ and $(0,0)$ (zoomed in)

$\Diamond$

18. (a) The gradient of f is $\nabla f = \begin{bmatrix} 2x \\ -2y \end{bmatrix}$ and this vanishes precisely at the origin. The value of f here is $f(0,0) = 0$. This is a saddle point since the Hessian here is $\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$, so it is not even a local extremum (let alone a global one) on the disk. The surface graph of f over the disk in Figure 5 below shows the saddle at the yellow origin.

(b) We calculate that $h(t) = (1 + 2\cos t)^2 - (2\sin t)^2 = 1 + 4\cos t + 4\cos^2 t - 4\sin^2 t$. Hence, $h'(t) = -4\sin t - 8\cos t \sin t - 8\sin t \cos t = -4\sin t(1 + 4\cos t)$.. This derivative vanishes precisely when one of the factors vanishes; i.e., when $t = 0, \pi$ (this is where $\sin t$ vanishes) and when $\cos t = -1/4$. This happens at two separate angles t, let's call them t_0 and t_1, corresponding to points where the unit circle meets the line $x = -1/4$ (so the corresponding y-coordinates, which are $\sin t_0$ and $\sin t_1$, are negatives of each other); let's say t_0 corresponds to the point in the second quadrant and t_1 corresponds to the point in the third quadrant.

It isn't actually necessary to determine t_0 and t_1 numerically, though if you're curious then you can use a calculator to see that $t_0 \approx 1.823$ radians, $t_1 \approx 4.460$ radians. What matters is that since $\sin^2 t + \cos^2 t = 1$, we know the values $\sin t_0 = \sqrt{1 - (-1/4)^2} = \sqrt{15}/4$ and $\sin t_1 = -\sqrt{15}/4$. At these various points,

$$h(0) = 1+4+4 = 9,\ h(\pi) = 1-4+4 = 1,\ h(t_0) = 1-1+4((1/16)-(15/16)) = -7/2,\ h(t_1) = -7/2.$$

(c) The only candidate for where the maximum might occur inside the disk is the origin, and we have already seen that this cannot be the maximum. The only other possibility is that the maximum of f on the entire region is its maximum on the boundary, which occurs at $(3,0)$ with the value 9 by comparing the values of h computed at the end of (b) to find the biggest.

(d) At the two points $(1/2, \sqrt{15}/2)$ and $(1/2, -\sqrt{15}/2)$ (where $t = t_0, t_1$), the function f equals $-7/2$ and this is the most negative that f can become anywhere in the disk due to comparing the values of h found at the end of (b). At the remaining critical point $(-1,0)$ on the boundary, where $t = \pi$, we have $f(-1,0) = 1$. We will show that this point is a local maximum, as the surface graph in Figure 5 illustrates.

To justify the local maximum property for $x^2 - y^2$ on the disk of interest at the leftmost boundary point $(-1,0)$, we observe that points near $(-1,0)$ in the disk have the form (a,b) where $b \approx 0$ and a is only slightly larger than (or equal to) -1, so $-1 \le a \le 0$. Hence, $a^2 - b^2 \le 1 - b^2 \le 1$

since $a^2 \leq 1$ and $b^2 \geq 0$, so f has values at most $1 = f(-1,0)$ near $(-1,0)$ in the disk. (You could alternatively try using the linear approximation $f(-1+h,k) \approx f(-1,0)+(\nabla f)(-1,0)\cdot \begin{bmatrix} h \\ k \end{bmatrix} = 1-2h$ with small $h \geq 0$ and k near 0, but that is just a plausibility argument since it gets mired in issues around the magnitude of the error in the approximation – is it of smaller size than $2h$? – to nail down that really $f \leq 1$ near $(-1,0)$ on the disk.)

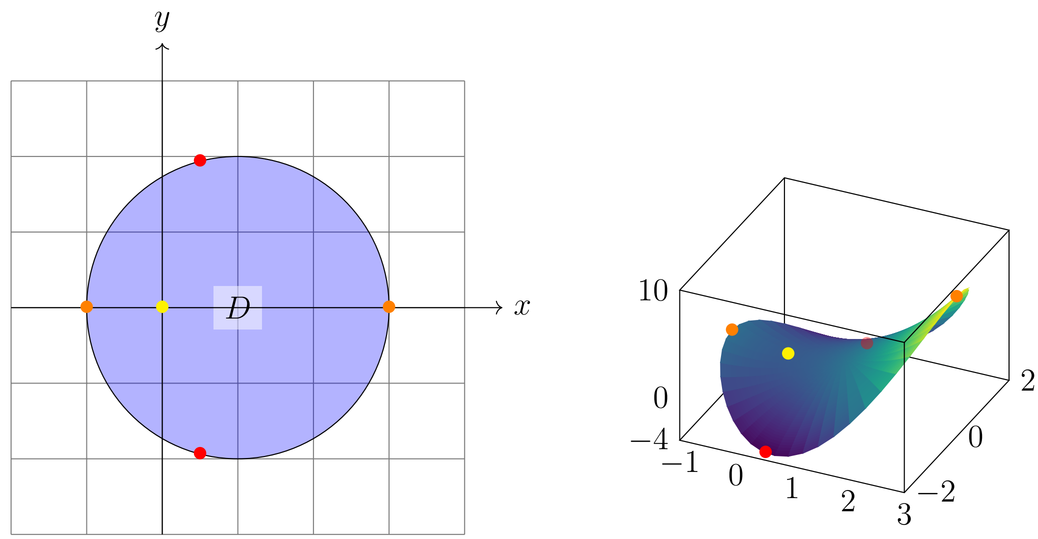

FIGURE 5. The graph of $x^2 - y^2$ over the blue disk $(x-1)^2 + y^2 \leq 4$. The red boundary points on $x = 1/2$ are the global minima, the yellow origin is a saddle point, the right orange dot is the global maximum, and the left orange dot is a local maximum.

$\diamondsuit$