1. (a) Since there is no movement other than the indicated one from I_1 to I_2 in the spring and from I_2 to I_1 in the fall, the spring transition matrix is $\begin{bmatrix} 0.9 & 0 \\ 0.1 & 1 \end{bmatrix}$, or B, and the fall transition matrix is $\begin{bmatrix} 1 & 0.1 \\ 0 & 0.9 \end{bmatrix}$, or D. Thus, $B\mathbf{p}$ is the population vector at the end of the spring, and D carries the spring population vector to the fall population vector. That is, the population vector at the end of the fall is $D(B\mathbf{p}) = (DB)\mathbf{p}$. Hence, we have $M = DB$. Multiplying these two matrices gives $M = \begin{bmatrix} .91 & .1 \\ .09 & .9 \end{bmatrix}$.

   (b) This part concerns the passage of two full years. Since the transition matrix for one full year is DB, for two full years it is $N = (DB)^2$. Indeed, beginning with a population vector $\mathbf{p}$ at the start of a year, after one full year it has become $(DB)\mathbf{p}$. Thus, after one more full year (so two full years from the start) we arrive at the population vector $(DB)((DB)\mathbf{p}) = (DB)^2\mathbf{p}$ (using the associative law). Hence, $N = (DB)^2$, which is $\begin{bmatrix} .8371 & .181 \\ .1629 & .819 \end{bmatrix}$.

   (c) In the long run, 52.63% of the birds are on island I_1 and 47.37% of the birds are on island I_2 at the end of every year. Birds keep moving around every year, but the proportions on each at the end of the year stabilize at these values.

◇

2. (a) The given rules tell us:

$$F_{i+1} = (0.9)F_i + (0.05)N_i, \;\; T_{i+1} = (0.9)T_i + (0.1)N_i, \;\; N_{i+1} = (0.1)F_i + (0.1)T_i + (0.85)N_i.$$

   Hence, we use

$$A = \begin{bmatrix} 0.9 & 0 & 0.05 \\ 0 & 0.9 & 0.1 \\ 0.1 & 0.1 & 0.85 \end{bmatrix}.$$

   (b) We have $M = A^{10}$, by the same iteration arguments as given in the course text for many circumstances (bird migration, gambler's ruin, etc.).

◇

3. (a) Both $\mathbf{v}$ and $\mathbf{w}$ are nonzero and not scalar multiples of each other, so they're linearly independent. Being 2 such vectors in the 2-dimensional $\mathbf{R}^2$, they must be a basis of $\mathbf{R}^2$. By direct calculation,

$$A\mathbf{v} = \begin{bmatrix} 1/2 & 3/4 \\ 1/2 & 1/4 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 3/2+3/2 \\ 3/2+1/2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \mathbf{v}$$

   and

$$A\mathbf{w} = \begin{bmatrix} 1/2 & 3/4 \\ 1/2 & 1/4 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/2+3/4 \\ -1/2+1/4 \end{bmatrix} = \begin{bmatrix} 1/4 \\ -1/4 \end{bmatrix} = -(1/4)\begin{bmatrix} -1 \\ 1 \end{bmatrix} = -(1/4)\mathbf{w}.$$

   (b) We compute via linearity that

$$A\mathbf{x} = \alpha(A\mathbf{v}) + \beta(A\mathbf{w}) = \alpha\mathbf{v} + \beta((-1/4)\mathbf{w}) = \alpha\mathbf{v} - (\beta/4)\mathbf{w},$$

$$\begin{aligned} A^2\mathbf{x} = A(A\mathbf{x}) = A(\alpha\mathbf{v} - (\beta/4)\mathbf{w}) &= \alpha(A\mathbf{v}) - (\beta/4)(A\mathbf{w}) \\ &= \alpha\mathbf{v} - (\beta/4)(-(1/4)\mathbf{w}) \\ &= \alpha\mathbf{v} + (\beta/16)\mathbf{w}, \end{aligned}$$

   and

$$\begin{aligned} A^3\mathbf{x} = A(A^2\mathbf{x}) = A(\alpha\mathbf{v} + (\beta/16)\mathbf{w}) &= \alpha(A\mathbf{v}) + (\beta/16)(A\mathbf{w}) \\ &= \alpha\mathbf{v} + (\beta/16)(-(1/4)\mathbf{w}) \\ &= \alpha\mathbf{v} - (\beta/64)\mathbf{w}. \end{aligned}$$

(c) The pattern in the calculations in (b) suggests that $A^m\mathbf{v} = \alpha\mathbf{v} \pm (\beta/4^m)\mathbf{w}$ for $m = 1, 2, 3, \ldots$. Working out a couple more fits the pattern, and we don't expect you to offer a more complete justification beyond observing this pattern for some small m. As m grows, the term $\alpha\mathbf{v}$ is unaffected but the coefficient $\beta/4^m$ approach 0 (since β is a constant having nothing to do with m). Hence, for large m the $\mathbf{w}$-term becomes negligible, leaving us with $A^m\mathbf{v} \approx \alpha\mathbf{v}$ for large m.

◇

4. (a) We calculate

$$\mathbf{v}_{n+1} = \begin{bmatrix} a_{n+2} \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} a_{n+1} - a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix} = A\mathbf{v}_n.$$

Also by direct multiplication, $A^2 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$, so

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

Hence, $A^6 = A^3A^3 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

(b) Since $\mathbf{v}_{n+1} = A\mathbf{v}_n$, feeding this into itself more times gives $\mathbf{v}_{n+2} = A\mathbf{v}_{n+1} = A(A\mathbf{v}_n) = A^2\mathbf{v}_n$ and more generally $\mathbf{v}_{n+k} = A^k\mathbf{v}_n$ for any $k \geq 1$. Setting $k = 6$ gives $\mathbf{v}_{n+6} = A^6\mathbf{v}_n = \mathrm{I}_2\mathbf{v}_n = \mathbf{v}_n$ for all n. Comparing bottom entries in these vectors gives $a_{n+6} = a_n$ for all n. We can also verify this final equality from the recurrence definition:

$$\begin{aligned} a_{n+6} = a_{n+5} - a_{n+4} = (a_{n+4} - a_{n+3}) - a_{n+4} = -a_{n+3} &= -(a_{n+2} - a_{n+1}) \\ &= -a_{n+2} + a_{n+1} \\ &= -(a_{n+1} - a_n) + a_{n+1} \\ &= -a_{n+1} + a_n + a_{n+1} \\ &= a_n. \end{aligned}$$

◇

5. We have

$$(Df)(x, y) = \begin{bmatrix} 3x^2y^2 & 2x^3y - 1 \\ y^3 - 1 & 3xy^2 \end{bmatrix},$$

so by the Chain Rule

$$D(f \circ f)(1, 1) = (Df)(f(1, 1))\,(Df)(1, 1) = (Df)(0, 0)\,(Df)(1, 1) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & -3 \\ -3 & -1 \end{bmatrix}.$$

Hence, the linear approximation is

$$\begin{aligned} (f \circ f)(1 + h, 1 + k) \approx (f \circ f)(1, 1) + (D(f \circ f))(1, 1) \begin{bmatrix} h \\ k \end{bmatrix} &= f(0, 0) + \begin{bmatrix} 0 & -3 \\ -3 & -1 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -3k \\ -3h - k \end{bmatrix} \\ &= \begin{bmatrix} -3k \\ -3h - k \end{bmatrix}. \end{aligned}$$

◇

6. (a) By definition, $(D_{\mathbf{v}}f)(\mathbf{a})$ is the entry in the 1×1 matrix

$$D(f \circ g)(0) = (Df)(g(0))\,(Dg)(0) = (Df)(\mathbf{a})\,(Dg)(0),$$

and the $n \times 1$ matrix (a "column vector") $(Dg)(0)$ is exactly $\mathbf{v}$ since we calculate its ith entry to be the derivative $g_i'(0)$ of the ith component function $g_i(t) = a_i + tv_i$ (and the t-derivative g_i' is the constant function with value v_i). In other words, $(D_{\mathbf{v}}f)(\mathbf{a}) = (Df)(\mathbf{a})\mathbf{v}$ when viewing the column vector $\mathbf{v}$ as a an $n \times 1$ matrix, and this in turn is exactly the matrix-vector product.

(b) For the specific f we have, $(Df)(x,y) = \begin{bmatrix} \pi y\cos(\pi xy) & \pi x\cos(\pi xy)\end{bmatrix}$. For $\mathbf{v} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2}\end{bmatrix}$, the hilliness when moving directly northeast amounts to looking at the height function f on the line $\mathbf{a} + t\mathbf{v}$. Hence, the slope is given by the t-derivative of this height at $t = 0$, which is

$$(D_{\mathbf{v}}f)(1,2) = ((Df)(1,2)) \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2}\end{bmatrix} = \begin{bmatrix} 2\pi & \pi \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2}\end{bmatrix} = 3\pi/\sqrt{2}.$$

◇

7. (a) This level set is the unit circle, parametrized by $g(t) = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}$. (Other answers are $(\sin t, \cos t)$, the somewhat exotic $(\cos(e^t), -\sin(e^t))$, and much else.)

(b) Differentiation yields

$$0 = \frac{d}{dt}((f \circ g)(t)) = (Df)(g(t))\,(Dg)(t) = ((\nabla f)(g(t))) \cdot g'(t),$$

where the final equality means that the scalar dot product on the right is the entry in the 1×1 matrix product on the left (as we verify by computing the 1×1 matrix product in terms of the entries of the row matrix $(Df)(g(t))$ and the column matrix $(Dg)(t)$). Hence, we see that these two vectors are perpendicular.

(c) Since $g'(t)$ is a nonzero vector pointing along the direction tangent to the curve $f = c$ at $g(t)$, the vector $(\nabla f)(g(t))$ perpendicular to $g'(t)$ must be perpendicular to the direction of the tangent line to $f = c$ at the point $g(t)$. Being orthogonal to the tangent direction at each point is a reasonable sense in which to say that the gradient is perpendicular to the level curve.

◇

8. (a) We have $h = f \circ F$ for $F(r,\theta) = (r\cos\theta, r\sin\theta)$, so F has component functions $F_1(r,\theta) = r\cos\theta$ and $F_2(r,\theta) = r\sin\theta$. Hence, by the Chain Rule

$$(Dh)(r,\theta) = (Df)(F(r,\theta))(DF)(r,\theta) = (Df)(x,y) \begin{bmatrix} (F_1)_r & (F_1)_\theta \\ (F_2)_r & (F_2)_\theta \end{bmatrix} = \begin{bmatrix} f_x & f_y \end{bmatrix} \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

Multiplying the matrices gives

$$\begin{bmatrix} h_r & h_\theta \end{bmatrix} = \begin{bmatrix} \cos\theta\, f_x + \sin\theta\, f_y & -r\sin\theta\, f_x + r\cos\theta\, f_y \end{bmatrix}.$$

Equating entries on the two sides gives the desired formulas for h_r and h_θ.
Alternatively, via (17.1.6) with $x(r,\theta) = r\cos\theta$ and $y(r,\theta) = r\sin\theta$ we have

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = \cos\theta \frac{\partial f}{\partial x} + \sin\theta \frac{\partial f}{\partial y}$$

and

$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} = -r\sin\theta\, \frac{\partial f}{\partial x} + r\cos\theta\, \frac{\partial f}{\partial y}.$$

In such equations, the left sides are informal notation for what are really $\partial h/\partial r$ and $\partial h/\partial \theta$ respectively, so we again get the result we wanted.

(b) From (a) we have the pair of equations

$$\begin{aligned} \frac{\partial h}{\partial \theta} &= -r\sin\theta\, \frac{\partial f}{\partial x} + r\cos\theta\, \frac{\partial f}{\partial y} \\ \frac{\partial h}{\partial r} &= \cos\theta\, \frac{\partial f}{\partial x} + \sin\theta\, \frac{\partial f}{\partial y}. \end{aligned}$$

We want to unwind this to express the x and y partials of f in terms of the partials of h.
Let's adapt the method used to solve "two (linear) equations in two unknowns": multiply by suitable expressions to then combine the equations and make f_x disappear (since we want to describe f_y). To make f_x disappear, let's multiply the first equation by $\cos\theta$, the second by $r\sin\theta$, and then add the equations. This gives

$$\cos\theta\, h_\theta + r\sin\theta\, h_r = (r\cos^2\theta + r\sin^2\theta) f_y = r f_y.$$

Dividing by r (as we may do since we're assuming $r > 0$) gives

$$f_y = \frac{\cos\theta}{r} h_\theta + (\sin\theta) h_r.$$

So $g_1(r,\theta) = (\cos\theta)/r$ and $g_2(r,\theta) = \sin\theta$.

◇

9. (a) Much like how $(AB)^{-1} = B^{-1}A^{-1}$, we guess that the inverse of ABC will be obtained by multiplying the inverses for A, B, and C in reverse order (for another explanation, we could use reasoning similar to that which appears at the end of Example 18.4.4). That is, we want to show that $C^{-1}B^{-1}A^{-1}$ is an inverse to ABC. We just have to check that it satisfies the two conditions in the definition of inverse:

$$\begin{aligned}(ABC)(C^{-1}B^{-1}A^{-1}) &= ABCC^{-1}B^{-1}A^{-1}\\ &= AB\mathrm{I}_n B^{-1}A^{-1}\\ &= ABB^{-1}A^{-1}\\ &= A\mathrm{I}_n A^{-1}\\ &= AA^{-1}\\ &= \mathrm{I}_n\end{aligned}$$

and

$$\begin{aligned}(C^{-1}B^{-1}A^{-1})(ABC) &= C^{-1}B^{-1}A^{-1}ABC\\ &= C^{-1}B^{-1}\mathrm{I}_n BC\\ &= C^{-1}B^{-1}BC\\ &= C^{-1}\mathrm{I}_n C\\ &= C^{-1}C\\ &= \mathrm{I}_n.\end{aligned}$$

(b) If we expand this expression all the way out, we notice that we can usefully regroup the terms (thanks to the "associative" law of multiplication):

$$(ADA^{-1})^{12} = (ADA^{-1})(ADA^{-1})\cdots(ADA^{-1}) = AD(A^{-1}A)D(A^{-1}A)\cdots(A^{-1}A)DA^{-1}.$$

Now, each of the $A^{-1}A$'s that appears within the parentheses is equal to I_n. Thus,

$$(ADA^{-1})^{12} = AD(\mathrm{I}_n)D(\mathrm{I}_n)\cdots(\mathrm{I}_n)DA^{-1} = ADD\cdots DA^{-1} = AD^{12}A^{-1},$$

since there were 12 copies of D in the full expansion and the I_n's have no effect under multiplication.

◇

10. (a) We just multiply:

$$M_a M_{-a} = \begin{bmatrix}1 & a & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \cdot \begin{bmatrix}1 & -a & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & a-a & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} = \mathrm{I}_3.$$

The same computation (but with a and $-a$ swapping roles) shows that $M_{-a}M_a = \mathrm{I}_3$.

(b) We can again just multiply everything out:

$$N_b N_{-b} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & b\\ 0 & 0 & 1\end{bmatrix} \cdot \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & -b\\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & b-b\\ 0 & 0 & 1\end{bmatrix} = \mathrm{I}_3,$$

and $N_{-b}N_b = \mathrm{I}_3$ similarly.

(c) Note that

$$\begin{bmatrix}1 & a & 0\\ 0 & 1 & b\\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & b\\ 0 & 0 & 1\end{bmatrix} \cdot \begin{bmatrix}1 & a & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} = N_b M_a.$$

(This doesn't work if we multiply N_b and M_a in the other order.) Since N_b and M_a are invertible, so is their product with inverse $(N_b M_a)^{-1} = M_a^{-1} N_b^{-1}$. By parts (a) and (b), this is equal to

$$M_{-a} N_{-b} = \begin{bmatrix}1 & -a & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \cdot \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & -b\\ 0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & -a & ab\\ 0 & 1 & -b\\ 0 & 0 & 1\end{bmatrix}.$$

◇

11. (a) We compute partial derivatives of component functions of $\mathbf{f}$ to obtain

$$(D\mathbf{f})(x, y) = \begin{bmatrix}3x^2 - 3y^2 & -6xy\\ 6xy & 3x^2 - 3y^2\end{bmatrix}.$$

[Just as an aside, let's check when this is invertible even though we don't need such generality for the rest of the problem. This will be invertible if its determinant is non-zero; that is, if $(3x^2 - 3y^2)(3x^2 - 3y^2) + 6xy(6xy) \neq 0$. The left side is equal to $9x^4 - 18x^2y^2 + 9y^4 + 36x^2y^2 = 9x^2 + 18x^2y^2 + 9y^4 = 9(x^2 + y^2)^2$. This vanishes precisely when $x^2 + y^2 = 0$, which happens exactly when $(x, y) = (0, 0)$. Hence, $(D\mathbf{f})(x, y)$ is invertible when $(x, y) \neq (0, 0)$.]

(b) Using the formula from Example 18.2.1, $(D\mathbf{f})(2/3, 1/3) = \begin{bmatrix}1 & -4/3\\ 4/3 & 1\end{bmatrix}$ is invertible with inverse

$$\frac{1}{1(1) - (4/3)(-4/3)} \begin{bmatrix}1 & 4/3\\ -4/3 & 1\end{bmatrix} = \frac{1}{25/9} \begin{bmatrix}1 & 4/3\\ -4/3 & 1\end{bmatrix} = \begin{bmatrix}9/25 & 12/25\\ -12/25 & 9/25\end{bmatrix}.$$

Also, $\mathbf{f}(\mathbf{a}_1) = (-25/27, 11/27)$. Thus,

$$\mathbf{a}_2 = \mathbf{a}_1 - \begin{bmatrix}9/25 & 12/25\\ -12/25 & 9/25\end{bmatrix} \cdot \begin{bmatrix}-25/27\\ 11/27\end{bmatrix} = \begin{bmatrix}2/3\\ 1/3\end{bmatrix} - \begin{bmatrix}-31/225\\ 133/225\end{bmatrix} = \begin{bmatrix}181/225\\ -58/225\end{bmatrix} \approx \begin{bmatrix}0.804\\ -0.258\end{bmatrix}.$$

(c) The point $\mathbf{a}_2$ differs from $\begin{bmatrix}1\\ 0\end{bmatrix}$ by $\begin{bmatrix}44/225\\ 58/225\end{bmatrix} = \frac{2}{225}\begin{bmatrix}22\\ 29\end{bmatrix}$. This difference has length

$$\frac{2}{225}\sqrt{22^2 + 29^2} = \frac{2}{225}\sqrt{1325} = \frac{2}{45}\sqrt{53} \approx 0.32,$$

whereas the distance from $\mathbf{a}_1$ to $\begin{bmatrix}1\\ 0\end{bmatrix}$ is $\sqrt{(1 - 2/3)^2 + (-1/3)^2} = \sqrt{2/9} = \sqrt{2}/3 \approx 0.47$. Since the first distance $2\sqrt{53}/45$ is smaller than the second distance $\sqrt{2}/3$ (either by comparing the decimal approximations or by cross-multiplying and squaring both sides), this step of Newton's method does bring us closer to the solution $\begin{bmatrix}1\\ 0\end{bmatrix}$. The points $\mathbf{a}_1$ and $\mathbf{a}_2$ are in yellow near the red point $(1, 0)$ in Figure 18.8.1 (showing that $\mathbf{a}_2$ is somewhat off the "direct route" from $\mathbf{a}_1$ to $\begin{bmatrix}1\\ 0\end{bmatrix}$).

◇

12. (a) Let's first pretend we already know B^{-1} exists and try to "solve" for it from the relation $A = BC$. Multiplying both sides on the left by B^{-1} would give $B^{-1}A = B^{-1}(BC) = (B^{-1}B)C = \mathrm{I}_n C = C$, and then multiplying both sides on the right by A^{-1} gives $CA^{-1} = (B^{-1}A)A^{-1} = B^{-1}(AA^{-1}) = B^{-1}\mathrm{I}_n = B^{-1}$. This suggests that we try CA^{-1} as a candidate for an inverse to B (the preceding doesn't show it works, since that relied on the *assumption* that B has an inverse, which we don't yet actually know to be the case!). Note that invertibility of A ensures that the expression CA^{-1} at least makes sense.

Now we calculate: $B(CA^{-1}) = (BC)A^{-1} = AA^{-1} = \mathrm{I}_n$. By Theorem 18.1.8, this confirms that CA^{-1} and B are inverse to each other, as desired.But if we multiply both sides of the equation $A = BC$ by A^{-1} on the right, we obtain $\mathrm{I}_n = BCA^{-1}$, as desired.

(If you were to try multiplication in the other order, you would want to show $(CA^{-1})B = \mathrm{I}_n$. This requires more effort to show directly, effort that is unnecessary via the solution above. But if you'd like to see it done, here is one way. We multiply both sides of $A = BC$ by A^{-1} on the left to obtain $\mathrm{I}_n = A^{-1}(BC) = (A^{-1}B)C$, with both $A^{-1}B$ and C are $n \times n$ matrices. Thus, by Theorem 18.1.8 , the $n \times n$ matrices C and $A^{-1}B$ are inverse to each other and hence also satisfy $C(A^{-1}B) = \mathrm{I}_n$. The left side is $(CA^{-1})B$, so we're done. Note that this approach also used Theorem 18.1.8 !)

(b) We can directly multiply $(A^{-1}B)C = A^{-1}(BC) = A^{-1}A = \mathrm{I}_n$, so by Theorem 18.1.8 we are done.

(If you were to try multiplication in the other order, you would want to show $C(A^{-1}B) = \mathrm{I}_n$. One want to obtain this is to use associativity to rewrite it as $(CA^{-1})B = \mathrm{I}_n$, which holds by (a).)

$\diamond$