$\underline{\star 1}$ $(\mathbf{A}+\mathbf{B})^2$ and $(\mathbf{B}+\mathbf{A})^2$ are equal because matrix addition is commutative (making $\mathbf{A}+\mathbf{B}$ the same as $\mathbf{B}+\mathbf{A}$). Since we are squaring equal things, the results are the same.

$(\mathbf{A}+\mathbf{B})^2$ and $(\mathbf{A}+\mathbf{B})(\mathbf{B}+\mathbf{A})$ are equal because the two factors in the right-hand expression are equivalent ways to write the same thing (commutativity) of addition).

$(\mathbf{A}+\mathbf{B})^2$, $\mathbf{A}(\mathbf{A}+\mathbf{B})+\mathbf{B}(\mathbf{A}+\mathbf{B})$, and $\mathbf{A}^2+\mathbf{AB}+\mathbf{BA}+\mathbf{B}^2$ are equal since

$$\begin{aligned}(\mathbf{A}+\mathbf{B})^2 &= (\mathbf{A}+\mathbf{B})(\mathbf{A}+\mathbf{B}) \\ &= \mathbf{A}(\mathbf{A}+\mathbf{B})+\mathbf{B}(\mathbf{A}+\mathbf{B}) && \text{(distributive law)} \\ &= \mathbf{A}^2+\mathbf{AB}+\mathbf{BA}+\mathbf{B}^2 && \text{(distributive law again).}\end{aligned}$$

But $\mathbf{A}^2+2\mathbf{AB}+\mathbf{B}^2$ is different from the others. Comparing it with $\mathbf{A}^2+\mathbf{AB}+\mathbf{BA}+\mathbf{B}^2$, we see that they would be equal only if $\mathbf{AB}=\mathbf{BA}$, which is rarely true.

$\underline{\star 2}$ (a) Here

$$\mathbf{AB}_1 = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} \qquad \text{while} \qquad \mathbf{B}_1\mathbf{A} = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}.$$

So, if $\mathbf{AB}_1 = \mathbf{B}_1\mathbf{A}$, then $b=0=c$. Similarly,

$$\mathbf{AB}_2 = \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix} \qquad \text{while} \qquad \mathbf{B}_2\mathbf{A} = \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix}.$$

So, if $\mathbf{AB}_2 = \mathbf{B}_2\mathbf{A}$, then $a=d$. Put together, this means

$$\mathbf{A} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = a\mathbf{I}.$$

(b) The matrix $\mathbf{A}=a\mathbf{I}$ from part (a) will, indeed, commute with all other 2-by-2 matrices, since

$$(a\mathbf{I})\mathbf{B} = a(\mathbf{IB}) = a\mathbf{B} = a(\mathbf{BI}) = \mathbf{B}(a\mathbf{I}).$$

(c) One can argue, as in part (b), that an n-by-n matrix $\mathbf{A}=a\mathbf{I}_n$ (i.e., a scalar multiple of the n-by-n identity matrix) will commute with any other n-by-n matrix $\mathbf{B}$. Part (a) might make one further suspect that if there are any nonzero entries in $\mathbf{A}$ *off* the main diagonal, then it will *not* commute with all $\mathbf{B}$.

$\underline{\star 3}$ (a) We have

$$\begin{aligned}\mathbf{AB}=\mathbf{AC} \quad &\Rightarrow \quad \mathbf{A}^{-1}(\mathbf{AB}) = \mathbf{A}^{-1}(\mathbf{AC}) \\ &\Rightarrow \quad \mathbf{B}=\mathbf{C}.\end{aligned}$$

(b) There are a *lot* of possible matrices **B**, **C** that satisfy the contraints of this problem. There are some requirements, however. If we write

$$\mathbf{B} = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} \qquad \text{and} \qquad \mathbf{C} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix},$$

then

$$\mathbf{AB} = \begin{bmatrix} b_1 & b_2 \\ 0 & 0 \end{bmatrix} \qquad \text{and} \qquad \mathbf{AC} = \begin{bmatrix} c_1 & c_2 \\ 0 & 0 \end{bmatrix}.$$

Thus, any correct pair **B**, **C** will have the same first row for both, but will have a different 2$^{\text{nd}}$ row (the one from the other).

★4 (a) It is

$$\begin{aligned} 2x_2 + x_3 + 3x_4 &= 3, \\ 2x_1 + x_2 + 2x_3 - x_4 &= 4, \\ x_1 - 3x_2 + x_3 + x_4 &= 7, \\ 2x_1 + x_3 - 2x_4 &= 2. \end{aligned}$$

(b) We have

$$\left[\begin{array}{cccc|c} 0 & 2 & 1 & 3 & 3 \\ 2 & 1 & 2 & -1 & 4 \\ 1 & -3 & 1 & 1 & 7 \\ 2 & 0 & 1 & -2 & 2 \end{array}\right] \underset{\sim}{\mathbf{r}_1 \leftrightarrow \mathbf{r}_3} \left[\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 7 \\ 2 & 1 & 2 & -1 & 4 \\ 0 & 2 & 1 & 3 & 3 \\ 2 & 0 & 1 & -2 & 2 \end{array}\right]$$

$$\underset{\sim}{(-2)r_1 + \mathbf{r}_2 \to \mathbf{r}_2} \left[\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 7 \\ 0 & 7 & 0 & -3 & -10 \\ 0 & 2 & 1 & 3 & 3 \\ 2 & 0 & 1 & -2 & 2 \end{array}\right] \underset{\sim}{(-2)r_1 + \mathbf{r}_4 \to \mathbf{r}_4} \left[\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 7 \\ 0 & 7 & 0 & -3 & -10 \\ 0 & 2 & 1 & 3 & 3 \\ 0 & 6 & -1 & -4 & -12 \end{array}\right]$$

$$\underset{\sim}{\mathbf{r}_2 \leftrightarrow \mathbf{r}_3} \left[\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 7 \\ 0 & 2 & 1 & 3 & 3 \\ 0 & 7 & 0 & -3 & -10 \\ 0 & 6 & -1 & -4 & -12 \end{array}\right] \underset{\sim}{(-7/2)\mathbf{r}_2 + \mathbf{r}_3 \to \mathbf{r}_3} \left[\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 7 \\ 0 & 2 & 1 & 3 & 3 \\ 0 & 0 & -3.5 & -13.5 & -20.5 \\ 0 & 6 & -1 & -4 & -12 \end{array}\right]$$

$$\underset{\sim}{(-3)\mathbf{r}_2 + \mathbf{r}_4 \to \mathbf{r}_4} \left[\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 7 \\ 0 & 2 & 1 & 3 & 3 \\ 0 & 0 & -3.5 & -13.5 & -20.5 \\ 0 & 0 & -4 & -13 & -21 \end{array}\right] \underset{\sim}{(-8/7)\mathbf{r}_3 + \mathbf{r}_4 \to \mathbf{r}_4} \left[\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 7 \\ 0 & 2 & 1 & 3 & 3 \\ 0 & 0 & -3.5 & -13.5 & -20.5 \\ 0 & 0 & 0 & 17/7 & 17/7 \end{array}\right]$$

(c) There is, unfortunately, not just one sequence of EROs leading to RREF, though the end result must always be the same. Here is one sequence the produces the desired result.

viii. ERO2: rescale row 2 by a factor of (1/2); i.e., $(1/2)\mathbf{r}_2 \to \mathbf{r}_2$

ix. ERO2: rescale row 3 by a factor of $(-2/7)$; that is, $(-2/7)\mathbf{r}_3 \to \mathbf{r}_3$

x. ERO2: rescale row 4 by a factor of (7/17); $(7/17)\mathbf{r}_4 \to \mathbf{r}_4$

xi. ERO3: $\mathbf{r}_1 - \mathbf{r}_4 \to \mathbf{r}_1$

xii. ERO3: $\mathbf{r}_2 - (3/2)\mathbf{r}_4 \to \mathbf{r}_2$

xiii. ERO3: $\mathbf{r}_3 - (27/7)\mathbf{r}_4 \to \mathbf{r}_3$

xiv. ERO3: $\mathbf{r}_2 - (1/2)\mathbf{r}_3 \to \mathbf{r}_2$

xv. ERO3: $\mathbf{r}_1 - \mathbf{r}_3 \to \mathbf{r}_1$

xvi. ERO3: $\mathbf{r}_1 + 3\mathbf{r}_2 \to \mathbf{r}_1$

$$\left[\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 7 \\ 0 & 2 & 1 & 3 & 3 \\ 0 & 0 & -3.5 & -13.5 & -20.5 \\ 0 & 0 & 0 & 17/7 & 17/7 \end{array}\right] \begin{array}{c} (1/2)\mathbf{r}_2 \to \mathbf{r}_2 \\ (-2/7)\mathbf{r}_3 \to \mathbf{r}_3 \\ \sim \\ (7/17)\mathbf{r}_4 \to \mathbf{r}_4 \end{array} \left[\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 7 \\ 0 & 1 & 0.5 & 1.5 & 1.5 \\ 0 & 0 & 1 & 27/7 & 41/7 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

$$\begin{array}{c} -\mathbf{r}_4 + \mathbf{r}_1 \to \mathbf{r}_1 \\ (-3/2)\mathbf{r}_4 + \mathbf{r}_2 \to \mathbf{r}_2 \\ \sim \\ (-27/7)\mathbf{r}_4 + \mathbf{r}_3 \to \mathbf{r}_3 \end{array} \left[\begin{array}{cccc|c} 1 & -3 & 1 & 0 & 6 \\ 0 & 1 & 0.5 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] \begin{array}{c} \mathbf{r}_2 - (1/2)\mathbf{r}_3 \to \mathbf{r}_2 \\ \sim \\ \mathbf{r}_1 - \mathbf{r}_3 \to \mathbf{r}_1 \end{array} \left[\begin{array}{cccc|c} 1 & -3 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

$$\begin{array}{c} \mathbf{r}_1 + 3\mathbf{r}_2 \to \mathbf{r}_1 \\ \sim \end{array} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

(d) The (only) solution is $\mathbf{x} = (1, -1, 2, 1)$.

★5 (a) To find out how many subscribers after 2 years, you could left-multiply the vector $(6000, 4000)$ by the same matrix as before. Alternatively, you could left-multiply $(8000, 2000)$ by $\mathbf{A}^2$, where `A = [0.7 0.2; 0.3 0.8]`. In the code below, I have chosen 25 years as long enough for the long-term behavior to appear. (The software I use below is Octave, a clone of Matlab.)

```
octave-3.0.0:40> A = [0.7  0.2;  0.3  0.8];
octave-3.0.0:41> A^25*[6000; 4000]
ans =
   4000.0
   6000.0
```

So, after 25 years, there are 4000 subscribing households and 6000 non-subscribers. To see that this is, in fact, a steady state, you can left-multiply by $\mathbf{A}$ again, and the figure does not change.

(b) Repeating the above with an initial vector of $(9000, 1000)$ still yields a long-term outlook of 4000 subscribing households. This is not surprising, since just one year gets the initial number of subscribing households from 9000 down to 6500.

$\underline{\star 6}$ (a) $\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$.

(b) Here, **P** should be the transpose of the **P** from part (a), namely $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.

(c) $\mathbf{B} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}$.

$\underline{\star 7}$ By exchanging rows 1 and 3, **H** gets to RREF:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Writing an element of null $(\mathbf{H})$ as $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$, we see that we can take x_3, x_5, x_6 and x_7 as "free" variables (free to take on either of the values 0 or 1), while

$$\begin{aligned} x_1 &= x_3 + x_5 + x_7 \\ x_2 &= x_3 + x_6 + x_7 \ , \\ x_4 &= x_5 + x_6 + x_7 \end{aligned}$$

so vectors in the null space take the form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_7 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Thus, one possible basis is

$$\{(1,1,1,0,0,0,0), (1,0,0,1,1,0,0), (0,1,0,1,0,1,0), (1,1,0,1,0,0,1)\}.$$

Some notes:

- That this collection spans null$(\mathbf{H})$ is clear from the solution process. That it is linearly independent perhaps calls for forming a 7-by-4 matrix with these as the columns, reducing to echelon form and seeing that that echelon form has no free columns. I don't necessarily expect students will do this.
- As is pretty much always the case, there are other bases for the same subspace—we use a different basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ for null$(\mathbf{H})$ in part (b). While the one I've given above is the most likely basis for students to find, there are yet others. The easiest way to check a strange-looking answer is to make sure the proposed collection contains 4 vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ all from $\mathbb{Z}_2^7$, check that each one is in null$(\mathbf{H})$ (i.e., that $\mathbf{H}\mathbf{v}_j = \mathbf{0}$ for $j = 1, 2, 3, 4$), and that they are linearly independent.

★8 (a) The transmitted $\mathbf{v}$ is given by

$$\mathbf{v} \;=\; 1\begin{bmatrix}1\\0\\0\\0\\0\\1\\1\end{bmatrix} + 1\begin{bmatrix}0\\1\\0\\0\\1\\0\\1\end{bmatrix} + 0\begin{bmatrix}0\\0\\1\\0\\1\\1\\0\end{bmatrix} + 1\begin{bmatrix}0\\0\\0\\1\\1\\1\\1\end{bmatrix} \;=\; \begin{bmatrix}1\\1\\0\\1\\0\\0\\1\end{bmatrix}.$$

(b) We have

$$\mathbf{H}\tilde{\mathbf{v}} \;=\; \begin{bmatrix}0&0&0&1&1&1&1\\0&1&1&0&0&1&1\\1&0&1&0&1&0&1\end{bmatrix}\begin{bmatrix}1\\0\\1\\1\\1\\0\\0\end{bmatrix} \;=\; \begin{bmatrix}0\\1\\1\end{bmatrix}.$$

Since $\mathbf{H}\tilde{\mathbf{v}}$ is not the zero vector, $\tilde{\mathbf{v}}$ is corrupted. If only corrupted in a single entry, it must be the the 3^{rd} entry (as $\mathbf{H}\tilde{\mathbf{v}}$ equals the 3^{rd} column of $\mathbf{H}$). Thus, the originally-intended 7-bit word is $\mathbf{v} = (1, 0, 0, 1, 1, 0, 0)$, from which we extract the 4-bit word $(1, 0, 0, 1)$, (or 1001).

(c) Sadly, the use of the Hamming (7,4) scheme for detecting and correcting errors breaks down if two (or more) bits from a 7-bit transmitted word are corrupt. To see this, notice that if the 7-bit word $\mathbf{v} = (1, 0, 0, 1, 1, 0, 0)$ is corrupted to $\tilde{\mathbf{v}} = (1, 1, 0, 1, 0, 0, 0)$ (*two* altered bits), then we will, indeed, *detect* an error (it is still the case, with this $\tilde{\mathbf{v}}$, that $\tilde{\mathbf{v}} \notin$ null$(\mathbf{H})$), but that our process for correction would make us think that the 7^{th} bit alone was faulty (not the pair of 2^{nd} and 5^{th} bits). With *three* altered bits we might not even detect the error!