1. (a) In the nth year from now there are x_n customers of P_1; by assumption 85% of those remain with P_1 into the following year, contributing $(0.85)x_n$ to the value of x_{n+1}, and 8% of the y_n customers for P_2 switch to P_1, contributing a further $(0.08)y_n$ to the value of x_{n+1}. The sum of these gives the stated formula for x_{n+1}.

   We derive the formula for y_{n+1} the same way. Now the roles of 85% and 8% are played by 92% and 15%, and this leads to the stated formula for y_{n+1}. We can express all of this in matrix form using the Markov matrix

$$M = \begin{bmatrix} .85 & .08 \\ .15 & .92 \end{bmatrix}.$$

   (b) The matrix M^2 tells us about the behavior 2 years into the future from a given year, since $\mathbf{v}_{n+2} = M\mathbf{v}_{n+1} = M(M\mathbf{v}_n) = (MM)\mathbf{v}_n = M^2\mathbf{v}_n$. Thus, 73.45% of P_1 customers at the start of a calendar year are with P_1 and the end of two years' time (they may have moved to P_2 and then moved back in the meantime), and 14.16% of P_2 customers at the start of a calendar year have switched to P_1 by the end of two years' time.

   (c) In the long run, 34.78% of the population is with P_1 and 65.22% of the population is with P_2. (Customers keep moving around, but the "total market share" for each company stabilizes at these proportions.) ♢

2. (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. ♢

3. (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.). ♢

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

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

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

(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:

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

$\diamondsuit$

5. (a) First consider city C. Next year's urban population consists of 93% of last year's x_U (namely, those not in the 7% of x_U that moved to the suburbs) and 3% of last year's x_S, so the new value of x_U at the end of this year will be $(0.93)x_U + (0.03)x_S$, using last year's x_U and x_S. Likewise, the new value of x_S at the end of this year will be $(0.07)x_U + (0.97)x_S$ since it has a contribution from the 7% of the urban residents who move to the suburbs during this year and the 97% of suburban residents not in the 3% who move to the urban part of the city. Putting it all together,

$$M = \begin{bmatrix} .93 & .03 \\ .07 & .97 \end{bmatrix}.$$

We can apply the same reasoning for C', with the roles of 7% and 3% being replaced by 5% and 4% respectively, yielding

$$M' = \begin{bmatrix} .95 & .04 \\ .05 & .96 \end{bmatrix}.$$

(b) For M, the meaning is that in the long run, 30% of the population of C is urban and 70% of the population of C is suburban. Likewise, for M' the meaning is that in the long run, 44.4% of its population is urban and 55.6% of its population is suburban. It is indeed surprising that a change of just a 1% or 2% in the proportions moving in either direction has an impact on the long-term population distribution on the order of around 15% each! $\diamondsuit$

6. As a prelude to using the Chain Rule, we first calculate the derivative matrices

$$(Df)(x, y) = \begin{bmatrix} \cos(x) & 0 \\ 2xye^{yx^2} & x^2e^{yx^2} \end{bmatrix}, \quad (Dg)(r, s, t) = \begin{bmatrix} st & rt & rs \end{bmatrix}, \quad (Dh)(v) = \begin{bmatrix} 1 + \cos(v) \\ 1 - \sin(v) \end{bmatrix}.$$

(a) This is the matrix product

$$\begin{bmatrix} 1+\cos(rst) \\ 1-\sin(rst) \end{bmatrix} \begin{bmatrix} st & rt & rs \end{bmatrix} = \begin{bmatrix} st+st\cos(rst) & rt+rt\cos(rst) & rs+rs\cos(rst) \\ st-st\sin(rst) & rt-rt\sin(rst) & rs-rs\sin(rst) \end{bmatrix}.$$

(b) Since $h(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we are computing the matrix product

$$(Df)(0,1)\,(Dh)(0) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

(c) Explicitly, $g(f(x,y),x) = x\sin(x)e^{yx^2}$, and its x-partial is computed by using the product rule and Chain Rule of single-variable calculus treating y as constant:

$$\begin{aligned} ((x\sin(x))e^{yx^2})' &= (x\sin(x))'e^{yx^2} + (x\sin(x))(e^{yx^2})' \\ &= (\sin(x) + x\cos(x))e^{yx^2} + 2xy(x\sin(x))e^{yx^2} \\ &= e^{yx^2}(\sin(x) + x\cos(x) + 2x^2y\sin(x)). \end{aligned}$$

For the second way, the x-partial is the first entry in the 1×3 matrix $D(g\circ F)(x,y)$ where $F(x,y) = (\sin(x), e^{yx^2}, x)$. By the Chain Rule, this derivative matrix is

$$\begin{aligned} (Dg)(F(x,y))\,(DF)(x,y) &= \begin{bmatrix} st & rt & rs \end{bmatrix}\Big|_{F(x,y)} \begin{bmatrix} \cos(x) & 0 \\ 2xye^{yx^2} & x^2e^{yx^2} \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} e^{yx^2}x & x\sin(x) & e^{yx^2}\sin(x) \end{bmatrix} \begin{bmatrix} \cos(x) & 0 \\ 2xye^{yx^2} & x^2e^{yx^2} \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} e^{yx^2}x\cos(x) + 2x^2y\sin(x)e^{yx^2} + e^{yx^2}\sin(x) & x^3\sin(x)e^{yx^2} \end{bmatrix}. \end{aligned}$$

Hence, looking at the first entry gives the desired x-partial as $e^{yx^2}(x\cos(x)+(2x^2y+1)\sin(x))$. This agrees with what we got the first way, but by an entirely different-looking procedure. $\diamondsuit$

7. (a) We can view $f(x(r,s), y(r,s))$ as just a function of r (by holding s fixed). We calculate $\partial x/\partial r = 2r$ and $\partial y/\partial r = s$. Formula (17.1.6) applies and we get

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = \frac{\partial f}{\partial x}\cdot(2r) + \frac{\partial f}{\partial y}\cdot s.$$

(b) We iteratively apply the above formula

$$\begin{aligned} \frac{\partial f}{\partial v} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} \\ &= \frac{\partial f}{\partial x}\left(\frac{\partial x}{\partial r}\frac{\partial r}{\partial v} + \frac{\partial x}{\partial s}\frac{\partial s}{\partial v}\right) + \frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial r}\frac{\partial r}{\partial v} + \frac{\partial y}{\partial s}\frac{\partial s}{\partial v}\right) \\ &= \frac{\partial f}{\partial x}\cdot(2r+4sv) + \frac{\partial f}{\partial y}\cdot(s+2rv) \end{aligned}$$

(c) We calculate

$$\begin{aligned}
\frac{\partial f}{\partial t} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \\
&= \frac{\partial f}{\partial x}\left(\frac{\partial x}{\partial r}\frac{\partial r}{\partial t} + \frac{\partial x}{\partial s}\frac{\partial s}{\partial t}\right) + \frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial r}\frac{\partial r}{\partial t} + \frac{\partial y}{\partial s}\frac{\partial s}{\partial t}\right) \\
&= \frac{\partial f}{\partial x}\left(\frac{\partial x}{\partial r}\left(\frac{\partial r}{\partial v}\frac{\partial v}{\partial t}\right) + \frac{\partial x}{\partial s}\left(\frac{\partial s}{\partial v}\frac{\partial v}{\partial t}\right)\right) + \frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial r}\left(\frac{\partial r}{\partial v}\frac{\partial v}{\partial t}\right) + \frac{\partial y}{\partial s}\left(\frac{\partial s}{\partial v}\frac{\partial v}{\partial t}\right)\right),
\end{aligned}$$

where in the final equality there are no terms involving $\partial w/\partial t$ since this vanishes. Evaluating all of the partial derivatives apart from the unknown partials of f with respect to x and y now yields

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\cdot(4rt + 8svt) + \frac{\partial f}{\partial y}\cdot(2st + 4rvt). \qquad \Diamond$$

8. (a) The Chain Rule applied to the function $F(x, y(x)) = (F \circ h)(x) = c$ that is constant gives that its derivative $0 = c' = (F \circ h)'(x)$ is the entry in the 1×1 matrix

$$(DF)(x, y(x))\,(Dh)(x) = \begin{bmatrix} F_x & F_y \end{bmatrix}\begin{bmatrix} 1 \\ y'(x) \end{bmatrix} = \begin{bmatrix} F_x + F_y\, y'(x) \end{bmatrix}.$$

This says $0 = F_x + F_y y'(x)$, so $y'(x) = -F_x/F_y$ wherever F_y is non-vanishing.

(b) Let $F(x, y) = 2x^3y - y^5x$. According to (a) we have

$$y'(x) = -\frac{F_x(x,y)}{F_y(x,y)} = -\frac{6x^2y - y^5}{2x^3 - 5y^4x}.$$

Hence the slope $(1, 1)$ is $-\frac{6-1}{2-5} = \frac{5}{3}$. $\Diamond$

9. (a) Since $f(x, y, z(x, y)) = c$ is constant, its y-partial derivative vanishes. But the Chain Rule computes the y-partial of $f(x, y, z(x, y)) = (f \circ g)(x, y)$ for $g(x, y) = (x, y, z(x, y))$ to be the second entry in the matrix

$$\begin{aligned}
D(f \circ g)(x, y) = (Df)(g(x, y))\,(Dg)(x, y) &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix}\Big|_{g(x,y)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \partial z/\partial x & \partial z/\partial y \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x} & \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y} \end{bmatrix}.
\end{aligned}$$

In other words, we have $0 = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial y}$, so

$$\frac{\partial z}{\partial y} = -\frac{\partial f/\partial y}{\partial f/\partial z}$$

as desired.

The exact same reasoning, but treating y as a function of (x, z) on the level surface $f = c$ and computing its x-partial yields $\partial y/\partial x = -f_x/f_y$. Likewise, treating x as a function of (y, z) on the level surface $f = c$ and computing its z-partial yields $\partial x/\partial z = -f_z/f_x$; doing similarly with the y-partial yields $\partial x/\partial y = -f_y/f_x$, which is the reciprocal of $-f_x/f_y = \partial y/\partial x$ as desired.

(b) Putting it all together,
$$\frac{\partial z}{\partial y}\frac{\partial y}{\partial x}\frac{\partial x}{\partial z} = \left(-\frac{f_y}{f_z}\right)\left(-\frac{f_x}{f_y}\right)\left(-\frac{f_z}{f_x}\right) = (-1)^3\,\frac{f_y}{f_z}\cdot\frac{f_x}{f_y}\cdot\frac{f_z}{f_x} = -1$$
(by cancellation of numerators and denominators in genuine fractions!).

(c) Since $P = cT/V$, we have $\partial P/\partial T = c/V$. Since $T = PV/c$, we have $\partial T/\partial V = P/c$. Since $V = cT/P$, we have $\partial V/\partial P = -cT/P^2$. Hence, the product of all three of these is $(c/V)(P/c)(-cT/P^2) = -cT/(VP) = -1$. ♢

10. Let $H : \mathbf{R}^n \to \mathbf{R}^{2m}$ be given by $H(\mathbf{x}) = (f(\mathbf{x}), g(\mathbf{x}))$ and $M : \mathbf{R}^{2m} \to \mathbf{R}$ be given by $M(\mathbf{u}, \mathbf{v}) = \mathbf{u}\cdot\mathbf{v}$. Then we have $h = M \circ H$. To calculate the derivative we calculate
$$(DH)(\mathbf{x}, \mathbf{y}) = \begin{bmatrix}(Df)(\mathbf{x})\\(Dg)(\mathbf{x})\end{bmatrix}$$
and
$$(DM)(\mathbf{u}, \mathbf{v}) = \begin{bmatrix}v_1 & \cdots & v_m & u_1 & \cdots & u_m\end{bmatrix}.$$
Thus,
$$\begin{aligned} D(M \circ H)(\mathbf{x}) &= (DM)(f(\mathbf{x}), g(\mathbf{x}))\begin{bmatrix}(Df)(\mathbf{x})\\(Dg)(\mathbf{x})\end{bmatrix} \\ &= \begin{bmatrix}g_1(\mathbf{x}) & \cdots & g_m(\mathbf{x}) & f_1(\mathbf{x}) & \cdots & f_m(\mathbf{x})\end{bmatrix}\begin{bmatrix}(Df)(\mathbf{x})\\(Dg)(\mathbf{x})\end{bmatrix} \\ &= g(\mathbf{x})^\top(Df)(\mathbf{x}) + f(\mathbf{x})^\top(Dg)(\mathbf{x}) \end{aligned}$$
(where the final equality is seen by considering how the matrix product works out entry by entry; think through the case $m = n = 2$ to see what is going on). ♢

11. (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
$$(D_{\mathbf{v}}f)(1, 2) = ((Df)(1, 2))\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix} = \begin{bmatrix}2\pi & \pi\end{bmatrix}\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix} = 3\pi/\sqrt{2}.$$ ♢

12. (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

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

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 $g_2(r,\theta) = \sin\theta$. $\diamondsuit$