SOLUTIONS TO MATH 240 EXERCISES

1.1-1.2: LINEAR SYSTEMS

(1) Recall that the set of solutions (x, y) to a single linear equation in 2 variables constitutes a line ℓ in $\mathbb{R}^2$. We denote this $\ell \colon ax + by = c$.

Similarly, the set of solutions (x, y, z) to a single linear equation in 3 variables $ax+by+cz = d$ constitutes a plane $\mathcal{P}$ in $\mathbb{R}^3$. We denote this $\mathcal{P} \colon ax + by + cz = d$.

   (a) Fix $m > 1$ and consider a system of m linear equations in the 2 unknowns x and y. What do solutions (x, y) to this *system* of linear equations correspond to geometrically?

   (b) Use your interpretation to give a *geometric* argument that a system of m equations in 2 unknowns will have either (i) 0 solutions, (ii) 1 solution, or (iii) infinitely many solutions.

   (c) Use your geometric interpretation to help produce explicit examples of systems in 2 variables satisfying these three different cases (i)-(iii).

   (d) Now repeat (a)-(b) for systems of linear equations in 3 variables x, y, z.

Solution 1:

(a) Geometrically, each equation in the system represents a line $\ell_i \colon a_i x + b_i y = c_i$. A solution (x, y) to the i-th equation corresponds to a point on ℓ_i. Thus a solution (x, y) to the system corresponds to a point lying on *all* of the lines: i.e., a point of intersection of the lines.

(b) First of all to prove the desired "or" statement it suffices to prove that if the number

of solutions is greater than 1, then there are infinitely many solutions.

Now suppose there is more than one solution. Then there are at least two different solutions: $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$. Take any of the two lines ℓ_i, ℓ_j. By above the intersection of ℓ_i and ℓ_j contains P_1 and P_2. But two *distinct* lines intersect in at most one point. It follows that ℓ_i and ℓ_j must be equal. Since ℓ_i and ℓ_j were arbitrary, it follows *all* of the lines ℓ_i are in fact the same line ℓ. But this means the common intersection of the lines is ℓ, which has infinitely many points. It follows that the system has infinitely many solutions.

(c) We will get 0 solutions if the system includes two different parallel lines: e.g., $\ell_1\colon x+y = 5$ and $\ell_2\colon x + y = 1$.

We will get exactly one solution when the slopes of each line in the system are distinct.

We will get infinitely many solutions when *all* equations in the system represent the *same line*. This happens when all equations are multiples of one another.

(d) Now each equation in our system defines a plane $\mathcal{P}_i\colon a_i x + b_i y + c_i z = d_i$. A solution (x, y, z) to the system corresponds to a point $P = (x, y, z)$ of intersection of the planes. We recall two facts from Euclidean geometry:

Fact 1 Given two distinct points, there is a unique line containing both of them.

Fact 2 Given any number of distinct planes, they either do not intersect, or intersect in a line.

We proceed as in part (b) above: that is show that if there are two distinct solutions to the system, then there are infinitely many solutions. First, for simplicity, we may assume that the equations $\mathcal{P}_i\colon a_i x + b_i y + c_i z = d_i$ define *distinct* planes; if we have two equations defining the same plane, we can delete one of them and not change the set of solutions to the system.

Now suppose $P = (x_1, y_1, z_1)$ and $Q = (x_2, y_2, z_2)$ are two distinct solutions to the system. Let ℓ be the unique line containing P and Q (Fact 1). I claim that ℓ is precisely the set of solutions to the system. To see this, take any two equations in the system $\mathcal{P}_i\colon a_i x + b_i y + c_i z = d_i$ and $\mathcal{P}_j\colon a_j x + b_j y + c_i z = d_j$. Since the two corresponding planes are distinct, and intersect in at least the points P and Q, they must intersect in a line (Fact 2); since this line contains P and Q, it must be the line ℓ (Fact 1). Thus any two planes in the system intersect in the line ℓ. From this it follows that: (a) a point satisfying the system must lie in ℓ; and (b) all points on ℓ satisfy the system (since we have shown that ℓ lies in all the planes). It follows that ℓ is precisely the set of solutions, and hence that there are infinitely many solutions.

(2) We made the claim that each of our three row operations

   i scalar multiplication ($e_i \mapsto c \cdot e_i$ for $c \neq 0$),
   ii swap ($e_i \leftrightarrow e_j$),
   iii addition ($e_i \mapsto e_i + c \cdot e_j$ for some c)

do not change the set of solutions of a linear system.

To prove this claim, let L be a general linear system

$$
\begin{array}{rlllllll}
e_1: & a_{11}x_1 & + & a_{12}x_2 & +\cdots+ & a_{1n}x_n & = & b_1 \\
e_2: & a_{21}x_1 & + & a_{22}x_2 & +\cdots+ & a_{2n}x_n & = & b_2 \\
 & \vdots & & \vdots & & \vdots & \vdots & \\
e_m: & a_{m1}x_1 & + & a_{m2}x_2 & +\cdots+ & a_{mn}x_n & = & b_m
\end{array}.
$$

Now consider each type of row operation separately, write down the new system L' you get by applying this row operation, and prove that an n-tuple $s = (s_1, s_2, \ldots, s_n)$ is a solution to the original system L $\boxed{\text{if and only if}}$ it is a solution to the new system L'.

Solution 2:

Let S be the original system with equations $e_1, e_2, \ldots, e_m$.

For each specified row operation, we will call the resulting new system S' and its equations $e'_1, e'_2, \ldots, e'_m$.

Row swap.: In this case systems S and S' have exactly the same equations, just written in a different order. Thus the n-tuple s satifies S if and only if it satisfies each of the equations e_i, if and only if it satisfies each of the equations e'_i, since these are the same equations! It follows that s is a solution of S if and only if it is a solution to S'.

Scalar mult.: In this case $e_j = e'_j$ for all $j \neq i$, and $e'_i = c \cdot e_i$ for some $c \neq 0$. Since only the i-th equation has changed, it suffices to show that s is a solution to e_i if and only if s is a solution to $c \cdot e_i$. Let's prove each direction of this if and only if separately.

$$s \text{ satisfies } e_i \Rightarrow a_{ii}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n = b_i \qquad \text{(by def.)}$$
$$\Rightarrow ca_{i1}s_1 + ca_{i2}s_2 + \cdots + ca_{in}s_n = cb_i \qquad \text{(mult both sides by } c\text{)}$$
$$\Rightarrow s \text{ satisfies } e'_i \qquad \text{(by def)}$$

Now for the other direction of the if and only if.

$$s \text{ satisfies } e'_i \Rightarrow ca_{ii}s_1 + ca_{i2}s_2 + \cdots ca_{in}s_n = cb_i \qquad \text{(by def)}$$
$$\Rightarrow \frac{1}{c}(ca_{i1}s_1 + ca_{i2}s_2 + \cdots + ca_{in}s_n) = \frac{1}{c} \cdot cb_i \qquad \left(\text{mult both sides by } \frac{1}{c}\right)$$
$$\Rightarrow a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}x_n = b_i \qquad \left(\text{distribute } \frac{1}{c}\right)$$
$$\Rightarrow s \text{ satisfies } e'_i \qquad \text{(by def.)}.$$

Row addition.: Here $e'_i = e_i + ce_j$. The only equation of S' that differs from S is

$$e'_i = e_i + ce_j.$$

Writing this equation out in terms of coefficients gives us

$$e'_i : a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + c(a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n) = b_i + cb_j.$$

Now if s satisfies S, then it satisfies e_i and e_j, in which case evaluating the RHS of the e'_i above at s yields

$$a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n + c(a_{j1}s_1 + a_{j2}s_2 + \cdots + a_{jn}s_n) \;\; = \;\; b_i + cb_j$$

showing that s satisfies e'_i.

Now suppose $s = (s_1, s_2, \ldots, s_n)$ satisfies S'. Since s satisfies $e'_j = e_j$, we have

$$(*) \qquad a_{j1}s_1 + a_{j2}s_2 + \cdots + a_{jn}s_n = b_j.$$

Since s satisfies e'_i, we have

$$a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n + c(a_{j1}s_1 + a_{j2}s_2 + \cdots + a_{jn}s_n) = b_i + cb_j$$

Substituting $(*)$ into the equation above we get

$$a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n + c(b_j) = b_i + cb_j,$$

and hence

$$a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n = b_i.$$

This shows that s satisfies e_i. It follows that s satisfies S.

(3) Consider the linear system

$$
\begin{aligned}
7x_5 &= 2x_3 + 12 \\
2x_1 + 4x_2 - 10x_3 + 6x_4 + 12x_5 &= 28 \\
2x_1 + 4x_2 - 5x_3 + 6x_4 - 5x_5 &= -1
\end{aligned}
$$

Using augmented matrices, try and solve this system by reducing it to a simpler one using row operations. How many solutions (0, 1, or ∞) does the system have?
Solution 3:

See my lecture notes on Gaussian elimination!

(4) Compute the set of solutions S to the following system of linear equations:

$$
\begin{aligned}
x_1 + x_2 - x_3 + x_4 &= 1 \\
-2x_1 - 2x_2 + 2x_3 - 2x_4 &= -2 \\
x_1 + x_2 + x_3 + 2x_4 &= 3
\end{aligned}
$$

Follow the Gaussian elimination algorithm *to the letter*, indicating what row operations are used. Make sure to indicate how the x_i are sorted into free and leading variables.
Your final answer should be described in set notation.

Solution 4:

$$
\begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
-2 & -2 & 2 & -2 & -2 \\
1 & 1 & 1 & 2 & 3
\end{bmatrix}
\xrightarrow{r_2+2r_1}
\begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 2 & 3
\end{bmatrix}
$$

$$
\xrightarrow{r_3-r_1}
\begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 1 & 2
\end{bmatrix}
$$

$$
\xrightarrow{r_2\leftrightarrow r_3}
\begin{bmatrix}
1 & 1 & -1 & 1 & 1 \\
0 & 0 & 2 & 1 & 2 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

$$
\xrightarrow{\frac{1}{2}r_2}
\begin{bmatrix}
\boxed{1} & 1 & -1 & 1 & 1 \\
0 & 0 & \boxed{1} & 1/2 & 1 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

The row echelon matrix tells us that $x_2 = s$ and $x_4 = t$ are the free variables. Back substitution then yields the general solution:

$$
\begin{aligned}
x_1 &= 2 - s - \frac{3t}{2} \\
x_2 &= s \\
x_3 &= 1 - \frac{t}{2} \\
x_4 &= t,
\end{aligned}
$$

Alternatively, the set of solutions is

$$
S = \left\{ \left(2 - s - \frac{3t}{2}, s, 1 - \frac{t}{2}, t\right) : s, t \in \mathbb{R} \right\}.
$$

(5) Determine all values of a for which the system has (a) no solutions, (b) exactly one solution, or (c) infinitely-many solutions.

$$\begin{aligned} x + 2y - \quad\;\; 3z &= \quad 4 \\ 3x - \;\; y + \quad\;\; 5z &= \quad 2 \\ 4x + \;\; y + (a^2 - 14)z &= a + 2 \end{aligned}$$

Solution 5:

The augmented matrix row reduces to

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & 0 & (a^2 - 16) & a - 4 \end{array}\right]$$

From this it follows that the system has:

a) 0 solutions iff $a^2 - 16 = 0$ and $a - 4 \neq 0$ iff $\boxed{a = -4}$;

(b) exactly one solution iff $a^2 - 16 \neq 0$ iff $\boxed{a \neq \pm 4}$;

(c) infinitely many solutions iff $a^2 - 16 = 0$ and $a - 4 = 0$ iff $\boxed{a = 4}$.

(6) Use Gaussian elimination to find the general solution to the following system of linear equations:

$$\begin{aligned} x_1 + 2x_2 &= x_3 + x_4 + 3 \\ 3x_1 + 6x_2 &= 2x_3 - 4x_4 + 8 \\ -x_1 + 2x_3 &= 2x_2 - x_4 - 1 \end{aligned}$$

Follow the Gaussian elimination algorithm to the letter, indicating what row operations are used. Make sure to indicate how the x_i are sorted into free and leading variables.

Solution 6:

$$\begin{bmatrix} 1 & 2 & -1 & -1 & 3 \\ 3 & 6 & -2 & 4 & 8 \\ -1 & -2 & 2 & 1 & -1 \end{bmatrix} \xrightarrow{r_2 - 3r_1} \begin{bmatrix} 1 & 2 & -1 & -1 & 3 \\ 0 & 0 & 1 & 7 & -1 \\ -1 & -2 & 2 & 1 & -1 \end{bmatrix}$$

$$\xrightarrow{r_3 + r_1} \begin{bmatrix} 1 & 2 & -1 & -1 & 3 \\ 0 & 0 & 1 & 7 & -1 \\ 0 & 0 & 1 & 0 & 2 \end{bmatrix}$$

$$\xrightarrow{r_3 - r_2} \begin{bmatrix} 1 & 2 & -1 & -1 & 3 \\ 0 & 0 & 1 & 7 & -1 \\ 0 & 0 & 0 & -7 & 3 \end{bmatrix}$$

$$\xrightarrow{-\frac{1}{7}r_3} \begin{bmatrix} \boxed{1} & 2 & -1 & -1 & 3 \\ 0 & 0 & \boxed{1} & 7 & -1 \\ 0 & 0 & 0 & \boxed{1} & -\frac{3}{7} \end{bmatrix}$$

The row echelon matrix tells us that $x_2 = t$ is the only free variable. Back substitution then yields the general solution:

$$\begin{aligned} x_1 &= \frac{32}{7} - 2t \\ x_2 &= t \\ x_3 &= 2 \\ x_4 &= -\frac{3}{7}. \end{aligned}$$

5

Alternatively, the set of solutions is

$$S = \left\{ \left(\frac{32}{7} - 2t, t, 2, -3/7 \right) : t \in \mathbb{R} \right\}.$$

(7) Show that a linear system with more unknowns than equations has has either 0 solutions or infinitely many solutions.
Solution 7:

Suppose we have a system of m equations in n unknowns $x_1, x_2, \ldots, x_n$. We assume $n > m$. Let A be the augmented matrix associated to the system, and suppose A is reduced to a matrix U in row echelon form.
Since U has m rows, there are *at most* m leading 1's in U, which means there are at most m leading variables among the x_i. Since $n > m$, not all the x_i can be leading. Thus the system *must* have a free variable.
What does this mean? Note that the system could still be inconsistent, meaning no solutions. However, the existence of a free variable means if there is a solution, then there are infinitely many, because the parametric equations for the x_i will involve at least one parameter.
We conclude that the system is either inconsistent, or has infinitely many solutions.

(8) True or false. If true you must prove it so; if false, you must give an explicit counterexample.
   (a) Every matrix has a unique row echelon form.
   (b) Any homogeneous linear system with more unknowns than equations has infinitely many solutions.
   (c) The only solution to a homogeneous system of n equations in n unknowns with n leading 1's is the trivial solution $s = (0, 0, \ldots, 0)$.

Solution 8:
(a) False. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Then A is already in row echelon form, but can be further reduced to $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, which is also in row echelon form. Thus A and I are two different row echelon forms of A.
(b) True. Since its homogeneous, it is consistent: thus 1 or ∞-many solutions. Furthermore, we showed above that a system with more variables than equations has a free variable. Thus it must have infinitely many solutions.
(c) True. First note that the given s is always a solution. Since the system has n leading ones, there will be no free variables, which implies this is the only solution.

(9) Use Gaussian elimination to solve the following system:

$$\begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 0 \\ -2x_1 + 5x_2 + 2x_3 &= 1 \\ 8x_1 + x_2 + 4x_3 &= -1 \end{aligned}$$

Solution 9:

$$\begin{bmatrix} 2 & 2 & 2 & \big| & 0 \\ -2 & 5 & 2 & \big| & 1 \\ 8 & 1 & 4 & \big| & -1 \end{bmatrix} \xrightarrow{\frac{1}{2}r_1} \begin{bmatrix} 1 & 1 & 1 & \big| & 0 \\ -2 & 5 & 2 & \big| & 1 \\ 8 & 1 & 4 & \big| & -1 \end{bmatrix}$$

$$\xrightarrow{r_2+2r_1} \begin{bmatrix} 1 & 1 & 1 & \big| & 0 \\ 0 & 7 & 4 & \big| & 1 \\ 8 & 1 & 4 & \big| & -1 \end{bmatrix}$$

$$\xrightarrow{r_3-8r_1} \begin{bmatrix} 1 & 1 & 1 & \big| & 0 \\ 0 & 7 & 4 & \big| & 1 \\ 0 & -7 & -4 & \big| & -1 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{7}r_1} \begin{bmatrix} 1 & 1 & 1 & \big| & 0 \\ 0 & 1 & \frac{4}{7} & \big| & \frac{1}{7} \\ 0 & -7 & -4 & \big| & -1 \end{bmatrix}$$

$$\xrightarrow{r_3+7r_2} \begin{bmatrix} 1 & 1 & 1 & \big| & 0 \\ 0 & 1 & \frac{4}{7} & \big| & \frac{1}{7} \\ 0 & 0 & 0 & \big| & 0 \end{bmatrix}$$

Now solve. We set the free variable $x_3 = r$ and do back substitution:

$$\begin{aligned} x_1 &= -\frac{1}{7} - \frac{3}{7}r \\ x_2 &= \frac{1}{7} - \frac{4}{7}r \\ x_3 &= r \end{aligned}$$

(10) Solve by Gaussian elimination:

$$\begin{aligned} -2b + 3c &= 1 \\ 3a + 6b - 3c &= -2 \\ 6a + 6b + 3c &= 5 \end{aligned}$$

Solution 10:

$$\begin{bmatrix} 0 & -2 & 3 & \big| & 1 \\ 3 & 6 & -3 & \big| & -2 \\ 6 & 6 & 3 & \big| & 5 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{bmatrix} 3 & 6 & -3 & \big| & -2 \\ 0 & -2 & 3 & \big| & 1 \\ 6 & 6 & 3 & \big| & 5 \end{bmatrix}$$

$$\xrightarrow{r_3-2r_1} \begin{bmatrix} 3 & 6 & -3 & \big| & -2 \\ 0 & -2 & 3 & \big| & 1 \\ 0 & -6 & 9 & \big| & 9 \end{bmatrix}$$

$$\xrightarrow{-\frac{1}{2}r_2} \begin{bmatrix} 3 & 6 & -3 & \big| & -2 \\ 0 & 1 & -\frac{3}{2} & \big| & -\frac{1}{2} \\ 0 & -6 & 9 & \big| & 9 \end{bmatrix}$$

$$\xrightarrow{r_3+6r_2} \begin{bmatrix} 3 & 6 & -3 & \big| & -2 \\ 0 & 1 & -\frac{3}{2} & \big| & -\frac{1}{2} \\ 0 & 0 & 0 & \big| & 6 \end{bmatrix}$$

And since $0 \neq 6$, the system is inconsistent.

(11) Solve the following system using Gauss-Jordan elimination: that is, first reduce the corresponding augmented matrix to reduced row echelon form.

$$\begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 0 \\ -2x_1 + 5x_2 + 2x_3 &= 1 \\ 8x_1 + x_2 + 4x_3 &= -1 \end{aligned}$$

Solution 11:

The corresponding augmented matrix is

$$\left[\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ -2 & 5 & 2 & 1 \\ 8 & 1 & 4 & -1 \end{array}\right],$$

which row reduces first to

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & \frac{4}{7} & \frac{1}{7} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and then further to

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{3}{7} & -\frac{1}{7} \\ 0 & 1 & \frac{4}{7} & \frac{1}{7} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The corresponding system has solution set

$$S = \left\{ \left(-\frac{1}{7} - \frac{3}{7}r, \frac{1}{7} - \frac{4}{7}r, r \right) : r \in \mathbb{R} \right\}$$

as in Exercise 9.

(12) Solve the following system of equations by any method:

$$\begin{aligned} Z_3 + Z_4 + Z_5 &= 0 \\ -Z_1 - Z_2 + 2Z_3 - 3Z_4 + Z_5 &= 0 \\ Z_1 + Z_2 - 2Z_3 - Z_5 &= 0 \\ 2Z_1 + 2Z_2 - Z_3 + Z_5 &= 0 \end{aligned}$$

Solution 12:

$$
\left[\begin{array}{ccccc|c}
0 & 0 & 1 & 1 & 1 & 0 \\
-1 & -1 & 2 & -3 & 1 & 0 \\
1 & 1 & -2 & 0 & -1 & 0 \\
2 & 2 & -1 & 0 & 1 & 0
\end{array}\right]
\xrightarrow{-r_2 \leftrightarrow r_1}
\left[\begin{array}{ccccc|c}
1 & 1 & -2 & 3 & -1 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 \\
1 & 1 & -2 & 0 & -1 & 0 \\
2 & 2 & -1 & 0 & 1 & 0
\end{array}\right]
$$

$$
\xrightarrow{r_3 - r_1}
\left[\begin{array}{ccccc|c}
1 & 1 & -2 & 3 & -1 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & -3 & 0 & 0 \\
2 & 2 & -1 & 0 & 1 & 0
\end{array}\right]
$$

$$
\xrightarrow{r_4 - 2r_1}
\left[\begin{array}{ccccc|c}
1 & 1 & -2 & 3 & -1 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & -3 & 0 & 0 \\
0 & 0 & 3 & -6 & 3 & 0
\end{array}\right]
$$

$$
\xrightarrow{r_1 - 2r_2}
\left[\begin{array}{ccccc|c}
1 & 1 & 0 & 5 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & -3 & 0 & 0 \\
0 & 0 & 3 & -6 & 3 & 0
\end{array}\right]
$$

$$
\xrightarrow{r_4 - 3r_2}
\left[\begin{array}{ccccc|c}
1 & 1 & 0 & 5 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & -3 & 0 & 0 \\
0 & 0 & 0 & -9 & 0 & 0
\end{array}\right]
$$

$$
\xrightarrow{-\frac{1}{3}r_3}
\left[\begin{array}{ccccc|c}
1 & 1 & 0 & 5 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & -9 & 0 & 0
\end{array}\right]
$$

$$
\xrightarrow{r_4 + 9r_3}
\left[\begin{array}{ccccc|c}
1 & 1 & 0 & 5 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
$$

$$
\xrightarrow{r_2 - r_3}
\left[\begin{array}{ccccc|c}
1 & 1 & 0 & 5 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
$$

$$
\xrightarrow{r_1 - 5r_3}
\left[\begin{array}{ccccc|c}
1 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
$$

Now solve. We set the free variables $x_2 = r$ and $x_5 = s$ and substitute:

$$
\begin{aligned}
Z_1 &= -r - s \\
Z_2 &= r \\
Z_3 &= -s \\
Z_4 &= 0 \\
Z_5 &= s
\end{aligned}
$$

(13) Interpret each matrix below as an augmented matrix of a linear system. Asterisks represent an unspecified real number. For each matrix, determine whether the corresponding system is consistent or inconsistent.

If the system is consistent, decide further whether the solution is unique or not.

If there is not enough information answer 'inconclusive' and back up your claim by giving an explicit example where the system is consistent, and an explicit example where the system is inconsistent.

(a) $\begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 0 & * \\ * & 1 & 0 & * \\ * & * & 1 & * \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & * & * & * \end{bmatrix}$

(d) $\begin{bmatrix} 1 & * & * & * \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

Solution 13:

(a) The corresponding system is consistent since the row echelon form of the augmented matrix has no leading 1 in the last column. Since the three columns corresponding to the three variables all have leading 1's, there are no free variables. Hence the system has a unique solution.
(b) This system has a unique solution. You can see this either by noting that the "reverse staircase pattern" allows us to do "forwards substitution", solving first for x_1, then for x_2, etc., or else by noting that the 1's along the diagonal (and 0's above them) allow us to row reduce the matrix further to one have exactly three leading 1's in the first three columns.
(c) Inconsistent. Rows 1 and 2 give

$$x_1 = 0 \qquad x_1 = 1$$

(d) Inconclusive. Consider

$$\begin{bmatrix} 1 & a & b & c \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

If $a = b = 0$ and $c = 2$ the system is inconsistent: the matrix row reduces to one with a leading 1 in the last column. If $a = b = 0$ and $c = 1$, the system has infinitely many solutions: the matrix row reduces to one with a leading 1 in the first column only.

(14) Determine the values of a for which the system below has no solutions, many solutions, or one solution.

$$\begin{aligned} x + 2y + z &= 2 \\ 2x - 2y + 3z &= 1 \\ x + 2y - (a^2 - 3)z &= a \end{aligned}$$

Solution 14:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & -2 & 3 & 1 \\ 1 & 2 & 3 - a^2 & a \end{bmatrix} \xrightarrow{r_1 - r_3} \begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & -2 & 3 & 1 \\ 0 & 0 & a^2 - 2 & 2 - a \end{bmatrix}$$

$$\xrightarrow{2r_1 - r_2} \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 6 & -1 & 3 \\ 0 & 0 & a^2 - 2 & 2 - a \end{bmatrix}$$

$$\xrightarrow{\frac{1}{6}r_2} \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & -\frac{1}{6} & \frac{1}{2} \\ 0 & 0 & a^2 - 2 & 2 - a \end{bmatrix}$$

The row echelon form, and thus the set of solutions, now depends on whether $a^2 - 2 = 0$ or not: equivalently, whether $a = \pm\sqrt{2}$ or not. This gives us two cases:

$a = \pm\sqrt{2}$: In this case $2 - a \neq 0$, which means the row echelon matrix will end up having a leading 1 in the last column, resulting in an inconsistent system. There are no solutions in this case.

$a \neq \pm\sqrt{2}$: In this case the third column of the row echelon form will have a leading 1, and all variables are leading variables. Thus there is a unique solution in this case, obtained by back substitution.

(15) What condition must a,b, and c satisfy for the system to be consistent?

$$\begin{aligned} x + 3y + z &= a \\ -x - 2y + z &= b \\ 3x + 7y - z &= c \end{aligned}$$

Solution 15:

$$\begin{bmatrix} 1 & 3 & 1 & a \\ -1 & -2 & 1 & b \\ 3 & 7 & -1 & c \end{bmatrix} \xrightarrow{r_1 + r_2} \begin{bmatrix} 1 & 3 & 1 & a \\ 0 & 1 & 2 & a + b \\ 3 & 7 & -1 & c \end{bmatrix}$$

$$\xrightarrow{3r_1 - r_3} \begin{bmatrix} 1 & 3 & 1 & a \\ 0 & 1 & 2 & a + b \\ 0 & 2 & 4 & 3a - c \end{bmatrix}$$

$$\xrightarrow{2r_2 - r_3} \begin{bmatrix} 1 & 3 & 1 & a \\ 0 & 1 & 2 & a + b \\ 0 & 0 & 0 & a - 2b - c \end{bmatrix}$$

We see the system is consistent as long as $a - 2b - c = 0$, which guarantees there is no leading 1 in the last column.

11

(16) Solve for x, y, and z:

$$\frac{1}{x} + \frac{2}{y} - \frac{4}{z} = 1$$

$$\frac{2}{x} + \frac{3}{y} + \frac{8}{z} = 0$$

$$-\frac{1}{x} + \frac{9}{y} + \frac{10}{z} = 5$$

Solution 16:

Start by replacing variables. Let $X = \frac{1}{x}$, $Y = \frac{1}{y}$, and $Z = \frac{1}{z}$. Now we can solve the new system as we normally would.

$$\begin{bmatrix} 1 & 2 & -4 & 1 \\ 2 & 3 & 8 & 0 \\ -1 & 9 & 10 & 5 \end{bmatrix} \xrightarrow{r_1 + r_3} \begin{bmatrix} 1 & 2 & -4 & 1 \\ 2 & 3 & 8 & 0 \\ 0 & 11 & 6 & 6 \end{bmatrix}$$

$$\xrightarrow{2r_1 - r_2} \begin{bmatrix} 1 & 2 & -4 & 1 \\ 0 & 1 & -16 & 2 \\ 0 & 11 & 6 & 6 \end{bmatrix}$$

$$\xrightarrow{11r_2 - r_3} \begin{bmatrix} 1 & 2 & -4 & 1 \\ 0 & 1 & -16 & 2 \\ 0 & 0 & -182 & 16 \end{bmatrix}$$

Now solve the system for X, Y, Z:

$$X = -\frac{7}{13}$$
$$Y = \frac{54}{91}$$
$$Z = -\frac{8}{91}$$

Now we solve for the original x, y, and z:

$$x = -\frac{13}{7}$$
$$y = \frac{91}{54}$$
$$z = -\frac{91}{8}$$

(17) If A is a matrix with three rows and five columns, then what is the maximum possible number of leading 1's in its reduced row echelon form?
Solution 17:

The maximum possible number of leading 1's in the reduced row echelon form of a matrix with 3 rows and 5 columns is 3. It is indeed possible to obtain this maximal number, as the matrix

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

illustrates.

(18) If B is a matrix with three rows and six columns, then what is the maximum possible number of parameters in the general solution of the linear system with augmented matrix B?

Solution 18:

The matrix B corresponds to a linear system of 3 equations in 5 unknowns $x_1, x_2, \ldots, x_5$.

Let U be a row echelon form of B, and let k be the number of leading 1's *among the first five columns* of U. Then the number of parameters in the general solution to the system corresponding to B is $5 - k$. Thus we see, that the number of parameters is at most 5 (when $k = 5$).

This case is indeed possible, as the matrix $B = \underset{3 \times 6}{\mathbf{0}}$ illustrates.

(19) If C is a matrix with five rows and three columns, then what is the minimum possible number of rows of zeros in any row echelon form of C?

Solution 19:

If a row echelon form of A has r zero rows, then all other rows have leading 1's. Thus there are $5 - r$ leading 1's in this case. Since the number of leading 1's is at most 3 (the number of columns), we have $5 - r \le 3$. It follows that $2 \le r$, and thus there are at least 2 zero rows in a row echelon form of A. It is indeed possible to achieve this minimum number of zero rows, as the matrix

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

illustrates.

(20) True or False. If true, prove it; if false, give an explicit counterexample.

(a) If a matrix is in reduced row echelon form, then it is also in row echelon form.

(b) If an elementary row operation is applied to a matrix that is in row echelon form, the resulting matrix will still be in row echelon form.

(c) Every matrix has a unique row echelon form.

(d) A homogeneous linear system in n unknowns whose corresponding augmented matrix has a reduced row echelon form with r leading 1's has $n - r$ free variables.

(e) All leading 1's in a matrix in row echelon form must occur in different columns.

(f) If every column of a matrix in row echelon form has a leading 1, then all entries that are not leading 1's are zero.

(g) If a homogeneous linear system of n equations in n unknowns has a corresponding augmented matrix with a reduced row echelon form containing n leading 1's, then the linear system has only the trivial solution.

(h) If the reduced row echelon form of the augmented matrix for a linear system has a row of zeros, then the system must have infinitely many solutions.

(i) If a linear system has more unknowns than equations, then it must have infinitely many solutions.

Solution 20:

(a) True. A matrix in reduced row echelon form satisfies all conditions for row echelon form, plus an additional condition.

(b) False. Consider

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{r_2+r_1} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

(c) False. Consider $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, which is already in row echelon form. It is further row equivalent to $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

(d) True. Let A be the augmented matrix corresponding to the system, let U be a row echelon form of A, and let r be the number of pivot columns of U. Both matrices have $n+1$ columns, and the last column of each is a zero column. This means the system is guaranteed to be consistent (i.e., that the r pivot columns are among the first n columns of U), and hence that the number of free variables is $n-r$.

(e) True. This is part of the definition of row echelon form. The leading 1's must form a staircase pattern.

(f) False. Consider $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

(g) True. The system is consistent, since homogeneous systems are always consistent, and the number of free variables is $n - \#(\text{pivot columns}) = n - n = 0$. Hence there are no free variables, which means the system has a unique solution.

(h) False. The system could be inconsistent, in which case there are 0 solutions! Consider the system corresponding to the matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

(i) False. Consider the system

$$0x_1 + 0x_2 = 1$$

The system is clearly inconsistent. Hence it has 0 solutions, not infinitely many.

1.3-1.4: MATRIX ARITHMETIC

(1) Use the row or column method to quickly compute the following product:

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

Solution 1:

We did this in class using the column method. For kicks, I'll describe the row method here. Call the product AB.

Note that all the rows of A are all identical, and equal to $\begin{bmatrix} 1 & -1 & 1 & -1 & 1 \end{bmatrix}$. From the row method it follows that each row of AB is given by

$$\begin{bmatrix} 1 & -1 & 1 & -1 & 1 \end{bmatrix} B.$$

Thus the rows of AB are all identical, and the row method computes the product above by taking the corresponding alternating sum of the rows of B:

$$\begin{bmatrix} 1 & -1 & 1 & -1 & 1 \end{bmatrix} B = \begin{bmatrix} 2 & 2 & -1 & 4 \end{bmatrix}.$$

Thus AB is the the 5×4 matrix, all of whose rows are $\begin{bmatrix} 2 & 2 & -1 & 4 \end{bmatrix}$.

(2) Each of the 3×3 matrices B_i below performs a specific row operation when multiplying a

$3 \times n$ matrix $A = \begin{bmatrix} -\mathbf{r}_1- \\ -\mathbf{r}_2- \\ -\mathbf{r}_3- \end{bmatrix}$ on the left; i.e., the matrix $B_i A$ is the result of performing a

certain row operation on the matrix A.

Use the row method of matrix multiplication to decide what row operation each B_i represents.

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, B_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, B_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Solution 2:

The matrix B_1, when multiplied on the left, replaces the third row of A with $\mathbf{r}_3 - 2\mathbf{r}_2$.
The matrix B_2, when multiplied on the left, replaces the second row of A with $\frac{1}{2}\mathbf{r}_2$.
The matrix B_3, when multiplied on the left, swaps $\mathbf{r}_1$ and $\mathbf{r}_3$.

(3) For each part below write down the most general 3×3 matrix $A = [a_{ij}]$ satisfying the given condition (use letter names a, b, c,etc. for entries).
   (a) $a_{ij} = a_{ji}$ for all i, j.
   (b) $a_{ij} = -a_{ji}$ for all i, j
   (c) $a_{ij} = 0$ for $i \neq j$.

Solution 3:

$$A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}$$

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

(4) Let A be a square $n \times n$ matrix. We define its **square** to be the matrix $A^2 := AA$.
   Given a square $n \times n$ matrix B, we define a **square-root** of B to be a matrix A such that $A^2 = B$.
   (a) Find *all* square-roots of $\underset{2 \times 2}{\mathbf{0}}$.
   (b) Find *all* square-roots of I_2.
   Your answers for both parts should be a useful description of the entries of an arbitrary square-root matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, ideally some sort of parametrization.

Solution 4:

For both parts let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an arbitrary 2×2 matrix. Then

$$A^2 = \begin{bmatrix} a^2 + bc & ab + bd \\ ac + cd & d^2 + cb \end{bmatrix}.$$

Setting $A^2 = \mathbf{0}_{2\times 2}$ yields the *nonlinear* system of equations

$$a^2 + bc = 0$$
$$ab + bd = b(a+d) = 0$$
$$ac + cd = c(a+d) = 0$$
$$d^2 + cb = 0$$

This system has infinitely many solutions. For example, if we set $a = r$ and $d = -r$, then the second and third equations are automatically satisfied, and the first and four equations reduce to $r^2 + bc = 0$. If we set $b = s$ for any $s \neq 0$, then we must have $c = -r^2/s$. Thus any

$$A = \begin{bmatrix} r & s \\ -r^2/s & -r \end{bmatrix}, r \in \mathbb{R}, s \neq 0 \in \mathbb{R}$$

satisfies $A^2 = \mathbf{0}_{2\times 2}$.

Similarly, setting $A^2 = I_2$ yields the *nonlinear* system of equations

(1) $$a^2 + bc = 1$$
(2) $$ab + bd = b(a+d) = 0$$
(3) $$ac + cd = c(a+d) = 0$$
(4) $$d^2 + cb = 1$$

Let's try to write down *all* solutions to this system.

Equation 2 implies either $b = 0$, or $b \neq 0$ and $a + d = 0$. Consider the two cases separately.
Case 1: $b = 0$. Then (1) and (4) imply $a^2 = d^2 = 1$. Furthermore, (3) implies either $c = 0$ or $a + d = 0$. This leads to the following mutually exclusive possibilities for A:

$$\begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ r & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ r & 1 \end{bmatrix},$$

where r can be any nonzero real number.
Case 2: $b \neq 0$ and $a + d = 0$. Then $a = -d$. Set $a = r$, $d = -r$, $b = s$ where r and s are any two real numbers. Equation 1 (or 4) then implies $c = (1 - r^2)/s$, leading to solutions of the form

$$\begin{bmatrix} r & s \\ (1 - r^2)/s & -r \end{bmatrix}$$

where r and s are any real numbers.

The matrices described in the two cases comprise *all* solutions to $A^2 = I_2$. In particular, note that I_2 has infinitely many square-roots!

(5) Suppose A is $n \times n$ and satisfies $A^2 + 3A - 5I = \mathbf{0}_n$. Prove A is invertible by providing an explicit inverse.
Solution 5:

Set $A^{-1} = \frac{1}{5}(A + 3I)$, and show directly (using the above equality) that $AA^{-1} = A^{-1}A = I$.

(6) Show that in general $(AB)^r \neq A^r B^r$. Why does the argument from real number algebra fail?
Solution 6:

Easy to come up with counterexamples. I'll leave that part to you. Look for 2×2 examples.

Why doesn't this work? In a word: matrix multiplication is not commutative. Look at the $r = 3$ case for concreteness.

$$(AB)^3 = (AB)(AB)(AB) = ABABAB.$$

Since matrix multiplication is *not commutative*, I cannot rearrange the A's and B's above to get $AAABBB = A^3B^3$.

(7) Show that $(A + B)^2 = A^2 + 2AB + B^2$ if and only if $AB = BA$.

Solution 7:

Let's do a chain of equivalences, where at each step I do something valid (and reversible) to the given equation:

$$(A+B)^2 = A^2 + 2AB + B^2 \Leftrightarrow (A+B)(A+B) = A^2 + 2AB + B^2 \qquad \text{(def.)}$$
$$\Leftrightarrow A^2 + AB + BA + B^2 = A^2 + 2AB + B^2 \quad \text{(expand LHS carefully)}$$
$$\Leftrightarrow AB + BA = 2AB \qquad \text{(additive cancel)}$$
$$\Leftrightarrow BA = AB \qquad \text{(additive cancel)}$$

Following the chain of equivalences we see that

$$(A + B)^2 = A^2 + 2AB + B^2 \text{ if and only if } BA = AB,$$

as desired.

(8) Let

$$A = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{bmatrix}$$

$$D = \begin{bmatrix} 1 & 5 & 2 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{bmatrix}, \quad E = \begin{bmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{bmatrix}.$$

Compute the following matrices, or else explain why the given expression is not well defined.
(a) $(2D^T - E)A$
(b) $(4B)C + 2B$
(c) $B^T(CC^T - A^TA)$

Solution 8:

(a)

$$\left(\begin{bmatrix} 2 & -2 & 6 \\ 10 & 0 & 4 \\ 4 & 2 & 8 \end{bmatrix} - \begin{bmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{bmatrix} \right) \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -4 & -3 & 3 \\ 11 & -1 & 2 \\ 0 & 1 & 5 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -6 & -3 \\ 36 & 0 \\ 4 & 7 \end{bmatrix}$$

(b)

$$\begin{bmatrix} 16 & -4 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{bmatrix} + \begin{bmatrix} 8 & -2 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 60 & 12 \\ 24 & 8 & 40 \end{bmatrix} + \begin{bmatrix} 8 & -2 \\ 0 & 4 \end{bmatrix}$$

The matrix sum on the right is not defined. Thus the operation is not defined.

(c)
$$\begin{bmatrix} 4 & 0 \\ -1 & 2 \end{bmatrix} \left(\begin{bmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 4 & 1 \\ 2 & 5 \end{bmatrix} - \begin{bmatrix} 3 & -1 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix} \right)$$
$$= \begin{bmatrix} 4 & 0 \\ -1 & 2 \end{bmatrix} \left(\begin{bmatrix} 21 & 17 \\ 17 & 35 \end{bmatrix} - \begin{bmatrix} 11 & -1 \\ -1 & 5 \end{bmatrix} \right) = \begin{bmatrix} 40 & 72 \\ 26 & 42 \end{bmatrix}$$

(9) Let
$$A = \begin{bmatrix} 3 & -2 & 7 \\ 6 & 5 & 4 \\ 0 & 4 & 9 \end{bmatrix}, \quad B = \begin{bmatrix} 6 & -2 & 4 \\ 0 & 1 & 3 \\ 7 & 7 & 5 \end{bmatrix}.$$
Compute the following using either the row or column method of matrix multiplication:
(a) the first column of AB;
(b) the second row of BB;
(c) the third column of AA.

Solution 9:

(a) Using expansion by columns, the first column of AB is given by A times the first column of B. We compute
$$\begin{bmatrix} 3 & -2 & 7 \\ 6 & 5 & 4 \\ 0 & 4 & 9 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 7 \end{bmatrix} = \begin{bmatrix} 67 \\ 64 \\ 63 \end{bmatrix}$$

(b) Using expansion by rows, the second row of BB is given by the second row of B times B. We compute
$$\begin{bmatrix} 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 6 & -2 & 4 \\ 0 & 1 & 3 \\ 7 & 7 & 5 \end{bmatrix} = \begin{bmatrix} 21 & 22 & 18 \end{bmatrix}$$

(c) Using expansion by columns, the third column of AA is given by A times the third column of A. We compute
$$A \begin{bmatrix} 7 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 7 \\ 6 & 5 & 4 \\ 0 & 4 & 9 \end{bmatrix} \begin{bmatrix} 7 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 76 \\ 98 \\ 97 \end{bmatrix}$$

(10) Find all values of k, if any, that satisfy the equation:
$$\begin{bmatrix} 2 & 2 & k \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ k \end{bmatrix} = 0$$

Solution 10:
Since:
$$\begin{bmatrix} 2 & 2 & k \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ k \end{bmatrix} = \begin{bmatrix} 6 & 4+3k & 6+k \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ k \end{bmatrix} = \begin{bmatrix} k^2 + 12k + 20 \end{bmatrix}$$
Set $k^2 + 12k + 20 = 0$ and solve for k.
$$0 = k^2 + 12k + 20 = (k+2)(k+10) \rightarrow k = -2, -10$$

18

(11) Let $\mathbf{0}_{2\times 2}$ denote the 2×2 matrix, each of whose entries is zero.

(a) Is there a 2×2 matrix A such that $A \neq \mathbf{0}$ and $AA = \mathbf{0}$?

(b) Is there a 2×2 matrix A such that $A \neq \mathbf{0}$ and $AA = A$?

Solution 11:

Yes, there are many examples. Here is one:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, AA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

(b) Again yes, here is one example:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, AA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(12) Answer true or false. If true, provide a proof; if false, provide an explicit counterexample.

(a) If B has a column of zeros, then so does AB if this product is defined.

(b) If B has a column of zeros, then so does BA if this product is defined.

(c) If A, B, C are *nonzero* $n \times n$ matrices and $AC = BC$, then $A = B$. (Remember: a matrix is nonzero if it is not the zero matrix.)

(d) If $AB + BA$ is well-defined, then A and B are square matrices of the same size.

Solution 12:

(a) True. Use the column method of matrix multiplication. Suppose column j of B is a zero column, and denote this column $\mathbf{b}_j$. Then according to the column method of multiplication, the j-th column of the product AB is given by $A\mathbf{b}_j$. But clearly, $A\mathbf{b}_j$ is the zero column vector, since $\mathbf{b}_j$ contains nothing but zeros. Thus the j-th column of AB is a zero column.

(b) False. Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Then

$$BA = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

(c) False. Take $A = C = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$. Then $AC = BC = \mathbf{0}_{2\times 2}$, but $A \neq B$.

(d) True. Suppose A is $m \times n$ and B is $r \times s$. For AB to be well defined, we must have $n = r$. Thus B is in fact $n \times s$. For BA to be well-defined we need $s = m$. Thus B is $n \times m$, and hence AB is $m \times m$ and BA is $n \times n$. For the sum of these matrices to be equal, we need $m = n$. Thus A and B are both $n \times n$ matrices!

(13) Let $A = \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix}$.

Compute A^3, A^{-3} and $A^2 - 2A + I$.

Solution 13:

19

$$A^3 = AAA$$

$$= \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 0 \\ 12 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 8 & 0 \\ 28 & 1 \end{bmatrix}$$

$$A^{-3} = (A^{-1})^3 \qquad \text{(by definition)}$$

$$= A^{-1}A^{-1}A^{-1}$$

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

$$= \begin{bmatrix} \frac{1}{4} & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ -2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{8} & 0 \\ -\frac{7}{2} & 1 \end{bmatrix}$$

$$A^2 - 2A + I = \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix} - 2\begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 0 \\ 12 & 1 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 8 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}$$

(14) Let $A = \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix}$. Compute $p(A)$, where $p(x) = x^3 - 2x + 5$.
Solution 14:

Replace x with A and the constant term 5 with $5I$:

$$\begin{bmatrix} 8 & 0 \\ 28 & 1 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 8 & 2 \end{bmatrix} + 5\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 20 & 4 \end{bmatrix}$$

(15) Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Find all choices of a, b, c, and d for which $AC = CA$.
Solution 15:

$$AC = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & 0 \\ d & 0 \end{bmatrix}$$

$$CA = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix}$$

Thus A and C commute when: $a = d$ and $b = 0$

(16) Consider the system
$$2x_1 - 2x_2 = 4$$
$$x_1 + 4x_2 = 4$$

(a) Find a matrix A and column vector $\mathbf{b}$ such that solutions (x_1, x_2) to the system correspond to solutions $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ to the matrix equation $A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}$.

(b) Now solve your matrix equation in (a) algebraically using inverses.

Solution 16:

We let $A = \begin{bmatrix} 2 & -2 \\ 1 & 4 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$. Using our formula for inverses of 2×2 matrices, we see that A is invertible and $A^{-1} = \begin{bmatrix} 2/5 & 1/5 \\ -1/10 & 1/5 \end{bmatrix}$. Then we solve the matrix equation as follows:

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A^{-1} \begin{bmatrix} 4 \\ 4 \end{bmatrix} \qquad \text{(mult. both sides on left by } A^{-1})$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2/5 & 1/5 \\ -1/10 & 1/5 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} 12/5 \\ 2/5 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 12 \\ 2 \end{bmatrix}.$$

(17) Find at least nine solutions to $A^2 = I_3$
Solution 17:

Matrices of the form
$$\begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 \end{bmatrix}$$
comprise eight of the desired nine solutions. For a ninth solution take
$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(18) Can a matrix with two identical rows or two identical columns have an inverse?
Solution 18:

No. We see this using the row or column method of multiplication.

For example, suppose A has two identical columns; that means we have $\mathbf{a}_i = \mathbf{a}_j$ for two columns $\mathbf{a}_i$ and $\mathbf{a}_j$. Now suppose by contradiction that B is the inverse of A. Then $BA = I_n$. The column method tells us that the the the i- and j-th columns of BA are $B\mathbf{a}_i$ and $B\mathbf{a}_j$; but then these two columns are equal, a contradiction since I_n does not have any identical columns.

A similar argument applies if A has two identical rows.

(19) Assuming A, B, C and D are all invertible square matrixes, solve the following equation for D in terms of A, B and C.

$$ABC^T DBA^T C = AB^T.$$

Solution 19:

$$
\begin{aligned}
ABC^T DBA^T C &= AB^T \\
(ABC^T)^{-1} ABC^T DBA^T C &= (ABC^T)^{-1} AB^T \\
DBA^T C &= (ABC^T)^{-1} AB^T \\
DBA^T C (BA^T C)^{-1} &= (ABC^T)^{-1} AB^T (BA^T C)^{-1} \\
D &= (ABC^T)^{-1} AB^T (BA^T C)^{-1}
\end{aligned}
$$

Side note: You can also work with one matrix at a time and multiply on the correct side of both sides by the inverse. It takes longer, but gives the same result.

(20) Prove: $A + B = B + A$, where A and B are any $m \times n$ matrices. (This is Theorem 1.18 (a).)
Solution 20:

Let $A = [a_{ij}]$, $B = [b_{ij}]$ and $(A + B)_{ij}$ denote the ij entry of $A + B$. Show that $(A + B)_{ij} = (B + A)_{ij}$.

$$
\begin{aligned}
(A + B)_{ij} &= a_{ij} + b_{ij} \\
&= b_{ij} + a_{ij} \\
&= (B + A)_{ij}
\end{aligned}
$$

Note that line two follows from the fact that addition of real numbers is commutative

(21) Prove: $(AB)^T = B^T A^T$, where A is any $m \times n$ matrix and B is any $n \times r$ matrix.
Solution 21:

Both the LHS and RHS are $n \times n$ matrices. It remains to show the ij-th entries are all equal: that is, $((AB)^T)_{ij} = (B^T A^T)_{ij}$ for all i, j.

$$((AB)^T)_{ij} = (AB)_{ji} \qquad\qquad \text{(def. of transp.)}$$

$$= \sum_{\ell=1}^{n} (A)_{j\ell}(B)_{\ell j} \qquad\qquad \text{(def. of mult.)}$$

$$= \sum_{\ell=1}^{n} (A^T)_{\ell j}(B^T)_{i\ell} \qquad\qquad \text{(def. of transp.)}$$

$$= \sum_{\ell=1}^{n} (B^T)_{i\ell}(A^T)_{\ell j} \qquad\qquad \text{(real number commutativity)}$$

$$= (B^T A^T)_{ij}.$$

(22) Answer true or false. If true, provide a proof; if false, give an explicit counterexample.
   (a) If A and B are invertible $n \times n$ matrices, then $(AB)^{-1} = A^{-1}B^{-1}$.
   (b) A square matrix containing a row or column of zeros cannot be invertible.
   (c) If A and B are invertible $n \times n$ matrices, then $A + B$ is invertible.
   (d) If A is invertible, then A^T is invertible.

Solution 22:

(a) False. Let

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} : B = \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix}$$

$$A^{-1}B^{-1} = \begin{bmatrix} -5/2 & 2 \\ 4 & -3 \end{bmatrix}$$

$$(AB)^{-1} = B^{-1}A^{-1} = \begin{bmatrix} -2 & 3 \\ 5/2 & -7/2 \end{bmatrix}$$

(b) True. If A has a row of zeroes, then AB has a row of zeroes; it follows that there cannot be a matrix B such that $AB = I$, and thus that A is not invertible.
   Similarly, if A has a column of zeroes, then BA has a column of zeroes, and hence cannot be equal to I. Thus A cannot be invertible in this case.
(c) False. Take $A = I_2$ and $B = -I_2$. Both are invertible, yet $A + B = \mathbf{0}$ is not invertible.
(d) True. In fact, we have a theorem that says in this case that $(A^T)^{-1} = (A^{-1})^T$.

1.5-1.6: INVERSES, DIAGONAL AND SYMMETRIC MATRICES

(1) Let

$$B = \begin{bmatrix} 8 & 1 & 5 \\ 2 & -7 & -1 \\ 3 & 4 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 8 & 1 & 5 \\ -6 & 21 & 3 \\ 3 & 4 & 1 \end{bmatrix}, \quad F = \begin{bmatrix} 8 & 1 & 5 \\ 8 & 1 & 1 \\ 3 & 4 & 1 \end{bmatrix},$$

For each part below, find an explicit *elementary matrix* E satisfying the given matrix equation.
   Justify your answer by first explaining what row operation E should perform.

(a) $EB = D$
(b) $ED = B$
(c) $EB = F$
(d) $EF = B$.

Solution 1:

(a) To obtain D from B we must scale the second row of B by -3. Thus

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(c) To obtain F from B we must replace the second row of B with $r_2 + 2r_3$. Thus

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

(2) For each matrix below use the inversion algorithm to find the inverse, if it exists.

(a) $A = \begin{bmatrix} 1/5 & 1/5 & -2/5 \\ 1/5 & 1/5 & 1/10 \\ 1/5 & -4/5 & 1/10 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1/5 & 1/5 & -2/5 \\ 2/5 & -3/5 & -3/10 \\ 1/5 & -4/5 & 1/10 \end{bmatrix}$

(c) $A = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & 3 & 0 \\ 2 & 1 & 5 & -3 \end{bmatrix}$

Solution 2:

(a)

$$\left[\begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 1/5 & 1/5 & 1/10 & 0 & 1 & 0 \\ 1/5 & -4/5 & 1/10 & 0 & 0 & 1 \end{array}\right] \xrightarrow{r_1-r_2} \left[\begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & 1 & -1 & 0 \\ 1/5 & -4/5 & 1/10 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{r_1-r_3} \left[\begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & 1 & -1 & 0 \\ 0 & 1 & -1/2 & 1 & 0 & -1 \end{array}\right]$$

$$\xrightarrow{5r_1} \left[\begin{array}{ccc|ccc} 1 & 1 & -2 & 5 & 0 & 0 \\ 0 & 0 & -1/2 & 1 & -1 & 0 \\ 0 & 1 & -1/2 & 1 & 0 & -1 \end{array}\right]$$

$$\xrightarrow{r_2\leftrightarrow r_3} \left[\begin{array}{ccc|ccc} 1 & 1 & -2 & 5 & 0 & 0 \\ 0 & 1 & -1/2 & 1 & 0 & -1 \\ 0 & 0 & -1/2 & 1 & -1 & 0 \end{array}\right]$$

$$\xrightarrow{r_1-r_2} \left[\begin{array}{ccc|ccc} 1 & 0 & -3/2 & 4 & 0 & 1 \\ 0 & 1 & -1/2 & 1 & 0 & -1 \\ 0 & 0 & -1/2 & 1 & -1 & 0 \end{array}\right]$$

$$\xrightarrow{-2r_3} \left[\begin{array}{ccc|ccc} 1 & 0 & -3/2 & 4 & 0 & 1 \\ 0 & 1 & -1/2 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 & 2 & 0 \end{array}\right]$$

$$\xrightarrow{r_2+\frac{1}{2}r_3} \left[\begin{array}{ccc|ccc} 1 & 0 & -3/2 & 4 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -2 & 2 & 0 \end{array}\right]$$

$$\xrightarrow{r_1+\frac{3}{2}r_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 3 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -2 & 2 & 0 \end{array}\right]$$

(b)

$$\left[\begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 2/5 & -3/5 & -3/10 & 0 & 1 & 0 \\ 1/5 & -4/5 & 1/10 & 0 & 0 & 1 \end{array}\right] \xrightarrow{2r_1-r_2} \left[\begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 0 & 1 & -1/2 & 2 & -1 & 0 \\ 1/5 & -4/5 & 1/10 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{r_1-r_3} \left[\begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 0 & 1 & -1/2 & 2 & -1 & 0 \\ 0 & 1 & -1/2 & 1 & 0 & -1 \end{array}\right]$$

$$\xrightarrow{r_2-r_3} \left[\begin{array}{ccc|ccc} 1/5 & 1/5 & -2/5 & 1 & 0 & 0 \\ 0 & 1 & -1/2 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right]$$

Since we have row reduced our matrix to a matrix with a row of zeros, we can conclude that our original matrix was singular. Thus no inverse exists.

(c)

$$\left[\begin{array}{cccc|cccc} 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 2 & 1 & 5 & -3 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{r_1 \leftrightarrow r_2} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 2 & 1 & 5 & -3 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{r_4 - 2r_1} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -5 & 5 & 0 & 2 & 0 & -1 \end{array}\right]$$

$$\xrightarrow{-r_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 & 0 & 0 & -1 & 0 \\ 0 & -1 & -5 & 5 & 0 & 2 & 0 & -1 \end{array}\right]$$

$$\xrightarrow{r_4 + r_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -8 & 5 & 0 & 2 & -1 & -1 \end{array}\right]$$

$$\xrightarrow{r_2 \leftrightarrow r_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -3 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -8 & 5 & 0 & 2 & -1 & -1 \end{array}\right]$$

$$\xrightarrow{r_4 + 4r_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -3 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 4 & 2 & -1 & -1 \end{array}\right]$$

$$\xrightarrow{\frac{1}{2}r_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -3 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 4 & 2 & -1 & -1 \end{array}\right]$$

$$\xrightarrow{r_2 + 3r_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 3/2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 4 & 2 & -1 & -1 \end{array}\right]$$

$$\xrightarrow{\frac{1}{5}r_4} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 3/2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4/5 & 2/5 & -1/5 & -1/5 \end{array}\right]$$

$$\xrightarrow{r_1 - r_4} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -4/5 & 3/5 & 1/5 & 1/5 \\ 0 & 1 & 0 & 0 & 3/2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4/5 & 2/5 & -1/5 & -1/5 \end{array}\right]$$

(3) For each matrix below find the inverse using the inverse algorithm.
Assume $k_1, k_2, k_3, k_4, k \neq 0$.

(a) $A = \begin{bmatrix} 0 & 0 & 0 & k_1 \\ 0 & 0 & k_2 & 0 \\ 0 & k_3 & 0 & 0 \\ k_4 & 0 & 0 & 0 \end{bmatrix}$

(b) $A = \begin{bmatrix} k & 0 & 0 & 0 \\ 1 & k & 0 & 0 \\ 0 & 1 & k & 0 \\ 0 & 0 & 1 & k \end{bmatrix}$

Solution 3:

(a)

$$\left[\begin{array}{cccc|cccc} 0 & 0 & 0 & k_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & k_2 & 0 & 0 & 1 & 0 & 0 \\ 0 & k_3 & 0 & 0 & 0 & 0 & 1 & 0 \\ k_4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{r_1 \leftrightarrow r_4} \left[\begin{array}{cccc|cccc} k_4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & k_2 & 0 & 0 & 1 & 0 & 0 \\ 0 & k_3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & k_1 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{r_2 \leftrightarrow r_3} \left[\begin{array}{cccc|cccc} k_4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & k_3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & k_2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & k_1 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{\frac{1}{k_4} r_1} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1/k_4 \\ 0 & k_3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & k_2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & k_1 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{\frac{1}{k_3} r_2} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1/k_4 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1/k_3 & 0 \\ 0 & 0 & k_2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & k_1 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{\frac{1}{k_2} r_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1/k_4 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1/k_3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/k_2 & 0 & 0 \\ 0 & 0 & 0 & k_1 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{\frac{1}{k_1} r_4} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1/k_4 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1/k_3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/k_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1/k_1 & 0 & 0 & 0 \end{array}\right]$$

Thus $A^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1/k_4 \\ 0 & 0 & 1/k_3 & 0 \\ 0 & 1/k_2 & 0 & 0 \\ 1/k_1 & 0 & 0 & 0 \end{bmatrix}$

(b)

$\left[\begin{array}{cccc|cccc} k & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & k & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & k & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & k & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\frac{1}{k}r_1} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/k & 0 & 0 & 0 \\ 1 & k & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & k & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & k & 0 & 0 & 0 & 1 \end{array}\right]$

$\xrightarrow{r_1 - r_2} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/k & 0 & 0 & 0 \\ 0 & -k & 0 & 0 & 1/k & -1 & 0 & 0 \\ 0 & 1 & k & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & k & 0 & 0 & 0 & 1 \end{array}\right]$

$\xrightarrow{-\frac{1}{k}r_2} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/k & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/k^2 & 1/k & 0 & 0 \\ 0 & 1 & k & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & k & 0 & 0 & 0 & 1 \end{array}\right]$

$\xrightarrow{r_2 - r_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/k & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/k^2 & 1/k & 0 & 0 \\ 0 & 0 & -k & 0 & -1/k^2 & 1/k & -1 & 0 \\ 0 & 0 & 1 & k & 0 & 0 & 0 & 1 \end{array}\right]$

$\xrightarrow{-\frac{1}{k}r_3} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/k & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/k^2 & 1/k & 0 & 0 \\ 0 & 0 & 1 & 0 & 1/k^3 & -1/k^2 & 1/k & 0 \\ 0 & 0 & 1 & k & 0 & 0 & 0 & 1 \end{array}\right]$

$\xrightarrow{r_3 - r_4} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/k & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/k^2 & 1/k & 0 & 0 \\ 0 & 0 & 1 & 0 & 1/k^3 & -1/k^2 & 1/k & 0 \\ 0 & 0 & 0 & -k & 1/k^3 & -1/k^2 & 1/k & -1 \end{array}\right]$

$\xrightarrow{-\frac{1}{k}r_4} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/k & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/k^2 & 1/k & 0 & 0 \\ 0 & 0 & 1 & 0 & 1/k^3 & -1/k^2 & 1/k & 0 \\ 0 & 0 & 0 & 1 & -1/k^4 & 1/k^3 & -1/k^2 & 1/k \end{array}\right]$

Thus $A^{-1} = \begin{bmatrix} 1/k & 0 & 0 & 0 \\ -1/k^2 & 1/k & 0 & 0 \\ 1/k^3 & -1/k^2 & 1/k & 0 \\ -1/k^4 & 1/k^3 & -1/k^2 & 1/k \end{bmatrix}.$

(4) Find all values of c, if any, making the matrix $A = \begin{bmatrix} c & 1 & 0 \\ 1 & c & 1 \\ 0 & 1 & c \end{bmatrix}$ invertible.

Solution 4:

The matrix is row equivalent to $U = \begin{bmatrix} 1 & c & 1 \\ 0 & 1 & c \\ 0 & 0 & -2c + c^3 \end{bmatrix}.$

We know A is invertible if and only if U is invertible. Furthermore U is invertible if and

only if $-2c + c^3 = c(c^2 - 2) \neq 0$. This is the case if and only if $c \neq 0$ and $c \neq \pm\sqrt{2}$. Thus A is invertible if and only if $c \neq 0, \pm\sqrt{2}$.

(5) For each A below, express A and A^{-1} as products of elementary matrices.

(a) $A = \begin{bmatrix} 1 & 0 \\ -5 & 2 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

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

Solution 5:

(a)

$$A = \begin{bmatrix} 1 & 0 \\ -5 & 2 \end{bmatrix} \xrightarrow{r_2 + 5r_1} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$
$$\xrightarrow{\frac{1}{2}r_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus we have elementary matrices $E_1 = \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$ such that:

$E_2 E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix} = A^{-1}$ and $E_1^{-1} E_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -5/2 & -1 \end{bmatrix} = A$

(b)

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{r_2 - r_1} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{r_2 \leftrightarrow r_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right]$$

$$\xrightarrow{r_2 - r_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right]$$

$$\xrightarrow{r_1 - r_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right]$$

Then the elementary matrices are:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}, E_4 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

And

$A^{-1} = E_4 E_3 E_2 E_1$ and $A = E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1}$

(c)

$$
\left[\begin{array}{ccc|ccc}
1 & 1 & 1 & 1 & 0 & 0 \\
1 & -1 & 0 & 0 & 1 & 0 \\
1 & 0 & -1 & 0 & 0 & 1
\end{array}\right]
\xrightarrow{r_2 - r_1}
\left[\begin{array}{ccc|ccc}
1 & 1 & 1 & 1 & 0 & 0 \\
0 & -2 & -1 & -1 & 1 & 0 \\
1 & 0 & -1 & 0 & 0 & 1
\end{array}\right]
$$

$$
\xrightarrow{r_3 - r_1}
\left[\begin{array}{ccc|ccc}
1 & 1 & 1 & 1 & 0 & 0 \\
0 & -2 & -1 & -1 & 1 & 0 \\
0 & -1 & -2 & -1 & 0 & 1
\end{array}\right]
$$

$$
\xrightarrow{-\frac{1}{2}r_2}
\left[\begin{array}{ccc|ccc}
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1/2 & 1/2 & -1/2 & 0 \\
0 & -1 & -2 & -1 & 0 & 1
\end{array}\right]
$$

$$
\xrightarrow{r_3 + r_1}
\left[\begin{array}{ccc|ccc}
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1/2 & 1/2 & -1/2 & 0 \\
0 & 0 & -3/2 & -1/2 & -1/2 & 1
\end{array}\right]
$$

$$
\xrightarrow{-2/3 r_3}
\left[\begin{array}{ccc|ccc}
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1/2 & 1/2 & -1/2 & 0 \\
0 & 0 & 1 & 1/3 & 1/3 & -2/3
\end{array}\right]
$$

$$
\xrightarrow{r_2 - 1/2 r_3}
\left[\begin{array}{ccc|ccc}
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 1/3 & -2/3 & 2/3 \\
0 & 0 & 1 & 1/3 & 1/3 & -2/3
\end{array}\right]
$$

$$
\xrightarrow{r_1 - r_3}
\left[\begin{array}{ccc|ccc}
1 & 1 & 0 & 2/3 & -1/3 & 2/3 \\
0 & 1 & 0 & 1/3 & -2/3 & 2/3 \\
0 & 0 & 1 & 1/3 & 1/3 & -2/3
\end{array}\right]
$$

$$
\xrightarrow{r_1 - r_2}
\left[\begin{array}{ccc|ccc}
1 & 0 & 0 & 1/3 & 1/3 & 0 \\
0 & 1 & 0 & 1/3 & -2/3 & 2/3 \\
0 & 0 & 1 & 1/3 & 1/3 & -2/3
\end{array}\right]
$$

The corresponding elementary matrices are

$$
E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad
E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \quad
E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix},
$$

$$
E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad
E_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{2}{3} \end{bmatrix}, \quad
E_6 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix},
$$

$$
E_7 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad
E_8 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},
$$

We have $A^{-1} = E_8 E_7 E_6 E_5 E_4 E_3 E_2 E_1$ and $A = E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1} E_5^{-1} E_6^{-1} E_7^{-1} E_8^{-1}$.

(6) Let

$$
A = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 1 & 0 \\ 3 & 0 & -1 \end{bmatrix}, \quad
B = \begin{bmatrix} 6 & 9 & 4 \\ -5 & -1 & 0 \\ -1 & -2 & -1 \end{bmatrix}.
$$

Produce B from A using row operations, the use this sequence of row operations to find C such that $CA = B$.

Solution 6:

$$
\begin{bmatrix} 2 & 1 & 0 \\ -1 & 1 & 0 \\ 3 & 0 & -1 \end{bmatrix}
\xrightarrow{r_3 - 2r_1}
\begin{bmatrix} 2 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & -2 & -1 \end{bmatrix}
$$

$$
\xrightarrow{r_2 - 2r_1}
\begin{bmatrix} 2 & 1 & 0 \\ -5 & -1 & 0 \\ -1 & -2 & -1 \end{bmatrix}
$$

$$
\xrightarrow{r_1 - 4r_3}
\begin{bmatrix} 6 & 9 & 4 \\ -5 & -1 & 0 \\ -1 & -2 & -1 \end{bmatrix}
$$

$$
C = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}
=
\begin{bmatrix} 9 & 0 & -4 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}
$$

(7) Let A be $n \times n$. Show that the following are equivalent by proving a "cycle of implications".
   (i) A is invertible.
   (ii) For any column vector $\mathbf{b}$ the matrix equation $A\mathbf{x} = \mathbf{b}$ has a unique solution.
   (iii) For any column vector $\mathbf{b}$ the matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution.
This completes the proof of the invertibility theorem in its current form.

Solution 7:

$(i) \Rightarrow (ii)$. Suppose A is invertible. Given any $\underset{n \times 1}{\mathbf{b}}$, we have

$$A\mathbf{x} = \mathbf{b} \Rightarrow A^{-1}A = A^{-1}\mathbf{b}$$

$$\Rightarrow I\mathbf{x} = A^{-1}\mathbf{b}$$

$$\Rightarrow \mathbf{x} = A^{-1}\mathbf{b}$$

This shows that for any $\mathbf{b}$, there is a unique solution to the equation $A\mathbf{x} = \mathbf{b}$: namely, $\mathbf{x} = A^{-1}\mathbf{b}$.

$(ii) \Rightarrow (iii)$. Easy.

$(iii) \Rightarrow (i)$.

Assume $A\mathbf{x} = \mathbf{b}$ has a solution for any vector $\mathbf{b}$. For each $1 \le j \le n$ let $\mathbf{e}_j$ be the j-th column of the identity matrix I_n: i.e., $\mathbf{e}_j$ has a 1 in the j-th row, and 0's elsewhere. By assumption for each such $\mathbf{e}_j$ I can find a vector $\mathbf{c}_j$ such that $A\mathbf{c}_j = \mathbf{e}_j$.

I claim the matrix

$$
C = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & \cdots & | \end{bmatrix}
$$

31

is the inverse of A.

Indeed, using the column method of matrix multiplication we see that

$$AC = A \begin{bmatrix} | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & \cdots & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & \cdots & | \\ A\mathbf{c}_1 & A\mathbf{c}_2 & \cdots & A\mathbf{c}_n \\ | & | & \cdots & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & \cdots & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \\ | & | & \cdots & | \end{bmatrix}$$

$$= I_n,$$

where the last equality follows by construction: i.e, we picked $\mathbf{e}_j$ precisely to be the j-th column I_n !!

Note that this only shows $AC = I_n$; i.e., only that C is a *right-inverse* of A. Luckily we proved a corollary in the lecture notes that says you are a right inverse if and only if you are a left inverse. This allows us to claim that we also have $CA = I_n$, and hence that $C = A^{-1}$.

(You might object that the corollary cited was proved using the invertibility theorem, which the desired statement is a part of! However, if you look closely at the proof of the corollary, you see that it only uses statement (b) from the invertibility theorem.)

(8) Let A be $n \times n$ and suppose A has two identical columns. Use an appropriate statement from the invertibility theorem to show A is not invertible.
Solution 8:

Let $\mathbf{c}_i$ be the columns of A and suppose that columns $\mathbf{c}_i$ and $\mathbf{c}_j$ are equal. Define $\mathbf{x}$ to be the column vector with a 1 in the i-th row, a -1 in the j-th column, and 0's elsewhere. Then the column method of matrix multiplication tells us that

$$A\mathbf{x} = \mathbf{c}_i - \mathbf{c}_j = \mathbf{0}.$$

Since $\mathbf{x}$ is nontrivial, we have found a nontrivial solution to $A\mathbf{x} = \mathbf{0}$. By the invertibility theorem we now conclude A is not invertible.

(9) Prove that if A is an invertible matrix and B is row equivalent to A, then B is also invertible.
Solution 9:

Given that B is row equivalent to A, we can write B as the product of elementary matrices and A.

$$B = E_n \ldots E_1 A$$

By Theorem 1.5.2, all of the elementary matrices are invertible. And we know that the product of any number of invertible matrices is also invertible. Thus B is also invertible.

(10) Determine conditions on b_i's to guarantee a consistent system.

$$\begin{aligned} x_1 - 2x_2 - x_3 &= b_1 \\ -4x_1 + 5x_2 + 2x_3 &= b_2 \\ -4x_1 + 7x_2 + 4x_3 &= b_3 \end{aligned}$$

Solution 10:

$$
\begin{bmatrix} 1 & -2 & -1 & b_1 \\ -4 & 5 & 2 & b_2 \\ -4 & 7 & 4 & b_3 \end{bmatrix} \xrightarrow{4r_1+r_2} \begin{bmatrix} 1 & -2 & -1 & b_1 \\ 0 & -3 & -2 & 4b_1+b_2 \\ -4 & 7 & 4 & b_3 \end{bmatrix}
$$

$$
\xrightarrow{4r_1+r_3} \begin{bmatrix} 1 & -2 & -1 & b_1 \\ 0 & -3 & -2 & 4b_1+b_2 \\ 0 & -1 & 0 & 4b_1+b_3 \end{bmatrix}
$$

$$
\xrightarrow{r_2 \leftrightarrow r_3} \begin{bmatrix} 1 & -2 & -1 & b_1 \\ 0 & -1 & 0 & 4b_1+b_3 \\ 0 & -3 & -2 & 4b_1+b_2 \end{bmatrix}
$$

$$
\xrightarrow{-r_2} \begin{bmatrix} 1 & -2 & -1 & b_1 \\ 0 & 1 & 0 & -4b_1-b_3 \\ 0 & -3 & -2 & 4b_1+b_2 \end{bmatrix}
$$

$$
\xrightarrow{3r_2+r_3} \begin{bmatrix} 1 & -2 & -1 & b_1 \\ 0 & 1 & 0 & -4b_1-b_3 \\ 0 & 0 & -2 & -8b_1+b_2-3b_3 \end{bmatrix}
$$

$$
\xrightarrow{-1/2r_3} \begin{bmatrix} 1 & -2 & -1 & b_1 \\ 0 & 1 & 0 & -4b_1-b_3 \\ 0 & 0 & 1 & 4b_1-1/2b_2+3/2b_3 \end{bmatrix}
$$

Thus, this system is consistent for all values of b_i.

(11) Let $A\mathbf{x} = \mathbf{0}$ be a homogeneous system of n linear equations in n unknowns, and let Q be an invertible n x n matrix.
Prove that $A\mathbf{x} = \mathbf{0}$ has only the trivial solution if and only if $(QA)\mathbf{x} = \mathbf{0}$ has only the trivial solution.
Solution 11:

Claim: $A\mathbf{x} = \mathbf{0} \Leftrightarrow QA\mathbf{x} = \mathbf{0}$. From this claim it follows that the set of solutions to $A\mathbf{x} = \mathbf{0}$ is exactly equal to the set of solutions to $QA\mathbf{x} = \mathbf{0}$, and hence that the former equation has a unique solution if and only if the latter equation has a unique solution.
Proof of claim:

$$
QA\mathbf{x} = \mathbf{0} \Leftrightarrow Q^{-1}QA\mathbf{x} = Q^{-1}\mathbf{0}
$$
$$
\Leftrightarrow A\mathbf{x} = \mathbf{0}.
$$

(12) Answer true or false. If true, provide a proof; if false, give an explicit counterexample.
  (a) The product of two elementary matrices of the same size must be an elementary matrix.
  (b) If A is row equivalent to B, and B is row equivalent to C, then A is row equivalent to C.
  (c) A is a singular $n \times n$ matrix, then the linear system $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions.
  (d) If A is invertible, then if the second row of A is replaced with the second row plus the first row, the resulting matrix is invertible.
  (e) If A is a square matrix, and if the linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution, then the linear system $A\mathbf{x} = \mathbf{c}$ also must have a unique solution.
  (f) If A and B are row equivalent matrices, then the linear systems $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have the same solution set.

(g) Let A and B be n x n matrices. If A or B (or both) are not invertible, then neither is AB.

Solution 12:

(a) False. Consider:

$$\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 0 & -3 \end{bmatrix}$$

The matrix on the right is not an elementary matrix, as it performs TWO row operations on a matrix: (1) $r_1 \to r_1 + 4r_2$, and (2) $r_2 \to -3r_2$.

(b) True. Simply apply row operations to first transorm A to B, then B to C.

(c) True. Singular means noninvertible, btw. The invertibility theorem tells us that $A\mathbf{x} = \mathbf{0}$ has more than one solution. Since linear systems have either 0, 1 or infinitely-many solutions, we conclude that it has infinitely many solutions.

(d) True. The new matrix B is just EA, where E is an elementary matrix. Since elementary matrices are invertible, $B = EA$ is the product of invertible matrices, hence invertible.

(e) False. Take $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Then the equation $A\mathbf{x} = \mathbf{b}$ has a solution: namely $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. But the equation $A\mathbf{x} = \mathbf{c}$ does not have a solution.

(f) True. Since A and B are row equivalent, we can $B = E_r \cdots E_1 A = QA$, where $Q = E_r \cdots E_1$ is invertible, being the product of invertible matrices.

I claim $A\mathbf{x} = \mathbf{0}$ if and only if $QA\mathbf{x} = \mathbf{0}$. Indeed, the $\Rightarrow$ direction follows by multiplying both sides of $A\mathbf{x} = \mathbf{0}$ by Q; the $\Leftarrow$ direction follows by multiplying both sides of $QA\mathbf{x} = \mathbf{0}$ by Q^{-1}. Since $QA = B$, we conclude that $A\mathbf{x} = \mathbf{0}$ if and only if $B\mathbf{x} = \mathbf{0}$. This means the two equations have the same solution set.

(g) True. We know now that for any $n \times n$ matrices A, B we have the if and only if statement:

$$A \text{ and } B \text{ are invertible} \Leftrightarrow AB \text{ is invertible.}$$

Thus if either one of A or B is not invertible, then the product AB is not invertible.

(13) Let $D = \begin{bmatrix} d_1 & 0 & \cdots & \\ 0 & d_2 & 0 & \cdots \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$ be an $n \times n$ diagonal matrix.

(a) Let A be $n \times r$. Use either the row/column method to describe the product DA as on operation on the rows/columns of A. (You choose which is appropriate, rows or columns.)

(b) Let B be $m \times n$. Describe the product BD as on operation of the rows/columns of B. (You choose which is appropriate, rows or columns.)

Solution 13:

(a) To understand what multiplying on the left by D does, we use the row method. This tells us that DA returns the matrix whose i-th row is d_i times the i-th row of A.

(b) To understand what multiplying on the right by D does, we use the column method. This tells us that BD returns the matrix whose j-th column is d_j times the j-th column of B.

(14) Let $A = [a_{ij}]$ be a triangular $n \times n$ matrix.

     Prove: A is invertible if and only if $a_{ii} \neq 0$ for all $1 \leq i \leq n$.

     Hint: use a convenient equivalent statement of invertibility from the invertibility theorem.

Solution 14:

Note: this statement is extremely easy to prove once we can invoke the determinant. Without this tool, it is intuitively true but not so easy to prove as it turns out!

Case 1: assume A is upper triangular. We show the two implications separately.

     ($\Longleftarrow$): this is the easy direction. If all $a_{ii} \neq 0$, then since A is upper triangular, it is easy to see that A can be row reduced to the identity matrix. First, for all i we scale the i-th row by $\frac{1}{a_i}$; this transforms A into a row echelon matrix with a leading 1 in every column. Reducing the resulting matrix to reduced row echelon form then yields the identity matrix.

     ($\Longrightarrow$): we prove this implication by proving its contrapositive.

     Assume one of the diagonal entries of A is equal to 0. Let a_{jj} be the *first* zero diagonal entry, so that $a_{ii} \neq 0$ for all $i < j$. Consider the matrix equation $A\mathbf{x} = \mathbf{0}$, and its corresponding homogenous linear system. I claim there is a solution $(x_1, x_2, \ldots, x_n)$ where $x_j = 1$ and $x_k = 0$ for all $k > j$. This would yield a *non-trivial* solution to the homogenous system; the invertibility theorem then would imply A is not invertible.

     So why can we build such a solution? Since A is upper triangular, looking at the corresponding system, the equations j through n only involve the variables x_k for all $k \geq j$. Since $a_{jj} = 0$, clearly the choice $x_j = 1$, $x_k = 0$ for $k > j$ satisfies equations j through n. Next, by assumption the coefficients $a_{ii} \neq 0$ for all $i < j$. This allows me now to back-substitute and solve for each x_i, $i < j$.

     This concludes the proof that if A is invertible, then $a_{ii} \neq 0$ for all i.

Case 2: assume A is lower triangular. The result in this case follows from Case 1 as follows:

A is invertible $\Longleftrightarrow A^T$ is invertible   (since A inv. if and only if A^T inv.)

         $\Longleftrightarrow A^T$ has nonzero diagonal entries   (since A^T is upper triangular, using Case 1)

         $\Longleftrightarrow A$ has nonzero diagonal entries   (since A, A^T have the same diagonal entries).

(15) Prove that if A and B are symmetric, then $cA + dB$ is symmetric for any $c, d \in \mathbb{R}$.

Solution 15:

     We use the definition that A is symmetric if and only if $A^T = A$.

     Assume A and B are symmetric. Then $A^T = A$ and $B^T = B$. But then

$(cA + dB)^T = (cA)^T + (dB)^T$ (additive prop. of transp.)

       $= cA^T + dB^T$ (scaling prop. of transp. $= cA + dB$ (since A, B symmetric)

This shows $(cA + dB)^T = cA + dB$, which means $A + B$ is symmetric.

(16) Suppose A is symmetric and invertible. Prove A^{-1} is symmetric.

Solution 16:

     We use the alternative definition that A is symmetric if and only if $A^T = A$.

     Assume A is symmetric. Then $A^T = A$. But then

$$(A^{-1})^T = (A^T)^{-1} \qquad\qquad \text{(prop. of transp.)}$$

$$= A^{-1} \qquad\qquad\qquad \text{(since } A \text{ symmetric)}$$

This shows $(A^{-1})^T = A^{-1}$, which means A^{-1} is symmetric.

(17) Let A and B be symmetric. Prove that AB is symmetric if and only if $AB = BA$.
Solution 17:

I will prove this if and only if statement by showing both implications separately. Note that we assume throughout that A and B are symmetric: thus that $A^T = A$ and $B^T = B$.

($\Rightarrow$): suppose AB is symmetric. Then $(AB)^T = (AB)$. On the other hand, we have
$$(AB)^T = B^T A^T = BA,$$
where the last equality holds since A and B are symmetric.
  But then we have $AB = (AB)^T = BA$, showing $AB = BA$.

($\Leftarrow$): suppose $AB = BA$. Then
$$\begin{aligned}
(AB)^T &= B^T A^T && \text{(prop. of transp.)} \\
&= BA && (A,B \text{ are symmetric}) \\
&= AB && \text{(by assumption)}.
\end{aligned}$$
This shows that $(AB)^T = AB$, which proves AB is symmetric.

Having shown that
$$AB \text{ is symmetric} \Rightarrow AB = BA,$$
AND
$$AB = BA \Rightarrow AB \text{ is symmetric},$$
we conclude that
$$AB \text{ is symmetric} \Leftrightarrow AB = BA.$$

(18) For each matrix below (i) decide whether it is lower triangular, upper triangular, diagonal or neither, and (ii) decide whether it is invertible.

(a) $\begin{bmatrix} 4 & 0 \\ 1 & 7 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & -3 \\ 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 4 & 0 & 0 \\ 0 & \frac{3}{5} & 0 \\ 0 & 0 & -2 \end{bmatrix}$

(d) $\begin{bmatrix} 3 & 0 & 0 \\ 3 & 1 & 0 \\ 7 & 0 & 0 \end{bmatrix}$

Solution 18:

a. Lower Triangular and Invertible
b. Upper Triangular and Not Invertible
c. Diagonal (Upper and Lower Triangular) and Invertible
d. Lower Triangular and Not Invertible

(19) Find the product by inspection:

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

Solution 19:

Multiply the first row of the second matrix by 2, the second row by -1, and the third row by 4. Now we have:

$$\begin{bmatrix} 8 & -2 & 6 \\ -1 & -1 & 0 \\ -20 & 4 & -8 \end{bmatrix} \begin{bmatrix} -3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Now multiply the first column of the new matrix by -3, the second column by 5, and the third column by 2. Which gives the matrix:

$$\begin{bmatrix} -24 & -10 & 12 \\ 3 & -10 & 0 \\ 60 & 20 & -16 \end{bmatrix}$$

(20) Find A^2, A^{-2}, and A^{-k} by inspection:

$$A = \begin{bmatrix} -6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Solution 20:

$$A^2 = \begin{bmatrix} 36 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 25 \end{bmatrix}, A^{-2} = \begin{bmatrix} 1/36 & 0 & 0 \\ 0 & 1/9 & 0 \\ 0 & 0 & 1/25 \end{bmatrix}, A^{-k} = \begin{bmatrix} 1/(-6)^k & 0 & 0 \\ 0 & 1/(3)^k & 0 \\ 0 & 0 & 1/(5)^k \end{bmatrix}$$

(21) Find the diagonal entries of AB by inspection.

$$A = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 0 & 0 \\ -3 & 0 & 7 \end{bmatrix}, B = \begin{bmatrix} 6 & 0 & 0 \\ 1 & 5 & 0 \\ 3 & 2 & 6 \end{bmatrix}$$

Solution 21:

The diagonal entries of a product of two lower triangular matrices are obtained simply by multiplying the corresponding diagonal entries. Thus the diagonal entries of the product (in order) are $4 \cdot 6 = 24, 0 \cdot 5 = 0, 7 \cdot 6 = 42$.

(22) Find all values of x for which A is invertible.

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

Solution 22:

Since matrix A is lower triangular, we can apply Theorem 1.7.1, which states that a triangular matrix is invertible if and only if its diagonal entries are all non zero. Thus $x \neq 1/2, 1/3, -1/4$.

(23) Show that if A is a symmetric $n \times n$ matrix and B is any $n \times m$ matrix, the the following products are symmetric: $B^T B$, BB^T, $B^T AB$.

Solution 23:

We will use the general facts that: $(A^T)^T = A$ and $(AB)^T = B^T A^T$. As usual, to show a matrix C is symmetric, we must show $C^T = C$.

$$
\begin{aligned}
(B^T B)^T &= B^T (B^T)^T \\
&= B^T B
\end{aligned}
$$

$$
\begin{aligned}
(BB^T)^T &= (B^T)^T B^T \\
&= BB^T
\end{aligned}
$$

$$
\begin{aligned}
(B^T AB)^T &= B^T A^T B \\
&= B^T AB \text{ (since } A^T = A\text{)}
\end{aligned}
$$

(24) Find all 3×3 diagonal matrices A such that

$$A^2 - 3A - 4I = 0$$

Solution 24:

$\begin{bmatrix} a^2 & 0 & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & c^2 \end{bmatrix} - \begin{bmatrix} 3a & 0 & 0 \\ 0 & 3b & 0 \\ 0 & 0 & 3c \end{bmatrix} - \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ This produces three equations, each with one unknown.

$$
\begin{aligned}
a^2 - 3a - 4 &= 0 \\
b^2 - 3b - 4 &= 0 \\
c^2 - 3c - 4 &= 0
\end{aligned}
$$

Solving for one will solve the other two.
$a^2 - 3a - 4 = 0 \to (a-4)(a+1) = 0 \to a = 4, -1$
Thus $b = 4, -1$ and $c = 4, -1$. It is important to note that in matrix A, a need not be equal to b. So there are 8 solutions to A

$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} : \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -1 \end{bmatrix} : \begin{bmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix} : \begin{bmatrix} -1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

$\begin{bmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} : \begin{bmatrix} -1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -1 \end{bmatrix} : \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix} : \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

(25) Prove: If $A^T A = A$, then A is symmetric and $A = A^2$.
Solution 25:

First we prove $A^T A$ is symmetric by showing $(A^T A)^T = A^T A$:

$$(A^T A)^T = A^T (A^T)^T \qquad\qquad ((AB)^T = B^T A^T)$$
$$= A^T A \qquad\qquad (\text{since } (A^T)^T = A)$$

Thus $A^T A$ is symmetric. Since we assume $A^T A = A$, it follows that A is symmetric; thus $A^T = A$. This means

$$A = A^T A \qquad\qquad (\text{since } A^T A = A)$$
$$= AA \qquad\qquad (\text{since } A^T = A)$$
$$= A^2$$

2.1-2.3: THE DETERMINANT

(1) Let A be $n \times n$. Let c be any constant. State and prove a formula relating $\det(cA)$ with $\det(A)$.
(Look at the $n = 2$ case to see what this should be.)
Solution 1:

WARNING: $\det(cA) \neq c \det(A)$!!!
We claim $\det(cA) = c^n \det(A)$, where A is $n \times n$.

Proof: first observe that $cA = (cI_n)A = \begin{bmatrix} c & 0 & \dots & 0 \\ 0 & c & 0 & \dots \\ \vdots & & & \\ 0 & 0 & \dots & c \end{bmatrix} A$, using the result of Exercise 13

of the previous section. Then

$$\det(cA) = \det((cI_n)A)$$
$$= \det(cI_n) \det(A)$$
$$= c^n \det(A),$$

where the last equality follows since cI_n is a diagonal matrix with c's along the diagonal.

(2) Let A, B be $n \times n$, and suppose B is invertible. Prove the following:
 (a) $\det(B^{-1}) = \dfrac{1}{\det(B)}$.
 (b) $\det(B^{-1}AB) = \det(A)$.

Solution 2:

(a) Since $BB^{-1} = I$, we have

$$1 = \det(I)$$
$$= \det(BB^{-1})$$
$$= \det(B) \det(B^{-1})$$

Dividing both sides by $\det(B)$ (remember $\det(B)$ is just a number!) we get $\det(B^{-1}) = \frac{1}{\det(B)}$.

(b) Start with the LHS and compute:

$$\begin{aligned}
\det(B^{-1}AB) &= \det(B^{-1})\det(A)\det(B) &&\text{(det. is multiplicative)} \\
&= \det(A)\det(B^{-1})\det(B) &&\text{(real mult. is commutative)} \\
&= \det(A)\frac{1}{\det(B)}\det(B) &&\text{(prev. problem)} \\
&= \det(A)
\end{aligned}$$

(3) Let A be $n \times n$, and suppose two rows of A are identical. Show that $\det(A) = 0$. Now prove the same result for a matrix with two identical columns.
Solution 3:

Let the rows of A be $\mathbf{r}_k$, and suppose $\mathbf{r}_i = \mathbf{r}_j$.
Let E be the elementary matrix which replaces $\mathbf{r}_i$ with $\mathbf{r}_i - \mathbf{r}_j$. Then $EA = B$, where the i-th row of B is 0 (since $\mathbf{r}_i = \mathbf{r}_j$ in A). This clearly implies $\det(B) = 0$. Furthermore, we know $\det(E) = 1$. Though this is already a proven result, it is easy to see why it is true. Such an elementary matrix is triangular (maybe upper or lower depending), with 1's along the diagonal; thus its determinant, being the product of the diagonal elements, is 1.
    Now we compute

$$\begin{aligned}
0 &= \det(B) \\
&= \det(EA) \\
&= \det(E)\det(A) \\
&= \det(A),
\end{aligned}$$

showing $\det(A) = 0$.
    Now suppose A has two identical columns. Then A^T has two identical rows, and we have
$$\det(A) = \det(A^T) = 0,$$
where the last equality follows from the previous result.

(4) Suppose A is 3×3 and satisfies:

$$\underset{r_1-r_2}{E}\ \underset{2r_2}{E}\ \underset{r_1\leftrightarrow r_2}{E}\ A = \begin{bmatrix} 1 & 0 & 0 \\ 85 & 2 & 0 \\ -18 & 201 & 3 \end{bmatrix}.$$

Find $\det(A)$.
Solution 4:

We have

$$\det\left(\underset{r_1-r_2}{E}\right)\det\left(\underset{2r_2}{E}\right)\det\left(\underset{r_1\leftrightarrow r_2}{E}\right)\det(A) = \det\left(\begin{bmatrix} 1 & 0 & 0 \\ 85 & 2 & 0 \\ -18 & 201 & 3 \end{bmatrix}\right)$$

$$\Rightarrow 1(2)(-1)\det(A) = 1(2)(3)$$
$$\Rightarrow -2\det(A) = 6$$
$$\Rightarrow \det(A) = -3.$$

(5) State and prove a formula for $\det(A)$ where

$$A = \begin{bmatrix} a & b & b\dots & \\ b & a & b & \dots \\ \vdots & & \vdots & \\ b & b & \dots & a \end{bmatrix},$$

the matrix with a's along the diagonal and b's everywhere else.

Look at the $n = 2$ and $n = 3$ cases first. Your proof of the formula in the general case will involve row operations.

Solution 5:

We perform the following row operations: replace r_1 with $r_1 - r_2$; replace r_2 with $r_2 - r_3$; replace r_3 with $r_3 - r_4$, etc.These are row operations of the third type $(r_i + cr_j)$, and thus do NOT change the determinant. (Recall of course that the other types of row operatins DO change the determinant.)

These operations result in the matrix

$$B = \begin{bmatrix} (a-b) & -(a-b) & 0 & 0 & \dots \\ 0 & (a-b) & -(a-b) & 0 & \dots \\ \vdots & & & & \\ 0 & 0 & & \dots & (a-b) & -(a-b) \\ b & b & b & \dots & b & a \end{bmatrix}$$

Now replace column $\mathbf{c}_2$ with $\mathbf{c}_2 + \mathbf{c}_1$, column $\mathbf{c}_3$ with $\mathbf{c}_3 + \mathbf{c}_2, \dots$, column $\mathbf{c}_n$ with $\mathbf{c}_n + \mathbf{c}_{n-1}$. These operations, which also do not change the determinant, result in the lower triangular matrix

$$C = \begin{bmatrix} (a-b) & 0 & 0 & 0 & \dots \\ 0 & (a-b) & 0 & 0 & \dots \\ 0 & 0 & (a-b) & 0 & \dots\vdots \\ 0 & 0 & & \dots & (a-b) & 0 \\ b & 2b & 3b & \dots & (n-1)b & a+(n-1)b \end{bmatrix}$$

We conclude that $\det(A) = \det(B) = \det(C) = (a-b)^{n-1}(a+(n-1)b)$.

(6) Given $r_1, r_2, \dots, r_n \in \mathbb{R}$ the *Vandermonde* matrix is defined as

$$A_{r_1, r_2, \dots, r_n} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ r_1 & r_2 & \cdots & r_n \\ r_1^2 & r_2^2 & \cdots & r_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ r_1^{n-1} & r_2^{n-1} & \cdots & r_n^{n-1} \end{bmatrix} = [r_j^{i-1}]_{1 \le i,j \le n}.$$

Prove: $\det(A_{r_1, r_2, \dots, r_n}) = \displaystyle\prod_{1 \le i < j \le n} (r_j - r_i)$.

Hint. Do a proof by induction. For the induction step, first clear out the entries under the 1 of the first column, working from the bottom up. You will also use linearity in columns.

Solution 6:

This is a well known result. Look it up or else follow my instructions. The sequence of row operations (working from bottom up) is $\mathbf{r}_n - r_1\mathbf{r}_{n-1}, \mathbf{r}_{n-1} - r_1\mathbf{r}_{n-2}, \dots, \mathbf{r}_2 - r_1\mathbf{r}_1$.

(7) Let
$$A = \begin{bmatrix} 2 & 3 & -1 & 1 \\ -3 & 2 & 0 & 3 \\ 3 & -2 & 1 & 0 \\ 3 & -2 & 1 & 4 \end{bmatrix}$$

(a) Find M_{32} and C_{32}.

(b) Find M_{44} and C_{44}.

Solution 7:

(a) Delete row 3 and column 2 from A to and compute the determinant:
$$\begin{vmatrix} 2 & -1 & 1 \\ -3 & 0 & 3 \\ 3 & 1 & 4 \end{vmatrix} = 2(-3) + 1(-21) + -3 = -30.$$

Thus $M_{32} = -30$ and $C_{32} = (-1)^{3+2}M_{32} = 30$

(b) Delete the 4th row and column and compute the determinant:
$$\begin{vmatrix} 2 & 3 & -1 \\ -3 & 2 & 0 \\ 3 & -2 & 1 \end{vmatrix} = 2(2) - 3(-3) + (-1)(0) = 13.$$

Thus $M_{44} = 13$ and $C_{44} = (-1)^{4+4}M_{44} = 13$

(8) Let
$$A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & 0 & -5 \\ 1 & 7 & 2 \end{bmatrix}$$

(a) Compute $\det(A)$ by expanding along the second row.

(b) Compute $\det(A)$ by expanding along the third column.

Solution 8:

(9) Compute
$$\det \begin{bmatrix} 3 & 3 & 0 & 5 \\ 2 & 2 & 0 & -2 \\ 4 & 1 & -3 & 0 \\ 2 & 10 & 3 & 2 \end{bmatrix}$$

Solution 9:

Since column three has two zeros, lets expand by it:
$$\det(A) = -3 \begin{vmatrix} 3 & 3 & 5 \\ 2 & 2 & -2 \\ 2 & 10 & 2 \end{vmatrix} - 3 \begin{vmatrix} 3 & 3 & 5 \\ 2 & 2 & -2 \\ 4 & 1 & 0 \end{vmatrix}$$

Now we compute these 3×3 deterimants:
$$\begin{aligned} \det(A) &= -3(3(24) - 3(8) + 5(16)) - 3(4(-16) - 1(-16)) \\ &= -3(8 \cdot 16) - 3(-3 \cdot 16) \\ &= -15 \cdot 16 = -240. \end{aligned}$$

(10) Compute

$$\det \begin{bmatrix} 4 & 0 & 0 & 1 & 0 \\ 3 & 3 & 3 & -1 & 0 \\ 1 & 2 & 4 & 2 & 3 \\ 9 & 4 & 6 & 2 & 3 \\ 2 & 2 & 4 & 2 & 3 \end{bmatrix}$$

Solution 10:

Since row one has three zeros, lets expand by row one.

$$\det(A) = 4 \begin{vmatrix} 3 & 3 & -1 & 0 \\ 2 & 4 & 2 & 3 \\ 4 & 6 & 2 & 3 \\ 2 & 4 & 2 & 3 \end{vmatrix} - 1 \begin{vmatrix} 3 & 3 & 3 & 0 \\ 1 & 2 & 4 & 3 \\ 9 & 4 & 6 & 3 \\ 2 & 2 & 4 & 3 \end{vmatrix}$$

Expanding by the 4th columns gives:

$$= 4 \left(3 \begin{vmatrix} 3 & 3 & -1 \\ 4 & 6 & 2 \\ 2 & 4 & 2 \end{vmatrix} - 3 \begin{vmatrix} 3 & 3 & -1 \\ 2 & 4 & 2 \\ 2 & 4 & 2 \end{vmatrix} + 3 \begin{vmatrix} 3 & 3 & -1 \\ 2 & 4 & 2 \\ 4 & 6 & 2 \end{vmatrix} \right)$$

$$-1 \left(3 \begin{vmatrix} 3 & 3 & 3 \\ 9 & 4 & 6 \\ 2 & 2 & 4 \end{vmatrix} - 3 \begin{vmatrix} 3 & 3 & 3 \\ 1 & 2 & 4 \\ 2 & 2 & 4 \end{vmatrix} + 3 \begin{vmatrix} 3 & 3 & 3 \\ 1 & 2 & 4 \\ 9 & 4 & 6 \end{vmatrix} \right)$$

$$= 12((32-36)-(28-28)+(36-32))-3((138-168)-(54-48)+$$

$$= 12(-4+4)-3(-30-6+36)=0$$

(11) We know now how elementary row operations affect the determinant. State and prove analogous statements about how elementary *column operations* affect the determinant of a matrix.
Solution 11:

Multiplying a matrix A on the *right* by an elementary matrix E has the effect of performing the corresponding elementary *column operation*. In more detail:
$AE_{c\mathbf{r}_i}$ scales the i-th column of A by c;
$AE_{\mathbf{r}_i \leftrightarrow \mathbf{r}_j}$ swaps columns i and j;
$AE_{\mathbf{r}_i + c\mathbf{r}_j}$ replaces column i with column i plus c times column j.
    To see what happens to the determinant, we use the multiplicative property:
$\det(AE_{c\mathbf{r}_i}) = \det A \det E_{c\mathbf{r}_i} = c \det A$;
$\det(AE_{\mathbf{r}_i \leftrightarrow \mathbf{r}_j}) = \det A \det E_{\mathbf{r}_i \leftrightarrow \mathbf{r}_j} = -\det A$;
$\det(AE_{\mathbf{r}_i + c\mathbf{r}_j}) = \det A \det E_{\mathbf{r}_i + c\mathbf{r}_j} = \det A$.

(12) Let

$$A = \begin{bmatrix} -4 & 1 & 1 & 1 & 1 \\ 1 & -4 & 1 & 1 & 1 \\ 1 & 1 & -4 & 1 & 1 \\ 1 & 1 & 1 & -4 & 1 \\ 1 & 1 & 1 & 1 & -4 \end{bmatrix}$$

Show that $\det(A) = 0$ without directly computing the determinant.
Solution 12:

First solution: notice that each row sums to 0. Using the column method of matrix multiplication, this implies that

$$A \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \mathbf{0}_{5 \times 1}$$

Thus $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution: namey, $\mathbf{x} = (1,1,1,1,1)$. It follows from the invertibility theorem that A is not invertible. Thus $\det A = 0$.

Second solution: each column also sums to 0. This means if we apply successive row operations, replacing $\mathbf{r}_1$ with $\mathbf{r}_1 + \mathbf{r}_2$, then replacing the resulting $\mathbf{r}_1$ with $\mathbf{r}_1 + \mathbf{r}_2$, and so on, that we will eventually end up with a row of zeros:

$$
\begin{bmatrix}
-4 & 1 & 1 & 1 & 1 \\
1 & -4 & 1 & 1 & 1 \\
1 & 1 & -4 & 1 & 1 \\
1 & 1 & 1 & -4 & 1 \\
1 & 1 & 1 & 1 & -4
\end{bmatrix}
\xrightarrow{r_1 + r_2}
\begin{bmatrix}
-3 & -3 & 2 & 2 & 2 \\
1 & -4 & 1 & 1 & 1 \\
1 & 1 & -4 & 1 & 1 \\
1 & 1 & 1 & -4 & 1 \\
1 & 1 & 1 & 1 & -4
\end{bmatrix}
$$

$$
\xrightarrow{r_1 + r_3}
\begin{bmatrix}
-2 & -2 & -2 & 3 & 3 \\
1 & -4 & 1 & 1 & 1 \\
1 & 1 & -4 & 1 & 1 \\
1 & 1 & 1 & -4 & 1 \\
1 & 1 & 1 & 1 & -4
\end{bmatrix}
$$

$$
\xrightarrow{r_1 + r_4}
\begin{bmatrix}
-1 & -1 & -1 & -1 & 4 \\
1 & -4 & 1 & 1 & 1 \\
1 & 1 & -4 & 1 & 1 \\
1 & 1 & 1 & -4 & 1 \\
1 & 1 & 1 & 1 & -4
\end{bmatrix}
$$

$$
\xrightarrow{r_1 + r_5}
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 \\
1 & -4 & 1 & 1 & 1 \\
1 & 1 & -4 & 1 & 1 \\
1 & 1 & 1 & -4 & 1 \\
1 & 1 & 1 & 1 & -4
\end{bmatrix}
$$

Let B be the matrix we row reduced A to. Since all of our row operations were of the row addition type, we have $\det A = \det B$. Since B has a row of zeros, $\det B = 0$. It follows that $\det A = \det B = 0$.

(13) Use determinants to decide whether the given matrix is invertible.

$$A = \begin{bmatrix} 2 & 0 & 3 \\ 0 & 3 & 2 \\ -2 & 0 & -4 \end{bmatrix}$$

Solution 13:

Expanding by the first column gives:

$$\det(A) = 2 \begin{bmatrix} 3 & 2 \\ 0 & -4 \end{bmatrix} - 2 \begin{bmatrix} 0 & 3 \\ 3 & 2 \end{bmatrix} = -24 + 18 = -6$$

Since the determinant is non-zero, the matrix is invertible.

(14) Use determinants to decide whether the given matrix is invertible.

$$A = \begin{bmatrix} \sqrt{2} & -\sqrt{7} & 0 \\ 3\sqrt{2} & -3\sqrt{7} & 0 \\ 5 & -9 & 0 \end{bmatrix}$$

Solution 14:

Expanding by the third column gives:

$$\det(A) = 0$$

Thus A is not invertible.

(15) Find the values of k for which the matrix A is invertible.

$$\begin{bmatrix} 1 & 2 & 0 \\ k & 1 & k \\ 0 & 2 & 1 \end{bmatrix}$$

Solution 15:

Expanding by the first row gives:

$$\det(A) = \begin{bmatrix} 1 & k \\ 2 & 1 \end{bmatrix} - k \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix} = 1 - 2k - 2k = 1 - 4k.$$

Now we set $1 - 4k = 0$. Solving for k gives: $k = \frac{1}{2}$. Thus matrix A is invertible for all $k \neq \frac{1}{2}$.

(16) Use the adjoint formula to compute the inverse of

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

Solution 16:

First compute $\det A = 1(-1) - 2(-2) - 3(-1) = 6$. Next compute

$$\operatorname{adj} A = ([C_{ij}])^T = \left(\begin{bmatrix} -1 & 2 & -1 \\ 2 & 2 & 2 \\ 5 & -4 & -1 \end{bmatrix} \right)^T = \begin{bmatrix} -1 & 2 & 5 \\ 2 & 2 & -4 \\ -1 & 2 & -1 \end{bmatrix}.$$

We conclude that

$$A^{-1} = \frac{1}{6} \begin{bmatrix} -1 & 2 & 5 \\ 2 & 2 & -4 \\ -1 & 2 & -1 \end{bmatrix}.$$

Verify this yourself by checking $AA^{-1} = I_3$!!

(17) In each part, find the determinant given that A is a 4x4 matrix for which $\det(A) = -2$.
 (a) $\det(-A)$
 (b) $\det(A^{-1})$
 (c) $\det(2A^T)$
 (d) $\det(A^3)$

Solution 17:

(a)
$$\det(-A) = \det(-1 \cdot A) = (-1)^4 \det(A) = \det(A) = -2$$

(b)
$$\det(-A) = \frac{1}{\det(A)} = \frac{1}{-2} = -\frac{1}{2}$$

(c)
$$\det(2A^T) = 2^4 det(A) = 2^4(-2) = -32$$

(d)
$$\det(A^3) = \det(AAA) = \det(A)\det(A)\det(A) = -2 \cdot -2 \cdot -2 = -8$$

(18) Prove that a square matrix A is invertible if and only if $A^T A$ is invertible.
Solution 18:

($\Rightarrow$) Assume that A is invertible:
Since A is square, we know that $\det(A^T) = \det(A)$. Since A is invertible, $\det(A) \neq 0$. Now:
$$\det(A^T A) = \det(A^T)\det(A) = \det(A)\det(A) \neq 0$$

Thus $A^T A$ is invertible.
($\Leftarrow$) Assume that $A^T A$ is invertible:
Since $A^T A$ is invertible, we know that $\det(A^T A) \neq 0$. Thus
$$0 \neq \det(A^T A) = \det(A^T)\det(A) = \det(A)\det(A)$$

which makes $\det(A) \neq 0$. So matrix A is invertible.
  Alternatively, we have

$$A \text{ inv.} \Leftrightarrow A \text{ and } A^T \text{ inv.} \qquad\qquad (A \text{ inv.} \Leftrightarrow A^T \text{ inv.})$$
$$\Leftrightarrow A^T A \text{ inv.} \qquad\qquad (A, B \text{ inv.} \Leftrightarrow A \cdot B \text{ inv.})$$

That paranthetical justification of the first $\Leftrightarrow$ was never explicity stated thus far; we only proved the direction A invertible $\Rightarrow A^T$ invertible. However, we can prove the reverse direction using the first implication:

$$A^T \text{ invertible} \Rightarrow (A^T)^T \text{ invertible} \qquad\qquad \text{(by first implication)}$$
$$\Rightarrow A \text{ invertible} \qquad\qquad ((A^T)^T = A).$$

(19) True or false. If true, provide a proof; if false, give an explicit counterexample.
  (a) If A is a 4x4 matrix and B is obtained from A by interchanging the first two rows and then interchanging the last two rows, then $\det(A) = \det(B)$.
  (b) If A is a 3x3 matrix and B is obtained from A by multiplying the first column by 4 and multiplying the third column by $\frac{3}{4}$, then $\det(B) = 3 \det(A)$.
  (c) If A is a 3x3 matrix and B is obtained from A by adding 5 times the first row to each of the second and third rows, then $\det(B) = 25 \det(A)$.
  (d) If the sum of the second and fourth row vectors of a 6x6 matrix A is equal to the last row vector, then $\det(A) = 0$
  (e) If A is a 3x3 matrix, then $\det(2A) = 2 \det(A)$.
  (f) If A and B are square matrices of the same size and A is invertible, then $\det(A^{-1}BA) = \det(B)$.

(g) If A is a square matrix and the linear system $A\mathbf{x} = \mathbf{0}$ has multiple solutions for $\mathbf{x}$, then $\det(A) = 0$.

(h) If E is an elementary matrix, then $E\mathbf{x} = \mathbf{0}$ has only the trivial solution.

Solution 19:

(a) True. The row operations and determinant theorem states that interchanging two rows of a matrix flips the sign on the determinant. Since we did the interchanging twice, the sign flips twice. Thus $\det(A) = \det(B)$.

(b) True. We saw in Exercise 11 that performing elementary column operations on a matrix has the same effect on the determinant as the corresponding row operation. Thus in this case we have $\det(B) = \frac{3}{4} \cdot 4 \det(A) = 3 \det(A)$.

(c) False. By the row operations and determinant theorem we know that multiplying one row and adding it to another row does not change the value of the determinant. Thus $\det(A) = \det(B)$.

(d) True. By GE: Add r_2 to r_4 to replace r_2. Then subtract r_6 from r_2 to replace r_6. Now r_6 is all zeros, thus $\det(A) = 0$.

(e) False. In fact we have

$$\det(2A) = 2^3 \det(A) = 8 \det(A) \neq 2 \det(A),$$

as long as $\det A \neq 0$. For an explicit counterexample, take $A = I_3$. Then $\det(2I_3) = 8$ and $2 \det I_n = 2$.

(f) True. We have

$$
\begin{aligned}
\det(A^{-1}BA) &= \det(A^{-1}) \det(B) \det(A) \\
&= \frac{1}{\det(A)} \det(B) \det(A) \\
&= \det(B)
\end{aligned}
$$

(g) True. By the invertibility theorem know that the following statements are equivalent:
$$A\mathbf{x} = \mathbf{0} \text{ has only the trivial solution}$$
$$\det(A) \neq 0$$
By our assumptions, the first statement is false, thus the second statement is also false.

(h) True. Elementary matrices are invertible.

3.1: VECTOR SPACES

(1) Let $V = \{A \in M_{nn} : A \text{ is invertible}\}$. Show that the usual matrix addition and scalar multiplication do NOT make V into a vector space.
Solution 1:

We know already that matrix addition and scalar multiplication satisfy *most* of the vector space axioms: e.g., commutativity, distributivity, etc.

However the operations fail to be well-defined defined on the given set! That is the sum of two matrices in V does not necessarily lie in V.

As a concrete counterexample, take I_n and $-I_n$. Both are invertible matrices, and thus lie in V. However, $I_n + (-I_n) = \mathbf{0}_n$, the zero matrix, which does not lie in V.

Note also that scalar multiplication is also not well-defined on V: if we scale any invertible matrix A by the scalar $c = 0$, then we get the zero matrix, which is not in V.

(2) Let $V = \mathbb{R}^n$. Define vector addition on V to be the usual n-vector addition, but define scalar multiplication as

$$r(v_1, v_2, \ldots, v_n) = (r^2 v_1, r^2 v_2, \ldots, r^2 v_n).$$

Show that V is NOT a vector space with these choices of vector addition and scalar multiplication.

Solution 2:

This choice of vector addition and scalar multiplication fails the following distributivity axiom:

$$(r+s)\mathbf{v} = r\mathbf{v} + s\mathbf{v}.$$

Indeed take $r = s = 1$, and $\mathbf{v} = (1, 1, \ldots, 1)$. Then on the one hand we have

$$(r+s)\mathbf{v} = (1+1)\mathbf{v} = 2(1, 1, \ldots, 1) = (4, 4, \ldots, 4),$$

using the given definition of scalar multiplication.
On the other hand we have

$$r\mathbf{v} + s\mathbf{v} = 1(1, 1, \ldots, 1) + 1(1, 1, \ldots, 1) = (1, 1, \ldots, 1) + (1, 1, \ldots, 1) = (2, 2, \ldots, 2),$$

again using the given definitions of scalar mult. and vector addition.
Thus we see that

$$(r+s)\mathbf{v} \neq r\mathbf{v} + s\mathbf{v}$$

for this choice of r, s and $\mathbf{v}$, and conclude the given set and operations do not form a vector space.

(3) Let $V = \{(x, y) \in \mathbb{R}^2 : x > 0, \ y < 0\}$: i.e., V is the set of pairs whose first component is positive, and whose second component is negative. The following operations satisfy the axioms of a vector space.
Vector addition: $(x_1, \ y_1) + (x_2, \ y_2) = (x_1 x_2, \ -y_1 y_2)$.
(In other words, vector addition takes the product of the first components, and takes the negative of the product of the second components.)

Scalar multiplication: $r(x, \ y) = (x^r, \ -|y|^r)$.
Below you are asked to verify some (but not all!) of the vector space axioms.
  (a) Explicitly identify the element of V that acts as the additive identity $\mathbf{0}$ with respect to the above vector operations.
     Show directly that your choice of $\mathbf{0}$ satisfies $\mathbf{0} + \mathbf{v} = \mathbf{v}$ for all $\mathbf{v} = (x, y) \in V$.
     Caution: at the very least, the $\mathbf{0}$ you provide should be an actual element of the given V. For example, note that $(0, 0)$ is not even an element of V !!
  (b) Given $\mathbf{v} = (x, y) \in V$, give a formula in terms of x and y for its additive inverse $-\mathbf{v}$ with respect to the above vector operations.
     Show directly that your formula for $-\mathbf{v}$ satisfies $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$. Again, you must always use the definition of the vector space operations given above.
  (c) Prove that Axiom (h) holds: i.e., that $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$.

Solution 3:
(a) The additive identity for the given vector space is $\mathbf{0} = (1, -1)$.
Proof:

$$\begin{aligned}
(1, -1) + (x, y) &= (1x, -(-1)y) &&\text{(def. of vect. add.)}\\
&= (x, y)
\end{aligned}$$

(b) Given element $\mathbf{v} = (x, y) \in V$, its additive inverse is $-\mathbf{v} = \left(\frac{1}{x}, \frac{1}{y}\right)$.

Proof:

$$(x,y) + \left(\frac{1}{x}, \frac{1}{y}\right) = \left(x\,\frac{1}{x}, -y\,\frac{1}{y}\right) \qquad \text{(def. of vect. add.)}$$
$$= (1, -1)$$
$$= \mathbf{0} \qquad\qquad\qquad \text{(since } \mathbf{0} = (1, -1))$$

(c) Proof:

$$1(x,y) = (x^1, -|y|^1) \qquad \text{(def. of scalar mult.)}$$
$$= (x, -|y|)$$
$$= (x, y).$$

The last equality in the proof above follows from the definition of absolute value: for $y < 0$, we have by definition $|y| = -y$.

(4) Let $V = \{f(x) = 1 + ax \colon a \in \mathbb{R}\}$ be the set of all linear polynomials $f(x)$ with constant coefficient 1: i.e., $f(0) = 1$. Define:

$$\text{Vector addition: } (1 + ax) + (1 + bx) := (1 + ax)(1 + bx) - abx^2$$
$$\text{Scalar multiplication: } r \cdot (1 + ax) := 1 + rax$$

Prove V along with these operations forms a vector space. Make explicit what the additive identity and inverses are.

Solution 4:

If you expand and simplify the expression $(1 + ax)(1 + bx) - abx^2$ you get $1 + (a + b)x$. This shows that our vector addition is well-defined on the set V; it is clear that our scalar multiplication is also well-defined.

An element $f(x) = 1 + ax$ is completely determined by the coefficient a, so let's denote $f_a(x) = 1 + ax$. Our vector operations have a very simple description using this notation:

$$f_a(x) + f_b(x) = f_{a+b}(x)$$
$$r \cdot f_a(x) = f_{ra}(x).$$

In other words we see that our "fancy" vector operations boil down to the usual addition $a + b$, and multiplication $r \cdot a$. Once this it is clear, it is obvious that the operations satisfy the vector space axioms.

The additive identity is $\mathbf{0} = f_0(x) = 1 + 0x = 1$, the constant function equal to 1. Additive inverses are given by the formula $-f_a(x) = f_{-a}(x)$; i.e., the additive inverse of $f(x) = 1 + ax$ is $g(x) = 1 - ax$.

(5) Let $V = \mathbb{R}^2$. The following operations do **NOT** satisfy all of the vector space axioms.

Vector addition: $(x_1, y_1) + (x_2, y_2) = (x_1 + y_1, x_2 + y_2)$. (The usual vector addition on $\mathbb{R}^2$.)

Scalar multiplciation: $r(x, y) = (rx, 0)$.

(a) Verify that these operations satisfy Axiom (f): i.e., that $(r + s)\mathbf{v} = r\mathbf{v} + s\mathbf{v}$ for all $r, s \in \mathbb{R}$ and all $(x, y) \in \mathbb{R}^2$.

(b) Find an axiom that these operations do **NOT** satisfy, and provide an explicit counterexample showing that the axiom fails.

Solution 5:

(a) We have

$$
\begin{aligned}
(r+s)(x,y) &= ((r+s)x, 0) && \text{(by def. of scal. mult.)} \\
&= (rx + sx, 0) && \text{(real number arith.)} \\
&= (rx, 0) + (sx, 0) && \text{(by def. of vector add.)} \\
&= r(x, 0) + s(x, 0) && \text{(by def. of scal. mult.).}
\end{aligned}
$$

(b) The given operations do not satisfy axiom (h). By definition we have $1(x, y) = (x, 0)$. This will not be equal to (x, y) unless $y = 0$.

(6) Prove the vector space properties theorem. (Some of these parts have been proved in the lecture notes or elsewhere. No matter, recreate those proofs or produce new ones!)
Solution 6:

The theorem states that, given a vector space V, the properties below hold.
(a) $0\mathbf{v} = \mathbf{0}$ for all $\mathbf{v} \in V$.
(b) $k\mathbf{0} = \mathbf{0}$ for all $k \in \mathbb{R}$.
(c) $(-1)\mathbf{v} = -\mathbf{v}$ for all $\mathbf{v} \in V$.
(d) If $k\mathbf{u} = \mathbf{0}$, then $k = 0$ or $\mathbf{u} = \mathbf{0}$.
We have proved (a) and (c) somewhere (lecture notes, video examples, class, etc.).
It remains to prove (b) and (d). I may make use of properties (a) and (c), since these are already proven.
Let's prove (b). From (a) we know that if $k = 0$, then $k\mathbf{0} = 0\mathbf{0} = \mathbf{0}$. Suppose $k \neq 0$. Then we have

$$
\begin{aligned}
k\mathbf{0} &= k(\mathbf{0} + \mathbf{0}) && (\mathbf{0} + \mathbf{0} = \mathbf{0}) \\
&= k\mathbf{0} + k\mathbf{0} && \text{(dist. prop.)}
\end{aligned}
$$

Thus $k\mathbf{0} = k\mathbf{0} + k\mathbf{0}$. Now add the additive identity $-k\mathbf{0}$ to both sides of this equation to "cancel" one of the $k\mathbf{0}$. We are left with $\mathbf{0} = k\mathbf{0}$, as desired.

Finally, we prove (d). Suppose $k\mathbf{u} = \mathbf{0}$ and that $k \neq 0$. Then k has a multiplicative inverse $\frac{1}{k}$. Then

$$
\begin{aligned}
k\mathbf{u} = \mathbf{0} &\Rightarrow \frac{1}{k}(k\mathbf{u}) = \frac{1}{k}\mathbf{0} \\
&\Rightarrow (\frac{1}{k}k)\mathbf{u} = \mathbf{0} && \text{(assoc. and property (b))} \\
&\Rightarrow 1\mathbf{u} = \mathbf{0} \\
&\Rightarrow \mathbf{u} = \mathbf{0} && \text{(since } 1\mathbf{u} = \mathbf{u}).
\end{aligned}
$$

We have shown that if $k\mathbf{u} = \mathbf{0}$ and $k \neq 0$, then $\mathbf{u} = \mathbf{0}$. Logically, this is equivalent to showing if $k\mathbf{u} = \mathbf{0}$, then $k = 0$ or $\mathbf{u} = \mathbf{0}$.

(7) Prove that a vector space V is either trivial (i.e., $V = \{\mathbf{0}\}$) or infinite. Hint: if $\mathbf{v} \neq \mathbf{0}$, show that $r\mathbf{v} \neq s\mathbf{v}$ for two distinct real numbers $r \neq s$.
Solution 7:

Suppose V is nontrivial. Then there is a nonzero $\mathbf{v} \neq \mathbf{0}$ in V. Then for each $r \in \mathbb{R}$ the vector $r\mathbf{v}$ is also an element of V. To show that V infinite we need only show that all of these $r\mathbf{v}$ are distinct; i.e., we must show that if $r \neq s$, then $r\mathbf{v} \neq s\mathbf{v}$.
(Note: we cannot take this for granted, but must actually prove it. For all we know we could have $r\mathbf{v} = \mathbf{v}$ for all $r \in \mathbb{R}$.)

We prove the contrapositive. Suppose $r\mathbf{v} = s\mathbf{v}$. Then $r\mathbf{v} - s\mathbf{v} = \mathbf{0}$. Then $(r - s)\mathbf{v} = \mathbf{0}$. I now claim that since $\mathbf{v} \neq \mathbf{0}$, we must have $r - s = 0$, and hence $r = s$.

Indeed suppose $a\mathbf{v} = \mathbf{0}$ and $a \neq 0$. Since $a \neq 0$, we can multiply both sides of this equation by $\frac{1}{a}$:

$$\frac{1}{a}(a\mathbf{v}) = \frac{1}{a}\mathbf{0}$$
$$(\frac{1}{a} \cdot a)\mathbf{v} = \mathbf{0} \qquad \text{(assoc. of scalar mult; } c\mathbf{0} = \mathbf{0})$$
$$1\mathbf{v} = \mathbf{0}$$
$$\mathbf{v} = \mathbf{0} \qquad \text{(Axiom: } 1\mathbf{v} = \mathbf{v})$$

That concludes our (longer than expected) proof that if $r \neq s$, then $r\mathbf{v} \neq s\mathbf{v}$, which proves that V is infinite.

(8) Let V be a vector space. Show that for all $\mathbf{v} \in V$ there is a unique additive inverse for $\mathbf{v}$. That is, show if $\mathbf{v} + \mathbf{w} = \mathbf{0}$, then $\mathbf{w} = -\mathbf{v}$.
Solution 8:

Fix $\mathbf{v}$ and suppose $\mathbf{w}$ also satisfies $\mathbf{v} + \mathbf{w} = \mathbf{0}$. Then

$$\mathbf{v} + -\mathbf{v} = \mathbf{0} = \mathbf{v} + \mathbf{w}.$$

Then we have

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{v} + \mathbf{w}$$
$$-\mathbf{v} + \mathbf{v} + (-\mathbf{v}) = -\mathbf{v} + \mathbf{v} + \mathbf{w}$$
$$\mathbf{0} + (-\mathbf{v}) = \mathbf{0} + \mathbf{w}$$
$$-\mathbf{v} = \mathbf{w},$$

as desired.

(9) Let V be a vector space. Prove, using only the axioms of a vector space: if $\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}$, then $\mathbf{u} = \mathbf{v}$.
Solution 9:

$$
\begin{aligned}
\mathbf{u} + \mathbf{w} &= \mathbf{v} + \mathbf{w} && \text{(Hypothesis)} \\
(\mathbf{u} + \mathbf{w}) + (-\mathbf{w}) &= (\mathbf{v} + \mathbf{w}) + (-\mathbf{w}) && \text{(Add -}\mathbf{w}\text{ to both sides)} \\
\mathbf{u} + [\mathbf{w} + (-\mathbf{w})] &= \mathbf{v} + [\mathbf{w} + (-\mathbf{w})] && \text{(Axiom 3)} \\
\mathbf{u} + \mathbf{0} &= \mathbf{v} + \mathbf{0} && \text{(Axiom 5)} \\
\mathbf{u} &= \mathbf{v} && \text{(Axiom 4)}
\end{aligned}
$$

3.2: LINEAR TRANSFORMATIONS

(1) For each of the following functions T, decide whether T is linear. If yes, provide a proof; if no, given an explicit example that violates one of the axioms.
   (a) $T\colon M_{nn} \to M_{nn}$, $T(A) = A^2$
   (b) $T\colon M_{nn} \to \mathbb{R}$, $T(A) = \operatorname{tr} A$
   (c) $T\colon M_{nn} \to \mathbb{R}$, $T(A) = \det A$.
   (d) $T\colon F((-\infty, \infty)) \to F((-\infty, \infty))$, $T(f) = 1 + f$
   (e) $T\colon F((-\infty, \infty)) \to F((-\infty, \infty))$, $T(f(x)) = f(x+3)$
   (f) $T\colon \mathbb{R}^3 \to \mathbb{R}^2$, $T(x, y, z) = (xy, yz)$.

Solution 1:
(a) Nonlinear. Take $A = I_2$. Then $T(2A) = 4I \neq 2T(A) = 2I$.
(b) Linear. Easy proof.
(c) Nonlinear. Let $A = I_2$. Then $\det(2A) = 4 \neq 2\det(A) = 2$.
(d) Nonlinear. Notice that $T(\mathbf{0}) = 1 \neq \mathbf{0}$; yet any linear transformation sends the zero vector to the zero vector.
(e) Linear.

$$\begin{aligned} T(cf + dg) &= (cf + g)(x + 3) && \text{(by def)} \\ &= cf(x+3) + dg(x+3) && \text{(function arith.)} \\ &= cT(f) + dT(g) && \text{(by def)} \end{aligned}$$

(f) Nonlinear. $T(2(1,1,1)) = T(2,2,2) = (4,4)$ and $2T(1,1,1) = 2(1,1) = (2,2)$, so T does not respect scalar multiplication. (Nor does it respect vector addition.)

(2) Define $T\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ by $T(f) = g$, where $g(x) = f(x) + f(-x)$.
Show that T is a linear transformation.
Solution 2:

We show T is linear by showing $T(cf + dg) = cT(f) + dT(g)$ for any $c, d \in \mathbb{R}$ and any $f, g \in C^\infty(\mathbb{R})$.

$$\begin{aligned} T(cf + dg) &= (cf + dg)(x) + (cf + dg)(-x) && \text{(by def)} \\ &= c(f(x) + f(-x)) + d(g(x) + g(-x)) && \text{(arith)} \\ &= cT(f) + dT(g) && \text{(by def)} \end{aligned}$$

(3) *Reflections in $\mathbb{R}^2$.* Given a fixed angle α, $0 \leq \alpha < \pi$, let ℓ_α be the line through the origin that makes an angle of α with the positive x-axis.
We define $T_\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ as $T_\alpha(P) = $ the reflection P' of P through ℓ_α
In more detail, given point $P \in \mathbb{R}^2$, let ℓ' be the line passing through P that is perpendicular to ℓ_α, and let Q be the intersection of ℓ and ℓ_α. If P already lies on ℓ_α, then $T_\alpha(P) = P$; otherwise $T_\alpha(P)$ is the *other* point P' lying on ℓ, whose distance to Q is equal to the distance between P and Q.
   (a) Give a geometric argument (i.e., draw a picture) that strongly suggests T_α is linear.
   (b) Now prove T_α is linear by finding a 2×2 matrix A such that $T_\alpha = T_A$.

Solution 3:
(a) As the diagram below illustrates, reflection preserves sums of vectors and scalar multiples of vectors. This is a result of the fact that reflection preserves angles between position vectors.

(b) As suggestive as the diagram is, it falls somewhat short of a proof. We will find the matrix $\underset{2\times 2}{A}$ such that $T_\alpha = T_A$.

It is easiest to see what reflection does by thinking in terms of polar coordinates. Recall: any point $P = (x, y)$ can be expressed in the form $(r\cos\theta, r\sin\theta)$. The values r and θ are called polar coordinates of the vector. In general there are many choices for r and θ for any given vector (x, y). We will stipulate that $\mathbf{r} = \sqrt{x^2 + y^2}$ (the length of the vector), and that θ is the measure of any oriented angle that *starts from* the positive x-axis and *ends at* $\overrightarrow{OP}$. Note that for $(x, y) \neq (0, 0)$ the angle θ is defined only up to multiples of $2\pi n$, where n is an integer.

Suppose $P = (x, y)$ has polar coordinates (r, θ). As is illustrated in the diagram below, the oriented angle *starting from* the position vector $\overrightarrow{OP}$ and *ending at* ℓ_α can be described as $\alpha - \theta$ (up to a multiple of 2π). By symmetry (or congruent triangles), the oriented angle from the line ℓ_α to the position vector $\overrightarrow{OP'}$ of the reflection P' can also be described as $\alpha - \theta$. It follows that $\overrightarrow{OP'}$ has length r (reflection preserves length), and oriented angle $\theta + (\alpha - \theta) + (\alpha - \theta) = 2\alpha - \theta$. Thus, in terms of polar coordinates we have

$$T_\alpha(x, y) = T(r\cos\theta, r\sin\theta)$$

$$= (r\cos(2\alpha - \theta), r\sin(2\alpha - \theta)) \qquad \text{(from the discussion above)}$$

$$= (r\cos(2\alpha)\cos(\theta) + r\sin(2\alpha)\sin(\theta), r\sin(2\alpha)\cos(\theta) - r\cos(2\alpha)\sin(\theta)) \qquad \text{(trig. identities)}$$

$$= (\cos(2\alpha)r\cos(\theta) + \sin(2\alpha)r\sin(\theta), \sin(2\alpha)r\cos(\theta) - \cos(2\alpha)r\sin(\theta)) \qquad \text{(algebra)}$$

$$= \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix} \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} \qquad \text{(identifying 2-vectors with column vectors)}$$

$$= \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

This proves $T_\alpha = T_A$ where

$$A = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}.$$

Since any function of the form T_A is linear, we conclude T_α is linear.

(4) Let $V = \mathbb{R}^\infty = \{(a_1, a_2, \dots,) \colon a_i \in \mathbb{R}\}$, the space of all infinite sequences. Define the "shift left function", T_ℓ, and "shift right function", T_r, as follows:

$$T_\ell \colon \mathbb{R}^\infty \to \mathbb{R}^\infty \qquad\qquad\qquad T_r \colon \mathbb{R}^\infty \to \mathbb{R}^\infty$$

$$s = (a_1, a_2, a_3, \dots) \longmapsto T_\ell(s) = (a_2, a_3, \dots) \qquad s = (a_1, a_2, a_3, \dots) \longmapsto T_r(s) = (0, a_1, a_2, \dots)$$

Prove that T_ℓ and T_r are linear transformations.

Solution 4:

Let $\mathbf{s} = (a_1, a_2, a_3, \dots)$ and $\mathbf{t} = (b_1, b_2, b_3, \dots)$ be two infinite sequences. Then for any scalars $c, d \in \mathbb{R}$ we have

$$\begin{aligned} T_\ell(c\mathbf{s} + d\mathbf{r}) &= T_\ell(ca_1 + db_1, ca_2 + db_2, \dots) & \text{(arith. of sequences)} \\ &= (ca_2 + db_2, ca_3 + db_3, \dots) & \text{(def. of } T_\ell) \\ &= c(a_2, a_3, \dots) + d(b_2, b_3, \dots) & \text{(arith. of sequences)} \\ &= cT_\ell(\mathbf{s}) + dT_\ell(\mathbf{t}) & \text{(def. of } T_\ell), \end{aligned}$$

and

$$T_r(c\mathbf{s} + d\mathbf{r}) = T_r(ca_1 + db_1, ca_2 + db_2, \dots) \qquad \text{(arith. of sequences)}$$
$$= (0, ca_1 + db_1, ca_2 + db_2, \dots) \qquad \text{(def. of } T_r)$$
$$= c(0, a_1, a_1, \dots) + d(0, b_1, b_2, \dots) \qquad \text{(arith. of sequences)}$$
$$= cT_r(\mathbf{s}) + dT_r(\mathbf{t}) \qquad \text{(def. of } T_r).$$

This proves T_ℓ and T_r are both linear transformations.

(5) (Conjugation). Fix an invertible matrix $Q \in M_{nn}$. Define $T\colon M_{nn} \to M_{nn}$ as $T(A) = QAQ^{-1}$. Show that T is a linear transformation. (This operation is called **conjugation by** Q.)

Solution 5:

We show that $T(cA + dB) = cT(A) + dT(B)$ for any $c, d \in \mathbb{R}$ and any $A, B \in M_{nn}$:

$$T(cA + dB) = Q(cA + dB)Q^{-1} \qquad \text{(def. of } T)$$
$$= QcAQ^{-1} + QdBQ^{-1} \qquad \text{(left- and right-distributivity of matrix mult.)}$$
$$= cQAQ^{-1} + dQBQ^{-1} \qquad \text{(scalars commute)}$$
$$= cT(A) + dT(B) \qquad \text{(def. of } T)$$

3.3: SUBSPACES

(1) Given any square matrix $A = [a_{ij}] \in M_{nn}$, we define the **trace** of A as $\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}$. In other words, the trace $\operatorname{tr} A$ is the sum of the diagonal entries of A.
Let $V = M_{22}$, and define $W = \{A \in M_{22}\colon \operatorname{tr}(A) = 0\}$.
 (a) Prove that W is a subspace.
 (b) Find a spanning set for the subspace
$$W = \{A \in M_{22}\colon \operatorname{tr}(A) = 0\}.$$

Solution 1:

(2) Let $V = M_{22}$, and define $A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Compute $W = \operatorname{span}(\{A_1, A_2, A_3\})$, identifying it as a certain *familiar* set of matrices.

Solution 2:
Using the definition of span, we compute
$$\operatorname{span}\{A_1, A_2, A_3\} = \{c_1 A_1 + c_2 A_2 + c_3 A_3 \colon c_i \in \mathbb{R}\}$$
$$= \left\{ \begin{bmatrix} c_1 + c_3 & c_1 + c_2 + c_3 \\ c_1 + c_2 + c_3 & c_1 \end{bmatrix} \colon c_i \in \mathbb{R} \right\}.$$

We claim this last set is none other than the set of all symmetric matrices. It is clear that each element in the set above is symmetric, so it remains only to show that if A is symmetric, then A can be written in the above form for some choice of c_1, c_2, c_3.

Write $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. We need to find c_1, c_2, c_3 such that

$$\begin{bmatrix} c_1 + c_3 & c_1 + c_2 + c_3 \\ c_1 + c_2 + c_3 & c_1 \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

The bottom-right entry tells us that we must pick $c_1 = c$. Then the top-left entry tells us that $c_3 = a - c$. Lastly, the off diagonal entries tells us that $c_2 = b - a$.

Thus we can write $A = cA_1 + (b-a)A_2 + (a-c)A_3$, showing $A \in \text{span}\{A_1, A_2, A_3\}$, and proving that $\text{span}\{A_1, A_2, A_3\} = $ (set of symmetric matrices).

(3) Let $V = \mathbb{R}^4$ and define $W = \{(x_1, x_2, x_3, x_4): x_1 - x_4 = x_2 + x_3 = x_1 + x_2 + x_3 = 0\}$.
  (a) Show W is a subspace by identifying it as $\text{null}\,T$ for an explicit linear transformation.
  (b) Describe W in as simple a manner as possible.

Solution 3:
(a) The subset W can be described as the set of solutions to the linear system

(*)
$$\begin{aligned} x_1 + \quad\quad\quad - x_4 &= 0 \\ x_2 + x_3 \quad\quad &= 0 \\ x_1 + x_2 + x_3 \quad\quad &= 0 \end{aligned}$$

We can express this with the matrix equation $A\mathbf{x} = \mathbf{0}$, where

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

We see then that $W = \text{null}\,T_A = \text{null}\,A$. This proves W is a subspace, since T_A is a linear transformation.
(b) We use Gaussian elimination to help describe W. The system (*) has corresponding augmented maqtrix $[A|\mathbf{0}]$, which row reduces to

$$\left[\begin{array}{cccc|c} \boxed{1} & 0 & 0 & -1 & 0 \\ 0 & \boxed{1} & 1 & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 0 \end{array}\right]$$

Thus $x_3 = t$ is the only free variable, and we back-substitute to find $(x_1, x_2, x_3, x_4) = (0, -t, t, 0)$. We conclude that $W = \{(0, -t, t, 0): t \in \mathbb{R}\}$.

(4) Determine which of the following subsets W are subspaces of M_{nn}.
  (a) $W = \{\text{diagonal } n \times n \text{ matrices}\}$
  (b) $W = \{A \in M_{nn}: \det A = 0\}$
  (c) $W = \{\text{upper triangular } n \times n \text{ matrices}\}$
  (d) $W = \{A \in M_{nn}: \text{tr}\,A = 0\}$. (See Exercise 1 above for the definition of tr A.)
  (e) $W = \{\text{symmetric } n \times n \text{ matrices}\} = \{A \in M_{nn}: A^T = A\}$.
  (f) $W = \{A \in M_{nn}: A^T = -A\}$. (These are called **skew-symmetric matrices**.)
  (g) The set W of all n by n matrices A for which $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  (h) The set W of all n by n matrices A such that $AB = BA$, where B is a fixed $n \times n$ matrix.

Solution 4:

(a) Yes, let

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}$$

Then,

$$A + B = \begin{bmatrix} a_{11} + b_{11} & 0 & \cdots & 0 \\ 0 & a_{22} + b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} + b_{nn} \end{bmatrix} \in M$$

And, for some scalar k

$$kA = k\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} ka_{11} & 0 & \cdots & 0 \\ 0 & ka_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & ka_{nn} \end{bmatrix} \in M$$

(b) No, consider

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} B = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} A + B = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

Thus $\det(A) = 0$ and $\det(B) = 0$, but $\det(A + B) = 16 - 4 = 12$. So addition is not closed.
(c) Yes. Since $\underset{n \times n}{\mathbf{0}}$ is upper triangular, we have $\underset{n \times n}{\mathbf{0}} \in W$. If $A = [a_{ij}]$ and $B = [b_{ij}]$ are upper triangular, then $a_{ij} = b_{ij} = 0$ for all $j > 1$. Then we also have $a_{ij} + b_{ij} = 0$ for all $j > i$, showing that $A + B$ is upper triangular. Hence W is closed under addition. A similar argument shows W is closed under scalar multiplication.
(d) Yes. We prove by first showing that the map $\operatorname{tr}\colon M_{nn} \to \mathbb{R}$ is a linear transformation. Since $W = \{A \in M_{nn}\colon \operatorname{tr} A = 0\} = \operatorname{null} \operatorname{tr}$, our null space theorem implies W is a subspace.
To prove $\operatorname{tr}\colon M_{nn} \to \mathbb{R}$ is linear, we observe that

$$\begin{aligned} \operatorname{tr}(cA + dB) &= \sum_{i=1}^{n}(cA + dB)_{ii} \\ &= \sum_{i=1}^{n} c(A)_{ii} + d(B)_{ii} \\ &= c\sum_{i=1}^{n}(A)_{ii} + d\sum_{i=1}^{n}(B)_{ii} \\ &= c\operatorname{tr} A + d\operatorname{tr} B \end{aligned}$$

(e) Yes. We prove this in the usual fashion. Remember: a matrix A is symmetric iff $A^T = A$. Thus $W = \{A \in M_{nn}\colon A^T = A\}$.
 (i) $\mathbf{0} = 0_{nn}$ is clearly symmetric. Hence $\mathbf{0} \in W$.
(ii) Suppose $A_1, A_2 \in W$. Then $(A_i)^T = A_i$ for $i = 1, 2$. But then

$$\begin{aligned} (A_1 + A_2)^T &= A_1^T + A_2^T && \text{(prop. of transp.)} \\ &= A_1 + A_2 && \text{(since } A_i \in W) \end{aligned}$$

Thus $(A_1 + A_2)^T = A_1 + A_2$, showing $A_1 + A_2 \in W$.

(iii) Suppose $A \in W$. Then $A^T = A$. But then

$$(kA)^T = kA^T \qquad \text{(prop. of transp.)}$$
$$= kA \qquad \text{(since } A \in W\text{)}$$

Thus $kA \in W$.

(f) Yes. The proof of this looks nearly exactly like the one above, only now we are always showing a matrix B satisfies $B^T = -B$. The exact same properties of transpose will be invoked.

(g) No. The zero vector in M_{nn} is $\underset{n \times n}{\mathbf{0}}$, the matrix with all zeros. Clearly the matrix equation $\underset{n \times n}{\mathbf{0}} \, \mathbf{x} = \mathbf{0}$ has more than just the trivial solution; in fact all vectors $\mathbf{x}$ satisfy this equation! Thus $\underset{n \times n}{\mathbf{0}} \notin W$, showing that W is not a subspace.

(h) Yes.

(i) Clearly $\mathbf{0} = 0_{nn}$ satisfies $0_{nn} B = B 0_{nn}$. Thus $\mathbf{0} \in W$.

(ii) Next, let $A_1, A_2 \in W$. Then

$$(A_1 + A_2)B = A_1 B + A_2 B = B A_1 + B A_2 + B(A_1 + A_2)$$

So, $A_1 + A_2 \in W$.

(iii) Finally, suppose $A_1 \in W$ and let k be any scalar. Then

$$(kA_1)B = k(A_1 B) = k(B A_1) = (kB)A_1 = B(kA_1)$$

Thus $kA_1 \in W$.

(5) Determine which of the following are subspaces of P_3.

(a) The set W of all polynomials $a_0 + a_1 x + a_2 x^2 + a_3 x^3$ for which $a_0 = 0$

(b) The set W of all polynomials $a_0 + a_1 x + a_2 x^2 + a_3 x^3$ for which $a_0 + a_1 + a_2 + a_3 = 0$.

(c) All polynomials of the form $a_0 + a_1 x + a_2 x^2 + a_3 x^3$ in which a_0, a_1, a_2, a_3 are rational numbers.

(d) All polynomials of the form $a_0 + a_1 x$, where a_0, a_1 are real numbers.

Solution 5:

(a) Yes.

(i) First off the zero vector $\mathbf{0}$ in this case is the zero polynomial $p(x) = 0 + 0x + 0x^2 + 0x^3$. This is clearly an element of W.

(ii) Next, suppose $a_0 = b_0 = 0$. Then,

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + b_0 + b_1 x + b_2 x^2 + b_3 x^3 = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3$$
$$= 0 + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3$$

Thus addition is closed.

(iii) Finally, let k be some scalar. Then,

$$k(a_0 + a_1 x + a_2 x^2 + a_3 x^3) = ka_0 + ka_1 x + ka_2 x^2 + ka_3 x^3$$
$$= 0 + ka_1 x + ka_2 x^2 + ka_3 x^3$$

Thus scalar multiplication is also closed.

(b) Yes.

(i) First off, clearly the zero element $p(x) = 0 + 0x + 0x^2 + 0x^3$ is in W.

(ii) Next, suppose $a_0 + a_1 + a_2 + a_3 = 0$ and $b_0 + b_1 + b_2 + b_3 = 0$. Then,

$$(a_0 + a_1 x + a_2 x^2 + a_3 x^3) + (b_0 + b_1 x + b_2 x^2 + b_3 x^3)$$
$$= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3$$

So,

$$(a_0 + b_0) + (a_1 + b_1) + (a_2 + b_2) + (a_3 + b_3)$$
$$= (a_0 + a_1 + a_2 + a_3) + (b_0 + b_1 + b_2 + b_3)$$
$$= 0 + 0 = 0$$

Thus addition is closed.

(iii) Lastly, let k be some scalar. Then,

$$k(a_0 + a_1 x + a_2 x^2 + a_3 x^3) = ka_0 + ka_1 x + ka_2 x^2 + ka_3 x^3$$

So,

$$ka_0 + ka_1 + ka_2 + ka_3 = k(a_0 + a_1 + a_2 + a_3) = k(0) = 0$$

Thus scalar multiplication is also closed.

(c) No. Let $p(x) = 1 + x + x^2$, a polynomial with rational coefficients. Then $\sqrt{2}p(x) = \sqrt{2} + \sqrt{2}x + \sqrt{2}x^2$ does not have rational coefficients, hence does not lie in the set.

(d) Yes. This is just the subspace P_1, which we have already shown to be a subspace.

(6) Which of the following are subspaces of $F(-\infty, \infty)$?
   (a) The set W of all functions f in $F(-\infty, \infty)$ for which $f(0) = 0$.
   (b) All functions f in $F(-\infty, \infty)$ for which $f(0) = 1$.
   (c) The set W of all functions f in $F(-\infty, \infty)$ for which $f(-x) = x$.
   (d) All polynomials of degree 2.

Solution 6:

(a) Yes. First note that the zero function clearly satisfies the stated condition. Thus $\mathbf{0} \in W$.
   Now suppose $f(0) = 0$ and $g(0) = 0$. Then

$$(f + g)(0) = f(0) + g(0) = 0 + 0 = 0$$

And for some scalar k,

$$k(f(0)) = k(0) = 0$$

(b) No not closed under addition. Consider $f(0) = 1$ and $g(0) = 1$, then

$$(f + g)(0) = f(0) + g(0) = 1 + 1 = 2 \neq 1$$

(c) Yes. Again, clearly the zero function satisfies this condition. Thus $\mathbf{0} \in W$.
   Next, suppose $f(-x) = x$ and $g(-x) = x$, then

$$(f + g)(x) = f(x) + g(x) = f(-x) + g(-x) = (f + g)(-x)$$

And for some scalar k,

$$(kf)(x) = k(f(x)) = k(f(-x)) = (kf)(-x)$$

(d) No, consider $-x^2$ and $x^2 + 3x + 1$. Then

$$-x^2 + (x^2 + 3x + 1) = 3x + 1$$

which is not a polynomial of degree 2.

(7) Express the following as linear combinations of $\mathbf{u} = (2, 1, 4)$, $\mathbf{v} = (1, -1, 3)$, and $\mathbf{w} = (3, 2, 5)$
(a) $(-9, -7, -15)$
(b) $(6, 11, 6)$
(c) $(0, 0, 0)$

Solution 7:
(a)

$$\begin{bmatrix} 2 & 1 & 3 & -9 \\ 1 & -1 & 2 & -7 \\ 4 & 3 & 5 & -15 \end{bmatrix} \xrightarrow{1/2r_1 - r_2} \begin{bmatrix} 2 & 1 & 3 & -9 \\ 0 & 3/2 & -1/2 & 5/2 \\ 4 & 3 & 5 & -15 \end{bmatrix}$$

$$\xrightarrow{2r_1 - r_3} \begin{bmatrix} 2 & 1 & 3 & -9 \\ 0 & 3/2 & -1/2 & 5/2 \\ 0 & -1 & 1 & -3 \end{bmatrix}$$

$$\xrightarrow{1/2r_1} \begin{bmatrix} 1 & 1/2 & 3/2 & -9/2 \\ 0 & 3/2 & -1/2 & 5/2 \\ 0 & -1 & 1 & -3 \end{bmatrix}$$

$$\xrightarrow{2/3r_2} \begin{bmatrix} 1 & 1/2 & 3/2 & -9/2 \\ 0 & 1 & -1/3 & 5/3 \\ 0 & -1 & 1 & -3 \end{bmatrix}$$

$$\xrightarrow{r_2 + r_3} \begin{bmatrix} 1 & 1/2 & 3/2 & -9/2 \\ 0 & 1 & -1/3 & 5/3 \\ 0 & 0 & 2/3 & -4/3 \end{bmatrix}$$

$$\xrightarrow{3/2r_3} \begin{bmatrix} 1 & 1/2 & 3/2 & -9/2 \\ 0 & 1 & -1/3 & 5/3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

Solving by back substitution,

$$\begin{aligned} x_3 &= -2 \\ x_2 &= 5/3 - 2/3 = 1 \\ x_1 &= -9/2 + 3 - 1/2 = -2 \end{aligned}$$

Thus

$$-2\mathbf{u} + 1\mathbf{v} + (-2)\mathbf{w} = (-9, -7, -15)$$

(b)

$$\begin{bmatrix} 2 & 1 & 3 & | & 6 \\ 1 & -1 & 2 & | & 11 \\ 4 & 3 & 5 & | & 6 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{bmatrix} 1 & -1 & 2 & | & 11 \\ 2 & 1 & 3 & | & 6 \\ 4 & 3 & 5 & | & 6 \end{bmatrix}$$

$$\xrightarrow{2r_1 - r_2} \begin{bmatrix} 1 & -1 & 2 & | & 11 \\ 0 & -3 & 1 & | & 16 \\ 4 & 3 & 5 & | & 6 \end{bmatrix}$$

$$\xrightarrow{4r_1 - r_3} \begin{bmatrix} 1 & -1 & 2 & | & 11 \\ 0 & -3 & 1 & | & 16 \\ 0 & -7 & 3 & | & 38 \end{bmatrix}$$

$$\xrightarrow{-1/3r_2} \begin{bmatrix} 1 & -1 & 2 & | & 11 \\ 0 & 1 & -1/3 & | & -16/3 \\ 0 & -7 & 3 & | & 38 \end{bmatrix}$$

$$\xrightarrow{7r_2 + r_3} \begin{bmatrix} 1 & -1 & 2 & | & 11 \\ 0 & 1 & -1/3 & | & -16/3 \\ 0 & 0 & 2/3 & | & 2/3 \end{bmatrix}$$

Solving by back substitution,

$$\begin{aligned} x_3 &= 1 \\ x_2 &= -16/3 + 1/3 = -5 \\ x_1 &= 11 - 2 - 5 = 4 \end{aligned}$$

Thus

$$4\mathbf{u} + (-5)\mathbf{v} + 1\mathbf{w} = (6, 11, 6)$$

(c)

$$0\mathbf{u} + 0\mathbf{v} + 0\mathbf{w} = (0, 0, 0)$$

(8) Which are the following are linear combinations of:

$$A = \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix}, B = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}, C = \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix}?$$

(a) $\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} -1 & 5 \\ 7 & 1 \end{bmatrix}$

Solution 8:

(a) Yes, this is a linear combination.

$$
\begin{bmatrix} 4 & 1 & 0 & 6 \\ 0 & -1 & 2 & -8 \\ -2 & 2 & 1 & -1 \\ -2 & 3 & 4 & -8 \end{bmatrix}
\xrightarrow{r_1+2r_3}
\begin{bmatrix} 4 & 1 & 0 & 6 \\ 0 & -1 & 2 & -8 \\ 0 & 5 & 2 & 4 \\ -2 & 3 & 4 & -8 \end{bmatrix}
$$

$$
\xrightarrow{r_1+2r_4}
\begin{bmatrix} 4 & 1 & 0 & 6 \\ 0 & -1 & 2 & -8 \\ 0 & 5 & 2 & 4 \\ 0 & 7 & 8 & -10 \end{bmatrix}
$$

$$
\xrightarrow{-r_2}
\begin{bmatrix} 4 & 1 & 0 & 6 \\ 0 & 1 & -2 & 8 \\ 0 & 5 & 2 & 4 \\ 0 & 7 & 8 & -10 \end{bmatrix}
$$

$$
\xrightarrow{1/4r_1}
\begin{bmatrix} 1 & 1/4 & 0 & 3/2 \\ 0 & 1 & -2 & 8 \\ 0 & 5 & 2 & 4 \\ 0 & 7 & 8 & -10 \end{bmatrix}
$$

$$
\xrightarrow{5r_2-r_3}
\begin{bmatrix} 1 & 1/4 & 0 & 3/2 \\ 0 & 1 & -2 & 8 \\ 0 & 0 & -12 & 36 \\ 0 & 7 & 8 & -10 \end{bmatrix}
$$

$$
\xrightarrow{7r_2-r_4}
\begin{bmatrix} 1 & 1/4 & 0 & 3/2 \\ 0 & 1 & -2 & 8 \\ 0 & 0 & -12 & 36 \\ 0 & 0 & -22 & 66 \end{bmatrix}
$$

$$
\xrightarrow{-1/12r_3}
\begin{bmatrix} 1 & 1/4 & 0 & 3/2 \\ 0 & 1 & -2 & 8 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & -22 & 66 \end{bmatrix}
$$

$$
\xrightarrow{22r_3+r_4}
\begin{bmatrix} 1 & 1/4 & 0 & 3/2 \\ 0 & 1 & -2 & 8 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}
$$

Now by substitution:

$$
\begin{aligned}
x_3 &= -3 \\
x_2 &= 8 + 2(-3) = 2 \\
x_1 &= 3/2 - 1/4(2) = 1
\end{aligned}
$$

(b) Yes, this is a linear combination.

$$
0A + 0B + 0C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
$$

(c) No, this is not a linear combination

$$\begin{bmatrix} 4 & 1 & 0 & -1 \\ 0 & -1 & 2 & 5 \\ -2 & 2 & 1 & 7 \\ -2 & 3 & 4 & 1 \end{bmatrix} \xrightarrow{1/2r_1+r_3} \begin{bmatrix} 4 & 1 & 0 & -1 \\ 0 & -1 & 2 & 5 \\ 0 & 5/2 & 1 & 13/2 \\ -2 & 3 & 4 & 1 \end{bmatrix}$$

$$\xrightarrow{1/2r_1+r_4} \begin{bmatrix} 4 & 1 & 0 & -1 \\ 0 & -1 & 2 & 5 \\ 0 & 5/2 & 1 & 13/2 \\ 0 & 7/2 & 4 & 1/2 \end{bmatrix}$$

$$\xrightarrow{-r_2} \begin{bmatrix} 4 & 1 & 0 & -1 \\ 0 & 1 & -2 & -5 \\ 0 & 5/2 & 1 & 13/2 \\ 0 & 7/2 & 4 & 1/2 \end{bmatrix}$$

$$\xrightarrow{5/2r_2-r_3} \begin{bmatrix} 4 & 1 & 0 & -1 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & -6 & -19 \\ 0 & 7/2 & 4 & 1/2 \end{bmatrix}$$

$$\xrightarrow{7/2r_2-r_4} \begin{bmatrix} 4 & 1 & 0 & -1 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & -6 & -19 \\ 0 & 0 & -11 & -18 \end{bmatrix}$$

$$\xrightarrow{-1/6r_3} \begin{bmatrix} 4 & 1 & 0 & -1 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 1 & 19/6 \\ 0 & 0 & -11 & -18 \end{bmatrix}$$

$$\xrightarrow{11r_3+r_4} \begin{bmatrix} 4 & 1 & 0 & -1 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 1 & 19/6 \\ 0 & 0 & 0 & 101/6 \end{bmatrix}$$

The system is not consistent, thus it cannot be a linear combination.

(9) Express $p(x) = 7 + 8x + 9x^2$ as a linear combination of

$$p_1 = 2 + x + 4x^2, p_2 = 1 - x + 3x^2, p_3 = 3 + 2x + 5x^2$$

Solution 9:

We have p equal to a linear combination of p_1, p_2, p_3 if and only if $p = ap_1 + bp_2 + cp_3$ for some $a, b, c \in \mathbb{R}$ iff

$$7+8x+9x^2 = a(2+x+4x^2)+b(1-x+3x^2)+c(3+2x+5x^2) = (2a+b+3c)+(a-b+2c)x+(4a+3b+5c)x^2.$$

Equating coefficients, we see that p is a linear combination of p_1, p_2, p_3 iff the following system is consistent:

$$2a + b + 3c = 7$$
$$a - b + 2c = 8$$
$$4a + 3b + 5c = 9$$

Now perform Gaussian elimination:

$$
\begin{bmatrix} 2 & 1 & 3 & 7 \\ 1 & -1 & 2 & 8 \\ 4 & 3 & 5 & 9 \end{bmatrix}
\xrightarrow{r_1 \leftrightarrow r_2}
\begin{bmatrix} 1 & -1 & 2 & 8 \\ 2 & 1 & 3 & 7 \\ 4 & 3 & 5 & 9 \end{bmatrix}
$$

$$
\xrightarrow{2r_1 - r_2}
\begin{bmatrix} 1 & -1 & 2 & 8 \\ 0 & -3 & 1 & 9 \\ 4 & 3 & 5 & 9 \end{bmatrix}
$$

$$
\xrightarrow{4r_1 - r_3}
\begin{bmatrix} 1 & -1 & 2 & 8 \\ 0 & -3 & 1 & 9 \\ 0 & -7 & 3 & 23 \end{bmatrix}
$$

$$
\xrightarrow{-1/3 r_2}
\begin{bmatrix} 1 & -1 & 2 & 8 \\ 0 & 1 & -1/3 & -3 \\ 0 & -7 & 3 & 23 \end{bmatrix}
$$

$$
\xrightarrow{7r_2 + r_3}
\begin{bmatrix} 1 & -1 & 2 & 8 \\ 0 & 1 & -1/3 & -3 \\ 0 & 0 & 2/3 & 2 \end{bmatrix}
$$

$$
\xrightarrow{3/2 r_3}
\begin{bmatrix} 1 & -1 & 2 & 8 \\ 0 & 1 & -1/3 & -3 \\ 0 & 0 & 1 & 3 \end{bmatrix}
$$

Now by substitution

$$
\begin{aligned}
c &= 3 \\
b &= -3 + 1 = -2 \\
a &= 8 - 6 - 2 = 0
\end{aligned}
$$

Thus

$$
7 + 8x + 9x^2 = 0 p_1 - 2 p_2 + 3 p_3
$$

(10) Determine whether the following polynomials span P_2.

$$
p_1 = 1 - x + 2x^2, \; p_2 = 3 + x, \; p_3 = 5 - x + 4x^2, \; p_4 = -2 - 2x + 2x^2
$$

Solution 10:

Reasoning exactly as in the previous exercise, we see that a polynomial $p(x) = a + bx + cx^2$ is a linear combination of the p_i if and only if we have $p = k_1 p_1 + k_2 p_2 + k_3 p_3 + k_4 p_4$ for some $k_i \in \mathbb{R}$. After expanding the expression on the right and equating coefficients, we conclude that p lies in the span of the p_i if and only if the system

$$
\begin{aligned}
k_1 + 3k_2 + 5k_3 - 2k_4 &= a \\
-k_1 + k_2 - k_3 - 2k_4 &= b \\
2k_0 + 0k_2 + 4k_3 + 2k_4 &= c
\end{aligned}
$$

63

Now apply Gaussian elimination.

$$\left[\begin{array}{cccc|c} 1 & 3 & 5 & -2 & a \\ -1 & 1 & -1 & -2 & b \\ 2 & 0 & 4 & 2 & c \end{array}\right] \xrightarrow{r_2+r_1} \left[\begin{array}{cccc|c} 1 & 3 & 5 & -2 & a \\ 0 & 4 & 4 & -4 & a+b \\ 2 & 0 & 4 & 2 & c \end{array}\right]$$

$$\xrightarrow{r_3-2r_1} \left[\begin{array}{cccc|c} 1 & 3 & 5 & -2 & a \\ 0 & 4 & 4 & -4 & a+b \\ 0 & 6 & 6 & -6 & c-2a \end{array}\right]$$

$$\xrightarrow{r_3-3/2r_2} \left[\begin{array}{cccc|c} 1 & 3 & 5 & -2 & a \\ 0 & 4 & 4 & -4 & a+b \\ 0 & 0 & 0 & 0 & -7/2a-3/2b+c \end{array}\right]$$

We see the system will be consistent if and only if $-7/2a - 3/2b + c = 0$ (to avoid having a leading 1 in the last column). For example, if we pick $a = 0, b = 0, c = 1$, the system is not consistent. This means the corresponding polynomial, $p(x) = 0 + 0x + 1x^2 = x^2$ is not in the span of the p_i.

(11) Let $f(x) = \cos^2 x$ and $g(x) = \sin^2 x$. Determine whether the following functions lie in the space spanned by f and g. If yes, provide an explicit linear combination; if no, show no linear combination exists. (You may want to review your trig identities.)

(a) $q(x) = \cos 2x$
(b) $q(x) = 3 + x^2$
(c) $q(x) = 1$
(d) $q(x) = \sin x$
(e) $q(x) = 0$

Solution 11:

(a) Yes.
$$f - g = \cos^2 x - \sin^2 x = \cos 2x$$

(b) No. Assume we can find k_1, k_2 such that
$$3 + x^2 = k_1 \cos^2 x + k_2 \sin^2 x$$
for all x.

Evaluating this identity at $x = 0$, we see that we must have $3 = k_1$.

However, evaluating the same identity at $x = \pi$ we see that $3 + \pi^2 = k_1$. Thus $k_1 = 3$, and $k_1 = 3 + \pi^2$, a contradiction.

We conclude there can be no such k_1, k_2.

(c) Yes.
$$f + g = \cos^2 x + \sin^2 x = 1$$

(d) No, assume we can find k_1, k_2 such that
$$\sin x = k_1 \cos^2 x + k_2 \sin^2 x$$
for all x. Then
$$\begin{aligned} \sin(-x) &= k_1 \cos^2(-x) + k_2 \sin^2(-x) \\ &= k_1 \cos^2(x) + k_2 \sin^2(x) \\ &= \sin(x) \end{aligned}$$

But it is not the case that $\sin(x) = \sin(-x)$ for all x. Thus there are no such k_i, and $\sin(x)$ is not in the span.

(e) Yes.

$$0f + 0g = 0$$

(12) Show that the solution vectors of a consistent nonhomogeneous system of m linear equations in n unknowns do not form a subspace of R^n.

Solution 12:

Represent the system as a matrix equation of the form

$$A\mathbf{x} = \mathbf{y},$$

where $\mathbf{y} \neq \mathbf{0}$, since the system is assumed to be nonhomogeneous.

We can describe the set of solutions as $W = \{\mathbf{x}\colon A\mathbf{x} = \mathbf{y}\}$. Then W is not a subspace as $\mathbf{0} \notin W$. Indeed $A\mathbf{0} = \mathbf{0} \neq \mathbf{y}$.

(13) Define the function $T\colon P_2 \to \mathbb{R}^2$ by

$$T(p(x)) = \begin{bmatrix} p(1) \\ p(-1) \end{bmatrix}.$$

(a) Prove T is a linear transformation.

(b) Compute $\mathrm{null}(T)$ and $\mathrm{range}(T)$ and describe these in as simple a manner as possible.

Solution 13:

(a)

$$
\begin{aligned}
T(cp_1 + dp_2) &= \begin{bmatrix} (cp_1 + dp_2)(1) \\ (cp_1 + dp_2)(-1) \end{bmatrix} \\
&= \begin{bmatrix} cp_1(1) + dp_2(1) \\ cp_1(-1) + dp_2(-1) \end{bmatrix} \\
&= c \begin{bmatrix} p_1(1) \\ p_1(-1) \end{bmatrix} + d \begin{bmatrix} p_2(1) \\ p_2(-1) \end{bmatrix} \\
&= cT(p_1) + dT(p_2)
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{null}(T) &= \{p \in P_2\colon T(p) = \mathbf{0}\} \\
&= \{p \in P_2\colon \begin{bmatrix} p(1) \\ p(-1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\} \\
&= \{p \in P_2\colon p(1) = p(-1) = 0\}.
\end{aligned}
$$

We must now describe all polynomials $p(x) = ax^2 + bx + c$ satisfying $p(1) = p(-1) = 0$. You could set up the system of equations by evaluating and then solve directly for a, b, c. Alternatively, you can observe that a degree 2 polynomial satisfying $p(1) = p(-1) = 0$ must *factor* as

$$p(x) = c(x - 1)(x + 1) = c(x^2 - 1) = cx^2 - c.$$

It follows that $\mathrm{null}\,T = \{p(x) \in P_2\colon p(x) = cx^2 - c \text{ for some } c \in \mathbb{R}\} = \mathrm{span}(\{x^2 - 1\})$.

The range of T is the set of vectors (b_1, b_2) such that $(b_1, b_2) = (p(1), p(-1))$ for some $p(x) = a_0 + a_1 x + a_2 x^2$: i.e., the set of pairs (b_1, b_2) for which the system

$$
\begin{aligned}
a_0 + a_1 + a_2 &= b_1 \\
a_0 - a_1 + a_2 &= b_2
\end{aligned}
$$

is consisent. Perform Gaussian elimination on the corresponding augmented matrix:

$$\begin{bmatrix} 1 & 1 & 1 & b_1 \\ 1 & -1 & 1 & b_2 \end{bmatrix} \xrightarrow{\text{row red.}} \begin{bmatrix} 1 & 1 & 1 & b_1 \\ 0 & 1 & 0 & \frac{b_1-b_2}{2} \end{bmatrix}$$

We see the system is always consistent, and hence that all 2-vectors are included in the range: i.e., $\operatorname{range} T = \mathbb{R}^2$.

(14) Define $T\colon P_2 \to P_3$ by $T(p) = q$, where $q(x) = \int_0^x p(t)\ dt$.

    (a) Prove that T is a linear transformation.

    (b) Compute $\operatorname{null}(T)$ and $\operatorname{range}(T)$ and describe these spaces in as simple a manner as possible.

Solution 14:

(a)

$$T(cp + dq) = \int_0^x cp(t) + dq(t)\ dt \qquad\qquad \text{(by def.)}$$

$$= c\int_0^x p(t)\ dt + d\int_0^x q(t)\ dt \qquad\qquad \text{(by calc.)}$$

$$= cT(p) + dT(q) \qquad\qquad \text{(by def)}$$

(b) For an arbitrary polynomial $p(x) = ax^2 + bx + c$, we have $T(p) = \frac{1}{3}ax^3 + \frac{1}{2}bx^2 + cx$. Clearly $T(p) = \mathbf{0}$ if and only if $a = b = c = 0$. Thus $\operatorname{null}(T) = \{\mathbf{0}\}$.
We clearly have $\operatorname{range}(T) \subseteq W = \{q(x) \in P_3\colon q(x) = dx^3 + ex^2 + fx\}$: the set of polynomials whose constant term is zero. Furthermore, given any $q(x) = dx^3 + ex^2 + fx \in W$, we have $q = T(3dx^2 + 2ex + f)$, showing that $W \subseteq \operatorname{range}(T)$. We conclude $W = \operatorname{range}(T)$.

(15) Define $S\colon M_{nn} \to M_{nn}$ by $S(A) = A^T - A$.

    (a) Show S is linear.

    (b) Compute $\operatorname{null}(S)$ and $\operatorname{range}(S)$ and identify both of these spaces as certain special named families of matrices. Make sure to prove your claims.

Solution 15:

(a)

$$S(cA + dB) = (cA + dB)^T - (cA + dB) \qquad\qquad \text{(by def)}$$

$$= cA^T + dB^T - cA - cB \qquad\qquad \text{(transpose prop.)}$$

$$= c(A^T - A) + d(B^T - B) \qquad\qquad \text{(matrix arith.)}$$

$$= cS(A) + dS(B) \qquad\qquad \text{(by def)}$$

(b) $S(A) = \mathbf{0}$ iff $A^T - A = \mathbf{0}$ iff $A^T = A$. Thus $\operatorname{null}(A)$ is the set of symmetric matrices. Now we identify $\operatorname{range} T$. Let $B = S(A) = A^T - A$. Then $B^T = (A - A^T) = -(A^T - A) = -B$. This shows $\operatorname{range}(T)$ is a subset of the space of skew-symmetric matrices. Does $\operatorname{range} T$ include *all* skew-symmetric matrices? Suppose B is skew, so that $B^T = -B$. Set $A = -\frac{1}{2}B$. Then

$$T(A) = (-\frac{1}{2}B)^T - \frac{1}{2}B) = -\frac{1}{2}(B^T - B) = -\frac{1}{2}(-B - B) = B.$$

This shows directly that $B \in \operatorname{range} T$, and thus that $\operatorname{range} T = \{\text{skew-symmetric } n \times n \text{ matrices}\}$.

(16) Define $T: P_3 \to M_{22}$ by $T(p) = \begin{bmatrix} p(1) & p(-1) \\ p(2) & p(-2) \end{bmatrix}$.

    (a) Show T is linear.

    (b) Compute null(T) and range(T).

       Hint: a nonzero polynomial of degree 3 can have at most 3 roots!

Solution 16:

(a)

$$T(cp + dq) = \begin{bmatrix} (cp+dq)(1) & (cp+dq)(-1) \\ (cp+dq)(2) & (cp+dq)(-2) \end{bmatrix} \quad\quad \text{(by def.)}$$

$$= \begin{bmatrix} cp(1)+dq(1) & cp(-1)+dq(-1) \\ cp(2)+dq(2) & cp(-2)+dq(-2) \end{bmatrix} \quad\quad \text{(function arith.)}$$

$$= c\begin{bmatrix} p(1) & p(-1) \\ p(2) & p(-2) \end{bmatrix} + d\begin{bmatrix} q(1) & q(-1) \\ q(2) & q(-2) \end{bmatrix} \quad\quad \text{(matrix arith.)}$$

$$= cT(p) + dT(q) \quad\quad \text{(by def)}$$

(b) We have $T(p) = \mathbf{0}$ iff $p(1) = p(-1) = p(2) = p(-2) = 0$ iff $p(x) = \mathbf{0}$, since a nonzero polynomial of degree at most 3 can have at most 3 distinct roots! Thus null$(T) = \{\mathbf{0}\}$ is the trivial subspace.

For given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, deciding whether there is a $p(x) = a_0 + a_1 x + a_2 x^2 + a_3$ such that $T(p) = A$ amounts to deciding whether the system

$$\begin{array}{rrrrl} a_0 + & a_1 + & a_2 + & a_3 &= a \\ a_0 - & a_1 + & a_2 - & a_3 &= b \\ a_0 + & 2a_1 + & 4a_2 + & 8a_3 &= c \\ a_0 - & 2a_1 + & 4a_2 - & 8a_3 &= d \end{array}$$

is consistent. This in turn can answered by deciding when the following matrix equation can be solved:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 2 & 4 & 8 \\ 1 & -2 & 4 & -8 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

A determinant computation shows that the given 4×4 matrix is invertible! (In fact this is an instance of the Vandermonde determinant!) By the invertibility theorem, for any choice of a, b, c, d there is *always* a solution. Thus range $T = M_{22}$.

(17) Define the function $T: C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ by $T(f) = f'$.

    (a) Show T is a linear transformation.

    (b) Compute null(T) and range T.

Solution 17:

(a) Shown in the notes. To repeat: $T(cf + dg) = (cf + dg)' = cf' + dg'$, using familiar properties of the derivative.

(b) null$(T) = \{f : f' = 0\} = \{$constant functions$\}$, again using a familiar fact from calculus. I claim range $T = C^\infty(\mathbb{R})$. Indeed, any $g \in C^\infty(\mathbb{R})$ is continuous, and thus has an antiderivative by the fundamental theorem of calculus. Thus there is a function f such that $f' = g$. Since g is infinitely differentiable, it follows that this f is also infinitely differentiable. Thus we have found an $f \in C^\infty(\mathbb{R})$ such that $T(f) = f' = g$.

(18) Define $T\colon F(-\infty,\infty) \to F(-\infty,\infty)$ by $T(f) = g$, where $g(x) = f(x) + f(-x)$. Show that T is linear, and identify $\mathrm{null}(T)$ and $\mathrm{range}(T)$ as certain special named types of functions. Make sure to prove your claims.

Solution 18:

Linearity:

$$
\begin{aligned}
T(cf + dg) &= (cf + dg)(x) + (cf + dg)(-x) && \text{(by def)} \\
&= c(f(x) + f(-x)) + d(g(x) + g(-x)) && \text{(arith)} \\
&= cT(f) + dT(g) && \text{(by def)}
\end{aligned}
$$

$T(f) = \mathbf{0}$ iff $f(x) + f(-x) = \mathbf{0}$ iff $f(-x) = -f(x)$. Thus $\mathrm{null}(T)$ is the set of odd functions.

I claim $\mathrm{range}(T)$ is the space of even functions. Indeed, given any $g = T(f) = f(x) + f(-x) \in \mathrm{range}(T)$, we have $g(-x) = f(-x) + f(x) = g(x)$. Thus $\mathrm{range}(T)$ is a subspace of the space of even functions. For the other direction, given even function $g(x)$, set $f(x) = \frac{1}{2}g(x)$. Then

$$
\begin{aligned}
T(f) &= \frac{1}{2}g(x) + \frac{1}{2}g(-x) \\
&= \frac{1}{2}g(x) + \frac{1}{2}g(x) && \text{(since } g \text{ is even)} \\
&= g(x)
\end{aligned}
$$

Thus the space of even functions is a subspace of $\mathrm{range}(T)$. We conclude the two spaces are equal.

3.4-3.5: LINEAR INDEPENDENCE AND BASES

(1) Let $V = \mathbb{R}^n$. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ be a set of r vectors. Prove (using some Gaussian elimination theory) that if $r > n$, then S is linearly dependent.

In other words: in $\mathbb{R}^n$ you can have *at most* n linearly independent vectors.

Solution 1:

Write

$$
\mathbf{v}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}
$$

Let $A = [a_{ij}]$ be the matrix obtained by placing $\mathbf{v}_j$ as the j-th column. This is a $n \times r$ matrix, $r > n$.

Now observe that

$$
c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r = \mathbf{0} \Leftrightarrow A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_r \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}
$$

$$
\Leftrightarrow (c_1, c_2, \ldots, c_r) \text{ is a solution to } A\mathbf{x} = \mathbf{0}
$$

But the linear system corresponding to $A\mathbf{x} = \mathbf{0}$ has n equations in r unknowns. Since $r > n$, there are more unknowns than equations. Gaussian elimination now tells us that

this system has a free variable, and hence infinitely many solutions. In particular, there is a nontrivial solution, which proves the vectors $\mathbf{v}_j$ are linearly dependent.

(2) Let $V = M_{22}$, and define $A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

    (a) Compute $W = \text{span}(\{A_1, A_2, A_3\})$, identifying it as a certain *familiar* set of matrices.

    (b) Decide whether $S = \{A_1, A_2, A_3\}$ is independent.

Solution 2:

(a) Using the definition of span, we compute

$$\text{span}\{A_1, A_2, A_3\} = \{c_1 A_1 + c_2 A_2 + c_3 A_3 \colon c_i \in \mathbb{R}\}$$
$$= \left\{ \begin{bmatrix} c_1 + c_3 & c_1 + c_2 + c_3 \\ c_1 + c_2 + c_3 & c_1 \end{bmatrix} : c_i \in \mathbb{R} \right\}.$$

We claim this last set is none other than the set of all symmetric matrices. It is clear that each element in the set above is symmetric, so it remains only to show that if A is symmetric, then A can be written in the above form for some choice of c_1, c_2, c_3.

    Write $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. We need to find c_1, c_2, c_3 such that

$$\begin{bmatrix} c_1 + c_3 & c_1 + c_2 + c_3 \\ c_1 + c_2 + c_3 & c_1 \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

The bottom-right entry tells us that we must pick $c_1 = c$. Then the top-left entry tells us that $c_3 = a - c$. Lastly, the off diagonal entries tells us that $c_2 = b - a$.

    Thus we can write $A = cA_1 + (b - a)A_2 + (a - c)A_3$, showing $A \in \text{span}\{A_1, A_2, A_3\}$, and proving that $\text{span}\{A_1, A_2, A_3\} = $ (set of symmetric matrices).

    (b) From the above description of a general linear combination $c_1 A_1 + c_2 A_2 + c_3 A_3$, it is clear that his will be the zero matrix iff $c_1 = c_2 = c_3 = 0$. Thus S is linearly independent.

(3) Let $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -2 \end{bmatrix}$. Find a basis for $W = \text{null}(A)$ and compute its dimension. You should first use GE to solve the defining matrix equation of $\text{null}(A)$.

Solution 3:

Using GE techniques we see that the set of solutions to $A\mathbf{x} = \mathbf{0}$ is given by

$$W = \{(\frac{t}{2} - s, s, -\frac{3t}{2}, t) \colon s, t \in \mathbb{R}\}$$

To find a basis for W we express this as a linear combination of two vectors:

$$W = \{(\frac{t}{2} - s, s, -\frac{3t}{2}, t) \colon s, t \in \mathbb{R}\}$$
$$= \{s(-1, 1, 0, 0) + t\left(\frac{1}{2}, 0, -\frac{3}{2}, 1\right) : s, t \in \mathbb{R}\}$$
$$= \text{span}\left\{(-1, 1, 0, 0), \left(\frac{1}{2}, 0, \frac{3}{2}, 1\right)\right\}$$

This shows that $\mathbf{v}_1 = (-1, 1, 0, 0)$ and $\mathbf{v}_2 = (1/2, 0, -3/2, 1)$ span $W = \text{null}(A)$. Since they are clearly linearly independent, they form a basis.

    Since $W = \text{null}(A)$ has a basis consisting of two vectors, we conclude $\dim(W) = 2$.

(4) Let $V = M_{nn}$. For each of the following subspaces $W \subset M_{nn}$, give a basis B of W and compute $\dim(W)$.
Justify your answer.

(a) W is the set of upper triangular $n \times n$ matrices.
(b) W is the set of symmetric $n \times n$ matrices.
(c) W is the set of skew-symmetric $n \times n$ matrices ($A^T = -A$).

Solution 4:

In what follows, E_{ij} is the matrix with a 1 in the ij-th entry, and 0's elsewhere. To come up with bases for the listed spaces, you should first write down what an arbitrary matrix of the given description looks like, then understand this description as a linear combination of some E_{ij}'s.

(a) An upper triangular matrix A can have arbitrary entries a_{ij} for all pairs i, j with $i \leq j$, and must have 0's elsewhere. That means we can write

$$A = \sum_{i \leq j} a_{ij} E_{ij}$$

as a linear combination *only* of the E_{ij} with $i \leq j$. This shows the set

$$B = \{E_{ij} : i \leq j\}$$

spans W. Since it is a subset of the standard basis for M_{nn}, B is automatically independent. Thus it is a basis.

We have $\dim(W) = \#B = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.

(b) An symmetric matrix A can have arbitrary diagonal entries a_{ii}, but once you pick an entry a_{ij} with $i < j$, then the corresponding entry a_{ji} is determined since $a_{ji} = a_{ij}$.
Let

$$S = \{E_{11}, \ldots, E_{nn}\} \cup \{D_{ij}\}_{1 \leq i < j \leq n}$$

where E_{ii} is exactly as above, and D_{ij} is the matrix with a 1 in the ij-th entry, a 1 in the ji-th entry, and 0's elsewhere. I claim S is a basis.

S spans. As noted above an arbitrary symmetric matrix A has any coefficients a_{ii} along the diagonal; if the ij-th entry is a_{ij}, then the ji-th entry is a_{ji}, and thus it is clear that

$$A = a_{11} E_{11} + \cdots a_{nn} E_{nn} + \sum_{1 \leq i < j \leq n} a_{ij} D_{ij}.$$

Thus S spans.

S is linearly independent. Suppose

$$a_{11} E_{11} + \cdots a_{nn} E_{nn} + \sum_{1 \leq i < j \leq n} a_{ij} D_{ij} = 0_{nn}.$$

From the definition of E_{ii} and D_{ij}, if follows that the ij-th entry of the matrix on the LHS is just a_{ij}. Thus $a_{ij} = 0$ for all i, j. This shows S is independent.

Finally, The number of elements in S is equal to the number of positions in an $n \times n$ matrix on and above the diagonal. Counting from the bottom of the matrix up, we see there are

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

of these. Thus $\dim(W) = \frac{n(n+1)}{2}$.

(c) Very similar to (b), except a skew-symmetric matrix A must $a_{ii} = 0$ for all i, and

$a_{ij} = -a_{ji}$ for all $i < j$. Setting S_{ij} to be the matrix with a 1 in the ij-th position, a -1 in the ji-th position and 0's elsewhere, we see that

$$B = \{S_{ij} : 1 \le i < j \le n\}$$

is a basis for W in this case. We count the basis in a similar manner to (b) to conclude

$$\dim(W) = 1 + 2 + \cdots + (n-1) = \frac{n(n-1)}{2}.$$

(5) Determine if the vectors are linearly independent or not in $\mathbb{R}^3$
   (a) $(-3, 0, 4), (5, -1, 2), (1, 1, 3)$
   (b) $(-2, 0, 1), (3, 2, 5), (6, -1, 1), (7, 0, -2)$

Solution 5:
The relevant system of linear equations has augmented matrix:

$$\begin{bmatrix} -3 & 5 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 4 & 2 & 3 & 0 \end{bmatrix} \xrightarrow{-1/3r_1} \begin{bmatrix} 1 & -5/3 & -1/3 & 0 \\ 0 & -1 & 1 & 0 \\ 4 & 2 & 3 & 0 \end{bmatrix}$$

$$\xrightarrow{4r_1-r_3} \begin{bmatrix} 1 & -5/3 & -1/3 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -26/3 & -13/3 & 0 \end{bmatrix}$$

$$\xrightarrow{-r_2} \begin{bmatrix} 1 & -5/3 & -1/3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -26/3 & -13/3 & 0 \end{bmatrix}$$

$$\xrightarrow{26/3r_2+r_3} \begin{bmatrix} 1 & -5/3 & -1/3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -13 & 0 \end{bmatrix}$$

Thus the only solution to the corresponding system is the trivial one, which means the vectors are linearly independent.
(b) Consider the matrix

$$\begin{bmatrix} -2 & 3 & 6 & 7 & 0 \\ 0 & 2 & -1 & 0 & 0 \\ 1 & 5 & 1 & -2 & 0 \end{bmatrix}$$

When you reduce it, you will have a free variable, which means we do not have just the trivial solution. Thus the vectors are linearly dependent.

(6) Determine if the vectors are linearly independent or not in $\mathbb{R}^4$
   (a) $(3, 8, 7, -3), (1, 5, 3, -1), (2, -1, 2, 6), (4, 2, 6, 4)$
   (b) $(3, 0, -3, 6), (0, 2, 3, 1), (0, -2, -2, 0), (-2, 1, 2, 1)$

Solution 6:
(a) Dependent. The relevant linear system has augmented matrix

$$\begin{bmatrix} 3 & 1 & 2 & 4 & 0 \\ 8 & 5 & -1 & 2 & 0 \\ 7 & 3 & 2 & 6 & 0 \\ -3 & -1 & 6 & 4 & 0 \end{bmatrix}$$

which row reduces to

$$\begin{bmatrix} 1 & 1/3 & 2/3 & 4/3 & 0 \\ 0 & 7/3 & -19/3 & -26/3 & 0 \\ 0 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Since the fourth column has no leading 1, the corresponding variable is free, which means the linear system has nontrivial solutions. This in turn implies the vectors are dependent.
(b) Independent. The relevant linear system has augmented matrix

$$\left[\begin{array}{cccc|c} 3 & 0 & 0 & -2 & 0 \\ 0 & 2 & -2 & 1 & 0 \\ -3 & 3 & -2 & 2 & 0 \\ 6 & 1 & 0 & 1 & 0 \end{array}\right]$$

which row reduces to

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & -2/3 & 0 \\ 0 & 1 & 0 & 5 & 0 \\ 0 & 0 & 1 & 15/2 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

Since the first four columns are all pivot columns, the corresponding linear system has no free variables, and hence only the trivial solution. This in turn implies the vectors are independent.

(7) Determine whether the vectors are linearly independent or are linearly dependent in P_2.
  (a) $2 - x + 4x^2$, $3 + 6x + 2x^2$, $2 + 10x - 4x^2$
  (b) $1 + 3x + 3x^2$, $x + 4x^2$, $5 + 6x + 3x^2$, $7 + 2x - x^2$

Solution 7:
Independent. The relevant system of equations arising from looking at linear combinations yielding the zero polynomial has augmented matrix:

$$\left[\begin{array}{ccc|c} 2 & 3 & 2 & 0 \\ -1 & 6 & 10 & 0 \\ 4 & 2 & -4 & 0 \end{array}\right]$$

which row reduces to

$$\left[\begin{array}{ccc|c} 1 & -6 & -10 & 0 \\ 0 & 1 & 22/15 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

Since the first three columns have leading 1's, the corresponding system has no free variables, and hence only the trivial solution. This implies the polynomials are independent.
(b) Dependent. The relevant system of equations arising from looking at linear combinations yielding the zero polynomial has augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 0 & 5 & 7 & 0 \\ 3 & 1 & 6 & 2 & 0 \\ 3 & 4 & 3 & -1 & 0 \end{array}\right]$$

Gaussian elimination shows the system has a free variable, which means there is a nontrivial linear combination yielding the zero polynomial, and hence that the polynomials are linearly dependent.

(8) Determine whether the matrices are linearly independent or dependent.

(a) In M_{22}

$$\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}; \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$$

(b) In M_{23}

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

Solution 8:

(a) Independent. Looking at linear combinations of the matrices equal to the zero matrix gives rise to the linear system with augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 1 & 1 & 0 \end{array}\right]$$

This row reduces to

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus $c_1 = c_2 = c_3 = 0$ is the only linear combination of the matrices yielding $\mathbf{0}$ and we can see that the matrices are independent.

(b) Independent. Consider

$$k_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + k_2 \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + k_3 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Since the first matrix is the only one to contribute to the zero in the a_{11} slot, k_1 must be zero. Since the second matrix is the only one to contribute to the zero in the a_{13} slot, k_2 must also be zero. Similarly, k_3 must also be zero. This means only the trivial solution exits, thus the 3 given matrices are independent.

(9) Determine all values of k for which the following matrices are linearly independent in M_{22}.

$$\begin{bmatrix} 1 & 0 \\ 1 & k \end{bmatrix}; \begin{bmatrix} -1 & 0 \\ k & 1 \end{bmatrix}; \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$$

Solution 9:

Looking at linear combinations of the matrices equal to the zero matrix gives rise to the linear system with augmented matrix

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & k & 1 & 0 \\ k & 1 & 3 & 0 \end{array}\right]$$

This row reduces to

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 0 & -k-1 & 2k-3 & 0 \\ 0 & 0 & 2k-4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The system has only the trivial solution iff the first three columns are pivot columns iff $k + 1 \neq 0$ and $2k - 4 \neq 0$ iff $k \neq -1$ and $k \neq 2$. Thus the matrices are independent for all values of k except for $k = -1$ and $k = 2$.

(10) Show that the following polynomials form a basis for P_3.

$$p_1 = 1 + x, p_2 = 1 - x, p_3 = 1 - x^2, p_4 = 1 - x^3$$

Solution 10:

To prove that the p_i form a basis, we must show that for any $p(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3$, we there is a *unique* choice of $c_1, c_2, \ldots, c_4$ satisfying $c_1 p_1 + c_2 p_2 + c_3 p_3 + c_4 p_4 = p(x)$. After combining like terms and equating like coefficients, this boils down to the matrix equation

$$A \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

having a unique solution for any choice of b_i, where

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

The invertibility theorem tells us that this is the case if and only if A is invertible if and only if $\det A \neq 0$. We compute

$$\det(A) = -1 \det \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -1 \left(-1 \det \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) = -1(-1(-1-1)) = -2 \neq 0$$

Thus A is invertible, and hence the p_i form a basis for P_3.

(11) Show that the following matrices form a basis for M_{22}.

$$A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}; \ A_2 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}; \ A_3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}; \ A_4 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Solution 11:

Reasoning as in Exercise 10, this boils down to the invertibility of the following matrix:

$$B = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & -1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

We test this using the determinant:

$$\det(B) = -1 \det \begin{bmatrix} 1 & 0 & 1 \\ -1 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} = -1 \left(1 \det \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix} \right) = -1(1(-1-0)) = 1 \neq 0$$

Thus B is not invertible, which implies the matrices A_i do not form a basis of M_{22}.

(12) Show that the following vectors do not form a basis for P_2

$$p_1 = 1 - 3x + 2x^2, p_2 = 1 + x + 4x^2, p_3 = 1 - 7x$$

Solution 12:

Reasoning as in Exercise 10, this amounts to showing

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 1 & -7 \\ 2 & 4 & 0 \end{bmatrix}$$

is invertible. Again, we can test this using the determinant:

$$\det(A) = 2\det\begin{bmatrix} 1 & 1 \\ 1 & -7 \end{bmatrix} - 4\det\begin{bmatrix} 1 & 1 \\ -3 & -7 \end{bmatrix} = 2(-7-1) - 4(-7+3) = -16 + 16 = 0$$

Thus A is not invertible, and hence the polynomials do not form a basis.

(13) Let V be the space spanned by $\mathbf{v}_1 = \cos^2 x$, $\mathbf{v}_2 = \sin^2 x$, $\mathbf{v}_3 = \cos 2x$.
(a) Show that $S = \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is not a basis for V.
(b) Find a basis for V.

Solution 13:
The trig identity $\cos^2 x - \sin^2 x = \cos(2x)$ provides a linear relation between the elements of S. Thus S is not linearly independent, and hence does not form a basis.
(b) Remove $\mathbf{v}_3$. The span remains the same and $\mathbf{v}_1$ and $\mathbf{v}_2$ are independent. So

$$S' = \{\cos^2 x, \sin^2 x\}$$

is a basis for V.

(14) Let $V = \mathbb{R}^n$, considered as column vectors. Prove: vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ form a basis of $\mathbb{R}^n$ if and only if the matrix $A = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & & \mathbf{v}_n \\ | & | & \cdots & | \end{bmatrix}$ is invertible.

Solution 14:

We prove the statement using a *chain of equivalences*:

$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis of $\mathbb{R}^n$

$\Leftrightarrow$ for all $\mathbf{b} \in \mathbb{R}^n$ we can write $\mathbf{b} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ in a *unique* way   (Theorem 3.15)

$\Leftrightarrow$ for all $\mathbf{b} \in \mathbb{R}^n$ the matrix equation $A\mathbf{c} = \mathbf{b}$ has a unique solution   (column method)

$\Leftrightarrow A$ is invertible   (invertibility theorem)

(15) Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be a basis for $\mathbb{R}^n$ (considered as column vectors), and let Q be an invertible $n \times n$ matrix.
Show that the vectors

$$\mathbf{w}_1 = Q\mathbf{v}_1$$
$$\mathbf{w}_2 = Q\mathbf{v}_2$$
$$\vdots$$
$$\mathbf{w}_n = Q\mathbf{v}_n$$

form a basis for $\mathbb{R}^n$.
Solution 15:

Let $A = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & & \mathbf{v}_n \\ | & | & \cdots & | \end{bmatrix}$.

Then A is invertible, by the previous exercise. Let $B = QA$. Using the column method of multiplication, we have

$$B = QA = \begin{bmatrix} | & | & \cdots & | \\ Q\mathbf{v}_1 & Q\mathbf{v}_2 & & Q\mathbf{v}_n \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{w}_1 & \mathbf{w}_2 & & \mathbf{w}_n \\ | & | & \cdots & | \end{bmatrix}$$

Since Q and A are invertible, so is B. The previous exercise then implies that the $\mathbf{w}_i$ form a basis.

(16) True or false. If true, provide a proof; if false, give an explicit counterexample.
   (a) If $V = \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V.
   (b) Every linearly independent subset of a vector space V is a basis for V.
   (c) If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V, then every vector in V can be expressed as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.
   (d) Every basis of P_4 contains at least one polynomial of degree 3 or less.
   (e) There is a basis for M_{22} consisting of invertible matrices.

Solution 16:
(a) False. A basis must also be linearly independent. Counter example: $V = R^2$ and the vectors $\{(1,0), (0,1), (1,1)\}$. Then $span\{(1,0), (0,1), (1,1)\} = R^2$ but since $(1,0) + (0,1) = (1,1)$ the vector set is not independent and therefore not a basis.

(b) False. The vector set must span V. Consider $V = R^3$ and the vector set $\{(1,0,0), (0,1,0)\}$. Clearly the set is independent, but it cannot span V.

(c) True. A basis spans V by definition.
(d) False. Consider the vector set

$$1 + x + x^2 + x^3 + x^4, x + x^2 + x^3 + x^4, x^2 + x^3 + x^4, x^3 + x^4, x^4$$

The degree of each polynomial is 4. To show this set forms a basis, we set up the usual linear combination equation $c_1 p_1 + c_2 p_2 + c_3 p_3 + c_4 p_4 = p$, where p is an arbitrary polynomial, which boils down to a matrix equation of the form

$$A \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \quad (A\mathbf{c} = \mathbf{a})$$

where

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

The set forms a basis if and only if this equation has a unique solution for each $\mathbf{a} = (a_0, a_1, a_2, a_3, a_4)$, if and only if A is invertible. Since $\det A = 1$ (lower triangular matrix), we conclude A is invertible, and hence the given polynomials form a basis.

(e) True. You can check for yourself that the following matrices are invertible and form a basis:

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

3.6: DIMENSION

(1) Prove the following statement from the "street smarts theorem".

Let V be a vector space with basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Then any collection of r vectors, with $r > n$, is linearly dependent.

Hint: suppose $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ with $r > n$. Begin by setting up the vector equation $c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_r\mathbf{w}_r = \mathbf{0}_V$; then write each $\mathbf{w}_j$ in this equation as a linear combination of the $\mathbf{v}_i$ (possible since $\mathbf{v}_i$ span); then collect like terms and use the fact that the $\mathbf{v}_i$ are linearly independent.

Solution 1:

Since B spans we can write each

$$\mathbf{w}_j = a_{1j}\mathbf{v}_1 + a_{2j}\mathbf{v}_2 + \cdots + a_{nj}\mathbf{v}_{nj} = \sum_{i=1}^{n} a_{ij}\mathbf{v}_i.$$

Now, we have

$$c_1\mathbf{w}_1 + c_2\mathbf{w}_2 \cdots + c_r\mathbf{w}_r = \mathbf{0}_V \Leftrightarrow c_1 \sum_{i=1}^{n} a_{i1}\mathbf{v}_i + c_2 \sum_{i=1}^{n} a_{i2}\mathbf{v}_i \cdots + c_r \sum_{i=1}^{n} a_{ir}\mathbf{v}_i = \mathbf{0}$$

$$\Leftrightarrow (\sum_{j=1}^{r} a_{1j}c_j)\mathbf{v}_1 + (\sum_{j=1}^{r} a_{2j}c_j)\mathbf{v}_2 + \cdots + (\sum_{j=1}^{r} a_{nj}c_j)\mathbf{v}_n = \mathbf{0} \quad \text{(grouping like terms)}$$

$$\Leftrightarrow (a_{i1}c_1 + a_{i2}c_2 + \cdots a_{ir}c_r) = 0 \text{ for all } i, \qquad (\mathbf{v}_i \text{ are lin. ind.})$$

$$\Leftrightarrow (c_1, \ldots, c_r) \text{ is a solution to system}$$

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1r}x_r = 0$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2r}x_r = 0$$
$$\vdots \qquad \vdots \qquad \vdots$$
$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nr}x_r = 0$$

Now this is a homogenous linear system of n equations in r unknowns, where $r > n$. Gaussian elimination theory tells us that this has a nontrivial solution. It follows that the $\mathbf{w}_i$ are linearly dependent.

(2) For each of the following choices of a vector space V and subset $S \subseteq V$ decide whether S is a basis for V.

Keep an eye out for potential shortcuts using the dimension theorem compendium.

(a) $V = P_3$, $S = \{1 - x^3, x - x^3, x^2 - x^3\}$.
(b) $V = \mathbb{R}^3$, $S = \{(1, 2, 3), (2, 5, 6), (1, -1, 3)\}$.

Solution 2:

(a) The set S is guaranteed not to span P_3. Indeed, we have $\dim P_3 = 4$, $\#S = 3$. Street smarts! (Dimension theorem compendium.)

(b) Since $\#S = \dim \mathbb{R}^3 = 3$, the dimension theorem compendium asserts that S is a basis

iff S is linearly independent iff S spans. This means I need only check one of the two usual conditions to being a basis: linear independence, or span.

Let's check whether S is linearly independent. We have

$$c_1(1,2,3) + c_2(2,5,6) + c_3(1,-1,3) = (0,0,0) \iff \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & -1 \\ 3 & 6 & 3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & -1 \\ 3 & 6 & 3 \end{bmatrix}$. We see via simple computation that $\det A = 0$ and thus that A is not invertible. The invertibility theorem then implies that there is a nontrivial solution $(c_1, c_2, c_3) \neq (0,0,0)$ to the matrix equation, and hence a nontrivial linear combination of the vectors yielding $\mathbf{0}$. This proves S is not linearly independent, and hence not a basis.

(3) In multivariable calculus, a plane in $\mathbb{R}^3$ is defined as the set of solutions to an equation of the form $ax + by + cz = d$, where at least one of a, b, c is nonzero.

  In particular, a plane passing through the origin $(0,0,0)$ is the set of solutions to an equation of the form $ax + by + cz = 0$, where at least one of a, b, c is nonzero.

  (a) Show that a plane passing through the origin is a subspace of $\mathbb{R}^3$ of dimension 2; conversely, show that a subspace of $\mathbb{R}^3$ of dimension 2 is a plane passing through the origin.

    In other words, show that the set of all planes passing through the origin is precisely the set of all 2-dimensional subspaces of $\mathbb{R}^3$.

  (b) Let $\mathcal{P}\colon ax + by + cz + d$ be any plane in $\mathbb{R}^3$, and let $P = (x_0, y_0, z_0)$ be a point of $\mathcal{P}$. Show that there is a 2-dimensional subspace $W \subseteq \mathbb{R}^3$ such that $\mathcal{P} = P + W := \{(x_0, y_0, z_0) + \mathbf{w}\colon \mathbf{w} \in W\}$.

    The set $P + W$ is called the **translate of W by P**. This result shows that all planes in $\mathbb{R}^3$ are translates of 2-dimensional subspaces.

Solution 3:

(a) Let $\mathcal{P}$ be the plane defined by $ax + by + cz = 0$. Gaussian elimination allows us to describe the points of $\mathcal{P}$ parametrically. For example, if $a \neq 0$, then $\mathcal{P} = \{(-(bs + ct)/a, s, t)\colon s, t\} = \text{span}\{(-b/a, 1, 0), (-c/a, 0, 1)\}$. The set $\{\{(-b/a, 1, 0), (-c/a, 0, 1)\}$ is clearly linearly independent: the vectors are not scalar multiples of one another. Thus $\mathcal{P}$ is a subspace of dimension 2 in this case. The cases $b \neq 0$ and $c \neq 0$ are dealt with similarly. In all cases, we see that $\mathcal{P}$ is a subspace of dimension 2.

  Going the other way, suppose $W \subseteq \mathbb{R}^3$ is a 2-dimensional subspace. By definition of dimension, W has a a basis $\{(x_1, y_1, z_1), (x_2, y_2, z_2)\}$ consisting of two linearly independent vectors. I claim that there is a nonzero vector (a, b, c) for which

$$ax_1 + by_1 + cz_1 = ax_2 + by_2 + cz_3 = 0.$$

To see this, consider the corresponding matrix equation

$$\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thinking of a, b, c as the unknowns here, we see that we have a homogenous linear system with more unknowns than equations. It follows that there is a nontrivial solution $(a, b, c) \neq (0,0,0)$ to this equation.

  Fix one such solution (a, b, c), and let $\mathcal{P}$ be the plane with defining equation $ax + by + cz = 0$. We showed above that $\mathcal{P}$ is a 2-dimensional subspace. Since (x_1, y_1, z_1) and (x_2, y_2, z_2)

both satisfy this equation we must have $W = \operatorname{span}\{(x_1, y_1, z_1), (x_2, y_2, z_2)\} \subseteq \mathcal{P}$. Lastly, since $\dim W = \dim \mathcal{P} = 2$, we conclude $W = \mathcal{P}$ (dimension theorem compendium).

(b) This claim follows from part (a) and the fact that an arbitrary plane $\mathcal{P}\colon ax+by+cz = d$ can be obtained by taking the parallel plane $\mathcal{P}_0\colon ax + by + cz = 0$ that passes through the origin, and translating it by $P = (x_0, y_0, z_0)$, where P is any point in $\mathcal{P}$.

(4) Let $V = M_{33}$, $W = \{A \in M_{33}\colon A^T = -A\}$ (i.e., the subspace of skew-symmetric matrices), and

$$W' = \operatorname{span}\left\{ A_1 = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \right\} \text{ Show that}$$

$W' = W$ as follows:
(a) Show that $W' \subseteq W$ (easy);
(b) Compute the dimensions of W' and W and use the dimension theorem (compendium).

Solution 4:
(a) Since each A_i is clearly skew-symmetric, then any linear combination of the A_i's is skew-symmetric. Thus $W' = \operatorname{span}(\{A_1, A_2, A_3\}) \subseteq W$.

(b) Since $W = \left\{ \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix} : a, b, c \in \mathbb{R} \right\}$, we see essentially by inspection that

$$B = \left\{ \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \right\}$$

is basis for W, and hence that $\dim W = 3$.

Since $W' \subseteq W$, it follows that $W = W'$ iff $\dim W' = 3$, by the dimension theorem (compendium). To show $\dim W' = 3$, we need only show that the given matrices, which span W' by definition, are linearly independent and hence a basis. This is done in the usual way. I leave it to you.

(5) Find a basis for the solution space of the homogeneous linear system, and find the dimension of that space.

$$
\begin{aligned}
x_1 - 4x_2 + 3x_3 - x_4 &= 0 \\
2x_1 - 8x_2 + 6x_3 - 2x_4 &= 0
\end{aligned}
$$

Solution 5:

Let W be the set of solutions to this homogenous system of equations. The augmented matrix we obtain to determine W is

$$\begin{bmatrix} 1 & -4 & 3 & -1 & 0 \\ 2 & -8 & 6 & -2 & 0 \end{bmatrix}$$

Since row two is just 2 times row one, the reduced matrix is:

$$\begin{bmatrix} 1 & -4 & 3 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Writing the solution with free variables $x_2 = t, x_3 = r, x_4 = s$

$$x_4 = s, x_3 = r, x_2 = t, x_1 = 4t - 3s + r$$

79

Writing the solution in vector form gives:

(∗) $$(x_1, x_2, x_3, x_4) = (4t - 3s + r, t, s, r)$$
(∗∗) $$= t(4, 1, 0, 0) + s(-3, 0, 1, 0) + r(1, 0, 0, 1)$$

Thus the vectors $\{(4, 1, 0, 0), (-3, 0, 1, 0), (1, 0, 0, 1)\}$ span the vector space. For linear independence, note that

$$t(4, 1, 0, 0) + s(-3, 0, 1, 0) + r(1, 0, 0, 1) = (0, 0, 0, 0) \Rightarrow (4t - 3s + r, t, s, r) = (0, 0, 0, 0) \Rightarrow r = s = t = 0.$$

This shows the three vectors form a basis of W, and hence that $\dim W = 3$.

(6) Find a basis for the solution space of the homogeneous linear system, and find the dimension of that space.

$$\begin{aligned} x + y + z &= 0 \\ 3x + 2y - 2z &= 0 \\ 4x + 3y - z &= 0 \\ 6x + 5y + z &= 0 \end{aligned}$$

Solution 6:

The matrix we obtain to find the basis is:

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

Which reduces to the matrix:

$$\begin{bmatrix} 1 & 0 & -4 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Writing the solution with free variable $x_3 = r$

$$x_3 = r, x_2 = -5r, x_1 = 4r$$

Writing the solution in vector form gives:

$$\begin{aligned} (x_1, x_2, x_3) &= (4r, -5, r) \\ &= r(4, -5, 1) \end{aligned}$$

Thus the vector $\{(4, -5, 1)\}$ spans the vector space, and by the remark at the end of example 3 on page 223, they also form a basis for the vector space. Since the basis has 1 vector, the dimension of the vector space is 1.

(7) Let
$$S = \{\mathbf{v}_1 = (1, 1, 1, 1), \mathbf{v}_2 = (2, 2, 2, 0)\} \subseteq \mathbb{R}^4.$$
Enlarge S into a basis of $\mathbb{R}^4$.
Solution 7:

We will have a systematic way of doing this later. For now we proceed in the spirit of the proof that every linearly independent set can be extended to a basis.

Namely, we search for a vector $\mathbf{v} \in \mathbb{R}^4$ such that $\mathbf{v} \notin \operatorname{span} S$. Consider $\mathbf{v}_3 = (1, 0, 0, 0)$. A simple application of Gaussian elimination shows that $(1, 0, 0, 0)$ is not a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. Thus $S' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

Next, I claim $\mathbf{v}_4 = (0, 1, 0, 0)$ is not in $\operatorname{span} S'$. Again you can confirm this with a Gaussian elimination argument. We conclude that

$$S'' = \{(1, 1, 1, 1), (2, 2, 2, 0), (1, 0, 0, 0), (0, 1, 0, 0)\}$$

is linearly independent. Since $\#S'' = \dim \mathbb{R}^4 = 4$, we conclude S'' is a basis for $\mathbb{R}^4$ extending our original set S.

(8) Compute bases and dimensions for the following subspaces of $\mathbb{R}^4$.
   (a) $W = \{(a, b, c, d) \in \mathbb{R}^4 : d = a + b, c = a - b\}$
   (b) $W = \{(a, b, c, d) \in \mathbb{R}^4 : a + b = c + d\}$

Solution 8:
Both spaces can be described as null spaces of a linear transformation.
   In (a) we have

$$W = \{(a, b, c, d) \in \mathbb{R}^4 : d = a + b, c = a - b\} = \{(a, b, c, d) \in \mathbb{R}^4 : a + b - d = 0 \text{ and } a - b - c = 0\}.$$

Thus $W = \operatorname{null} \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & -1 & -1 & 0 \end{bmatrix}$. After some Gaussian elimination, we see that $W = \{(\frac{1}{2}(s + t), \frac{1}{2}(-s + t), s, t) : s, t \in \mathbb{R}\}$. From this parametric description, we can then easily extract the basis $\{(1/2, -1/2, 1, 0), (1/2, 1/2, 0, 1)\}$. We conclude $\dim W = 2$.

In (b) we have $W = \{(a, b, c, d) \in \mathbb{R}^4 : a + b - c - d = 0\} = \operatorname{null}[1 \; 1 \; -1 \; -1]$. Again Gaussian elimination tells us that

$$W = \{(-r + s + t, r, s, t) : r, s, t \in \mathbb{R}\} = \operatorname{span} \underbrace{\{(-1, 1, 0, 0), (1, 0, 1, 0), (1, 0, 0, 1)\}}_{S} = \operatorname{span} S.$$

The spanning set S on the right is easily seen to be linearly independent: $a(-1, 1, 0, 0) + b(1, 0, 1, 0) + c(1, 0, 0, 1) = (-a + b + c, a, b, c) = (0, 0, 0, 0) \Rightarrow a = b = c = 0$. Thus S is a basis for W and we conclude $\dim W = 3$.

(9) Let $V = M_{nn}$, W_1 the set of diagonal matrices, W_2 the set of symmetric matrices, W_3 the set of upper triangular matrices. Compute the dimensions of all these spaces. You may use your results from Exercise 4 in Section 3.3-3.4.
Solution 9:

   (i) W_1, the set of diagonal matrices. Let

$$S = \{E_{11}, E_{22}, \ldots, E_{nn}\},$$

   where as usual E_{ij} denotes the matrix with a 1 in the ij-th entry, and 0's elsewhere. I claim S is a basis of W_1, and hence that $\boxed{\dim(W_1) = n}$.
   S spans. An arbitrary diagonal matrix A has entries $a_i i$ along the diagonal and 0 elsewhere. Then clearly

$$A = a_{11} E_1 1 + \cdots + a_{nn} E_n n.$$

   This shows S spans W_1.
   S is linearly independent. Note that S is a subset of the standard basis for M_{nn}. Since the standard basis is linearly independent, any subset of it is also linearly independent. Thus S is independent.

(ii) W_2, the set of symmetric matrices.

Let
$$S = \{E_{11}, \ldots, E_{nn}\} \cup \{F_{ij}\}_{1 \leq i < j \leq n}$$

where E_{ii} is exactly as above, and F_{ij} is the matrix with a 1 in the ij-th entry, a 1 in the ji-th entry, and 0's elsewhere. I claim S is a basis.

S spans. Similar argument as above. An arbitrary symmetric matrix A has any coefficients a_{ii} along the diagonal; if the ij-th entry is a_{ij}, then the ji-th entry is a_{ji}, and thus it is clear that

$$A = a_{11}E_{11} + \cdots a_{nn}E_{nn} + \sum_{1 \leq i < j \leq n} a_{ij}F_{ij}.$$

Thus S spans.

S is linearly independent. Suppose

$$a_{11}E_{11} + \cdots a_{nn}E_{nn} + \sum_{1 \leq i < j \leq n} a_{ij}F_{ij} = 0_{nn}.$$

From the definition of E_{ii} and F_{ij}, if follows that the ij-th entry of the matrix on the LHS is just a_{ij}. Thus $a_{ij} = 0$ for all i, j. This shows S is independent.

Finally, The number of elements in S is equal to the number of positions in an $n \times n$ matrix on and above the diagonal. Counting from the bottom of the matrix up, we see there are

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

of these. Thus $\boxed{\dim(W_2) = \dfrac{n(n+1)}{2}}$.

(iii) W_3, the set of upper triangular matrices.

The argument is very similar to the above arguments. The set

$$S = \{E_{ij}\}_{1 \leq i \leq j \leq n}$$

is a basis. This set also has $\frac{n(n+1)}{2}$ elements. Thus $\boxed{\dim(W_3) = \dfrac{n(n+1)}{2}}$.

(10) Let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a basis for a vector space V. Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is also a basis, where $\mathbf{u}_1 = \mathbf{v}_1$, $\mathbf{u}_2 = \mathbf{v}_1 + \mathbf{v}_2$, and $\mathbf{u}_3 = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3$.

Solution 10:

Since $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for V, any other basis for V must also have exactly 3 vectors. Since the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ also has 3 vectors, then by the dimension theorem (compendium), we only need to show that this set spans V or that it is linearly independent. Let's show that it is linearly independent. If we can show that if $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3 = 0$, then $c_1 = c_2 = c_3 = 0$, we will be done.

$$
\begin{aligned}
c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3 = 0 &\Rightarrow c_1\mathbf{v}_1 + c_2(\mathbf{v}_1 + \mathbf{v}_2) + c_3(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3) = 0 \\
&\Rightarrow (c_1 + c_2 + c_3)\mathbf{v}_1 + (c_2 + c_3)\mathbf{v}_2 + c_3\mathbf{v}_3 = 0 \\
&\Rightarrow (c_1 + c_2 + c_3) = (c_2 + c_3) = c_3 = 0 \qquad \text{(since } \mathbf{v}_i \text{ are independent)} \\
&\Rightarrow c_3 = c_2 = c_1 = 0 \qquad\qquad\qquad\quad \text{(work backwards)}
\end{aligned}
$$

This shows the $\mathbf{u}_i$ are independent, and hence constitute a basis.

(11) Let $V = M_{23}$ and define
$$W = \{A \in M_{23}: \text{ rows and columns of } A \text{ sum to } 0\}.$$
Find a basis for W and compute its dimension.
Solution 11:

We can proceed by inspection. The general element of W looks like
$$A = \begin{bmatrix} r & s & -(r+s) \\ -r & -s & (r+s) \end{bmatrix} = r \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} + s \begin{bmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$
From this description it is obvious that
$$B = \left\{ \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \right\}.$$
spans W. It is also fairly clear that the given matrices are linearly independent (look at the relevant matrix equation yourself). Thus B is a basis.
Since B has 2 elements, we have $\dim(W) = 2$.

(12) Let $C^\infty(-\infty, \infty)$ and consider the subspace
$$W = \text{span}(\{\cos(2x), \sin(2x), \sin(x)\cos(x), \cos^2(x), \sin^2(x)\}).$$
Compute the dimension of W by providing an explicit basis.
Solution 12:

Some trig identities:
$$\cos(2x) = \cos^2(x) - \sin^2(x)$$
$$\sin(2x) = 2\sin(x)\cos(x).$$
These identities show us that $\cos(2x)$ and $\sin(2x)$ are in the span of $\{\cos^2(x), \sin^2(x), \sin(x)\cos(x)\}$, and thus that $W = \text{span}(\{\cos^2(x), \sin^2(x), \sin(x)\cos(x)\}$. Is this new set linearly independent?
Suppose we have
$$k_1 \cos^2(x) + k_2 \sin^2(x) + k_3 \sin(x)\cos(x) = 0$$
for all x.
Evaluating at $x = 0$ yields $k_1 = 0$.
Evaluating at $x = \pi/2$ yields $k_2 = 0$.
Thus the identity becomes $k_3 \sin(x)\cos(x) = 0$, which is true for all x if and only if $k_3 = 0$.
This shows that $\{\cos^2(x), \sin^2(x), \sin(x)\cos(x)\}$ is linearly independent. Since it also spans W, it is a basis of W. Since W has a basis consisting of 3 elements, we conclude $\dim(W) = 3$.

(13) Let $A \in M_{nn}$. Show that there is a nonzero polynomial $p(x) = a_{n^2}x^{n^2} + a_{n^2-1}x^{n^2-1} + \cdots + a_1 x + a_0$ such that
$$p(A) = a_{n^2}A^{n^2} + a_{n^2-1}A^{n^2-1} + \cdots + a_1 A + a_0 I_n = \underset{n \times n}{\mathbf{0}}.$$
Hint: consider the set $S = \{I_n, A, A^2, \ldots, A^{n^2}\} \subseteq M_{nn}$ and use dimension theory.
Solution 13:

We are tempted to quote the dimension theorem compendium: $\dim M_{nn} = n^2$, $\#S = n^2 + 1$, thus S is linearly dependent. This would then imply there is a linear combination

of the elements of S equal to 0: i.e., there is a nontrivial choice of a_i such that $a_{n^2}A^{n^2} + a_{n^2-1}A^{n^2-1} + \cdots + a_1 A + a_0 I_n = \underset{n \times n}{\mathbf{0}}$, as desired.

The problem with this reasoning is that we cannot be sure that $\#S = n + 1$!! In other words we might have $A^i = A^k$ for some pair $0 \le i < j \le n^2$. But if this is the case, then we have $A^j - A^i = \underset{n \times n}{\mathbf{0}}$, and thus $p(A) = \underset{n \times n}{\mathbf{0}}$, where $p(x) = x^j - x^i$.

(14) Prove that $V = C^\infty(\mathbb{R})$ is infinite dimensional.

Hint: use a proof by contradiction. More specifically, assume $\dim V = n$ for some n, and try and derive a contradiction to one of our dimension theorem compendium statements.

Solution 14:

By contradiction. Assume $\dim V = n$ for some $n < \infty$. Consider the set $S = \{f_0, f_1, f_2, \ldots, f_n\} \subseteq C^\infty(\mathbb{R})$ where $f_j(x) = x^j$. The set S contains $n + 1$ elements, and is linearly independent: indeed it is the standard basis of P_n. This contradicts the dimension theorem compendium: a set containing more than n elements in an n-dimensional space is guaranteed to be linearly dependent! Thus we cannot have $\dim V = n$ for any n.

4.1-4.2: LINEAR TRANSFORMATIONS, BASES, RANK-NULLITY

(1) Suppose $T \colon \mathbb{R}^2 \to \mathbb{R}^3$ is linear, and satisfies $T(1,1) = (1,1,1)$, $T(1,-1) = (1,2,1)$.

Find the matrix A such that $T = T_A$.

Hint: first write $(1,0)$ and $(0,1)$ as linear combinations of $(1,1)$ and $(1,-1)$.

Solution 1:

Write $(1,0) = \frac{1}{2}(1,1) + \frac{1}{2}(1,-1)$, and $(0,1) = \frac{1}{2}(1,1) - \frac{1}{2}(1,-1)$. Then

$$
\begin{aligned}
T(1,0) &= T(\tfrac{1}{2}(1,1) + \tfrac{1}{2}(1,-1)) \\
&= \tfrac{1}{2}T(1,1) + \tfrac{1}{2}T(1,-1) && (T \text{ is linear}) \\
&= \tfrac{1}{2}(1,1,1) + \tfrac{1}{2}(1,2,1) && (\text{given}) \\
&= (1,3/2,1)
\end{aligned}
$$

and

$$
\begin{aligned}
T(0,1) &= T(\tfrac{1}{2}(1,1) - \tfrac{1}{2}(1,-1)) \\
&= \tfrac{1}{2}T(1,1) - \tfrac{1}{2}T(1,-1) && (T \text{ is linear}) \\
&= \tfrac{1}{2}(1,1,1) - \tfrac{1}{2}(1,2,1) && (\text{given}) \\
&= (0,-1/2,0).
\end{aligned}
$$

It follows, following our recipe, that $T = T_A$, where

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

(2) Suppose $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is linear, that $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for $\mathbb{R}^n$ and that $T(\mathbf{v}_i) = 2\mathbf{v}_i$ for all $1 \le i \le n$.

Compute the $n \times n$ matrix A such that $T = T_A$. Justify your answer.

Solution 2:

I claim $T = T_A$, where

$$A = 2I_n = \begin{bmatrix} 2 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 2 \end{bmatrix}.$$

Indeed, observe that T_A is linear (since it's a matrix equation), and that

$$T_A(\mathbf{x}) = 2I_n \underset{n \times 1}{\mathbf{x}} = 2\mathbf{x}$$

for any $\mathbf{x} \in \mathbb{R}^n$. In particular, $T_A(\mathbf{v}_i) = 2\mathbf{v}_i$ for the elements $\mathbf{v}_i$ of the given basis B. But then $T(\mathbf{v}_i) = T_A(\mathbf{v}_i)$ for all $\mathbf{v}_i \in B$. We have a theorem that states any two linear transformations that agree on a basis must agree everywhere, and hence are equal. We conclude $T = T_A$, as claimed.

(3) Use the rank-nullity theorem to compute the rank of the linear transformation T described.
  (a) $T\colon \mathbb{R}^7 \to M_{32}$ has nullity 2.
  (b) $T\colon P_3 \to \mathbb{R}$ has nullity 1.
  (c) The null space of $T\colon P_5 \to P_5$ is P_4.
  (d) $T\colon P_n \to M_{mn}$ has nullity 3.

Solution 3:
(a) $\operatorname{rank}(T) = \dim(R^7) - 2 = 7 - 2 = 5$
(b) $\operatorname{rank}(T) = \dim(P_3) - 1 = 4 - 1 = 3$
(c) $\operatorname{rank}(T) = \dim(P_5) - \operatorname{nullity}(T) = \dim(P_5) - \dim(P_4) = 1$
(d) $\operatorname{rank}(T) = \dim(P_n) - 3 = (n + 1) - 3 = n - 2$

(4) Give an explicit example of a matrix A such that $\operatorname{col}(A) \ne \operatorname{col}(U)$.

Solution 4:

Take $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, which row reduces to $U = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$.

Then $\operatorname{col}(A) = \operatorname{span}(\{(1, 1)\})$, which is the line $y = x$, and $\operatorname{col}(U) = \operatorname{span}(\{(1, 0)\})$, which is the x-axis. Clearly these two lines are not equal.

Note: any explicit counterexample you give must be non-invertible, since otherwise $\operatorname{col} A = \operatorname{col} U = \mathbb{R}^n$, by the invertibility theorem.

(5) Prove that $\operatorname{col}(A) = \{\mathbf{b} \in \mathbb{R}^m \colon A\mathbf{x} = \mathbf{b} \text{ has a solution}\}$.

Solution 5:

A complete proof of this is given in my lecture notes.

(6) Find bases and dimensions for all fundamental spaces of

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

Solution 6:

Routine. Follow the usual algorithms. The matrix row reduces to

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

From this we compute the following bases:

$\mathrm{null}(A)\colon B_1 = \{(1,1,0,-2),(-3,1,2,0)\}$ $\qquad\qquad$ $\dim(\mathrm{null}(A)) = 2$

$\mathrm{col}(A)\colon B_2 = \{(1,2,1),(-1,0,1)\}$ $\qquad\qquad$ $\dim(\mathrm{col}(A)) = 2$

$\mathrm{row}(A)\colon B_3 = \{(1,-1,2,0),(0,1,-\frac{1}{2},\frac{1}{2})$ $\qquad\qquad$ $\dim(RS(A)) = 2$

(7) For each matrix below use the rank-nullity theorem to help determine the fundamental spaces "by inspection"–i.e., without having to do Gaussian elimination.
   The various fundamental spaces will live either in $\mathbb{R}^2$ or $\mathbb{R}^3$. For each matrix give two sketches showing (a) $\mathrm{null}(A)$ and $\mathrm{row}(A)$ in one sketch, and (b) $\mathrm{col}(A)$ in another.

(a) $A = \begin{bmatrix} 3 & 2 & 1 \\ -6 & -4 & -2 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & 1 \\ 8 & 4 \\ 6 & 3 \end{bmatrix}$

Solution 7:
(a) The columns of A are all scalar multiples of each other, so clearly $B_1 = \{(1,-2)\}$ is a basis of $\mathrm{col}(A)$, and we have $\mathrm{rank}(A) = 1$. When you sketch $\mathrm{col}(A)$ you get the line in $\mathbb{R}^2$ connecting $(0,0)$ and $(1,-2)$.
   It follows that $\mathrm{row}(A)$ is also 1-dimensional, and thus that any nonzero row of A will do as a basis: e.g., $B_2 = \{(3,2,1)\}$. When you sketch $\mathrm{row}(A)$ you get the line in $\mathbb{R}^3$ passing through $(0,0,0)$ and $(3,2,1)$.
   Lastly rank-nullity implies $\dim(\mathrm{null}(A)) = 3 - 1 = 2$. Furthermore, by the orthogonality relation $\mathrm{null}(A)$ is the plane orthogonal to $\mathrm{row}(A)$. Having already sketched $\mathrm{row}(A)$ above, this is now easy to add to your sketch. You do it!
(b) The reasoning is similar to the above.
   A basis for $\mathrm{col}(A)$ is $B_1 = \{(1,4,3)\}$. Thus $\mathrm{col}(A)$ is the span of this vector, which defines a line in $\mathbb{R}^3$.
   A basis for $\mathrm{row}(A)$ is $B_2 = (2,1)$, and thus $\mathrm{row}(A)$ is the line passing through the origin and $(2,1)$. Then $\mathrm{null}(A)$ is just the orthogonal complement of this in $\mathbb{R}^2$, which now is the line perpendicular to this line. Sketch it!

(8) Let A be $m \times n$ with $n < m$. Use fundamental spaces to show that there is a $\mathbf{b} \in \mathbb{R}^m$ such that the system $A\mathbf{x} = \mathbf{b}$ is inconsistent.
Solution 8:

   Using the fact that $\mathrm{col}(A) = \{\mathbf{b} \in \mathbb{R}^m\colon A\mathbf{x} = \mathbf{b}$ has a solution$\}$, we see that we need only show that $\mathrm{col}(A) \subsetneq \mathbb{R}^m$. We do so using rank-nullity.
   We have $\mathrm{rank}(A) \leq \min\{m,n\}$. Since $n < m$, $\min\{m,n\} = n$. Thus $\mathrm{rank}(A) \leq n$.
   But $\mathrm{rank}(A) = \dim(\mathrm{col}(A))$. Thus

$$\dim(\mathrm{col}(A)) \leq n < m = \dim(\mathbb{R}^m).$$

Thus we have $\mathrm{col}(A) \subset \mathbb{R}^m$ and $\dim(\mathrm{col}(A)) < \dim(\mathbb{R}^m)$. By our theorem about dimensions of subspaces, it follows that $\mathrm{col}(A) \neq \mathbb{R}^m$, and we are done.

(9) Let $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. Use the rank-nullity theorem to find bases for all the fundamental spaces of A "by inspection".

Solution 9:

More of the same. There are clearly only 2 linearly independent columns. So a basis for $\text{col}(A)$ is $B_1 = \{(1, 2, 1), (1, 1, 1)\}$.

Since $\dim(\text{row}(A)) = \dim(\text{col}(A)) = 2$, we see easily that $B_2 = \{(1, 1, 1, 1), (2, 1, 2, 1)\}$ is a basis for $\text{row}(A)$.

Lastly, rank-nullity tells us that $\dim(\text{null}(A)) = n - 2 = 4 - 2 = 2$. Thus to come up with a basis for $\text{null}(A)$ I need only find two linearly independent elements of $\text{null}(A)$.

I can come up with two such solutions to $A\mathbf{x} = \mathbf{0}$ easily by inspection, thinking of $A\mathbf{x}$ as a linear combination of the columns of A. For example, since the first and third columns of A are identical, I see that $(1, 0, -1, 0)$ is an element of $\text{null}(A)$. (Do the multiplication yourself!) Since the second and fourth columns are identical, I see that $(0, 1, 0, -1)$ is an element of $\text{null}(A)$.

Since these two vectors are clearly linearly independent and $\text{null}(A)$ has dimension 2, we see that $B_3 = \{(1, 0, -1, 0), (0, 1, 0, -1)\}$ is a basis for $\text{null}(A)$.

(10) Find bases for all the fundamental spaces of the following matrices.

(a)
$$A = \begin{bmatrix} 1 & 4 & 5 & 2 \\ 2 & 1 & 3 & 0 \\ -1 & 3 & 2 & 2 \end{bmatrix}$$

(b)
$$A = \begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 3 & -2 & 1 & 4 & -1 \\ -1 & 0 & -1 & -2 & -1 \\ 2 & 3 & 5 & 7 & 8 \end{bmatrix}$$

Solution 10:

(a) The matrix A reduces to:
$$\begin{bmatrix} 1 & 0 & 1 & -2/7 \\ 0 & 1 & 1 & 4/7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The non-zero rows of the reduced matrix form a basis for the row space:
$$\left\{ \begin{bmatrix} 1 & 0 & 1 & -2/7 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 & 4/7 \end{bmatrix} \right\}$$

Since the null space is the solution space of the homogeneous linear system $A\mathbf{x} = \mathbf{0}$, we need to look at the first two rows of the reduced matrix.

$$\begin{aligned} x_1 + x_3 - 2/7x_4 &= 0 \\ x_2 + x_3 + 4/7x_4 &= 0 \end{aligned}$$

Setting the free variables $x_3 = r$ and $x_4 = s$.

$$\begin{aligned} (x_1, x_2, x_3, x_4) &= (-r + 2/7s, -r - 4/7s, r, s) \\ &= r(-1, -1, 1, 0) + s(2/7, -4/7, 0, 1) \end{aligned}$$

Thus the basis for the null space is:

$$\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2/7 \\ -4/7 \\ 0 \\ 1 \end{bmatrix} \right\}$$

(b) The matrix A reduces to:

$$\begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus the basis for the row space is:

$$\left\{ \begin{bmatrix} 1 & 0 & 1 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 1 & 1 & 2 \end{bmatrix} \right\}$$

Setting up the equations for the null space basis:

$$x_1 + x_3 + 2x_4 + x_5 = 0$$
$$x_2 + x_3 + x_4 + 2x_5 = 0$$

Setting the free variables $x_3 = r$, $x_4 = s$, and $x_5 = t$.

$$\begin{aligned} (x_1, x_2, x_3, x_4, x_5,) &= (-r - 2s - t, -r - s - 2t, r, s, t) \\ &= r(-1, -1, 1, 0, 0) + s(-2, -1, 0, 1, 0) + t(-1, -2, 0, 0, 1) \end{aligned}$$

Thus the basis for the null space is:

$$\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

(11) Find bases for the row space and for the column space of the given matrix.
(a)
$$A = \begin{bmatrix} 1 & 2 & 4 & 5 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

(b)
$$A = \begin{bmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Solution 11:
(a) The non-zero rows of the matrix for a basis for the row space:

$$\left\{ \begin{bmatrix} 1 & 2 & 4 & 5 \end{bmatrix}, \begin{bmatrix} 0 & 1 & -3 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & -3 \end{bmatrix} \right\}$$

The columns that have the leading 1's of the row vectors form a basis for the column space:

$$\left\{\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}2\\1\\0\\0\\0\end{bmatrix},\begin{bmatrix}4\\-3\\1\\0\\0\end{bmatrix},\begin{bmatrix}5\\0\\-3\\1\\0\end{bmatrix}\right\}$$

(b) The non-zero rows of the matrix for a basis for the row space:

$$\left\{\begin{bmatrix}1 & 2 & -1 & 5\end{bmatrix},\begin{bmatrix}0 & 1 & 4 & 3\end{bmatrix},\begin{bmatrix}0 & 0 & 1 & -7\end{bmatrix},\begin{bmatrix}0 & 0 & 0 & 1\end{bmatrix}\right\}$$

The columns that have the leading 1's of the row vectors form a basis for the column space:

$$\left\{\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}2\\1\\0\\0\\0\end{bmatrix},\begin{bmatrix}-1\\4\\1\\0\\0\end{bmatrix},\begin{bmatrix}5\\3\\-7\\1\\0\end{bmatrix}\right\}$$

(12) Let $A = \begin{bmatrix} 1 & -2 & 5 & 0 & 3 \\ -2 & 5 & -7 & 0 & -6 \\ -1 & 3 & -2 & 1 & -3 \\ -3 & 8 & -9 & 1 & -9 \end{bmatrix}$.

(a) Use the usual procedure to find bases for $\operatorname{col}(A)$ and $\operatorname{row}(A)$.
(b) Now compute a basis for $\operatorname{row}(A)$ consisting of a *subset of the rows* of A. Hint: look at A^T.

Solution 12:

The given matrix reduces to:

$$\begin{bmatrix} 1 & 0 & 11 & 0 & 3 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The non-zero rows of the reduced matrix for a basis for the row space:

$$\left\{\begin{bmatrix}1 & 0 & 11 & 0 & 3\end{bmatrix},\begin{bmatrix}0 & 1 & 3 & 0 & 0\end{bmatrix},\begin{bmatrix}0 & 0 & 0 & 1 & 0\end{bmatrix}\right\}$$

The columns of the original matrix that correspond the the columns of the reduced matrix that have the leading 1's of the row vectors form a basis for the column space:

$$\left\{\begin{bmatrix}1\\-2\\-1\\-3\end{bmatrix},\begin{bmatrix}-2\\5\\3\\8\end{bmatrix},\begin{bmatrix}0\\0\\1\\1\end{bmatrix}\right\}$$

(b) We have $\operatorname{row}(A) = \operatorname{col}(A^T)$. Applying the usual procedure to compute a basis of $\operatorname{col}(A^T)$ produces a basis for $\operatorname{row}(A) = \operatorname{col}(A^T)$ from among the columns of A^T, which are the rows of A.
We have

$$A^T = \begin{bmatrix} 1 & -2 & -1 & -3 \\ -2 & 5 & 3 & 8 \\ 5 & -7 & -2 & -9 \\ 0 & 0 & 1 & 1 \\ 3 & -6 & -3 & -9 \end{bmatrix}$$

which reduces to:

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

Since columns 1,2,and 3 have the leading 1's, the corresponding rows of the original matrix form a basis:

$$\left\{ \begin{bmatrix} 1 & -2 & 5 & 0 & 3 \end{bmatrix}, \begin{bmatrix} -2 & 5 & -7 & 0 & -6 \end{bmatrix}, \begin{bmatrix} -1 & 3 & -2 & 1 & -3 \end{bmatrix} \right\}$$

(13) Find a subset of the given vectors that forms a basis for the space spanned by those vectors, and then express each vector that is not in the basis as a linear combination of the basis vectors.

$$\mathbf{v}_1 = (1, 0, 1, 1), \mathbf{v}_2 = (-3, 3, 7, 1), \mathbf{v}_3 = (-1, 3, 9, 3), \mathbf{v}_4 = (-5, 3, 5, -1)$$

Solution 13:

Start by writing each vector as a column in a matrix.

$$\begin{bmatrix} 1 & -3 & -1 & -5 \\ 0 & 3 & 3 & 3 \\ 1 & 7 & 9 & 5 \\ 1 & 1 & 3 & -1 \end{bmatrix}$$

Like the previous problems we reduce the matrix.

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

Now the columns of the original matrix that correspond to the columns of the reduced matrix that have the leading 1's form a basis. Thus column 1 and 2 form the basis. Since our columns are vectors, the basis is:

$$\{\mathbf{v}_1, \mathbf{v}_2\}$$

Now we find a set of dependency equations for each column in the reduced matrix that does not have a leading 1. Let $\mathbf{w}_i$ be the columns of the reduced matrix. Then

$$\mathbf{w}_3 = 2\mathbf{w}_1 + \mathbf{w}_2$$
$$\mathbf{w}_4 = -2\mathbf{w}_1 + \mathbf{w}_2$$

Replacing each $\mathbf{w}_i$ with $\mathbf{v}_i$ gives the required result:

$$\mathbf{v}_3 = 2\mathbf{v}_1 + \mathbf{v}_2$$
$$\mathbf{v}_4 = -2\mathbf{v}_1 + \mathbf{v}_2$$

(14) Find the rank and nullity of each matrix by reducing it to row echelon form.
(a)

$$A = \begin{bmatrix} 1 & 0 & -2 & 1 & 0 \\ 0 & -1 & -3 & 1 & 3 \\ -2 & -1 & 1 & -1 & 3 \\ 0 & 1 & 3 & 0 & -4 \end{bmatrix}$$

(b)
$$A = \begin{bmatrix} 1 & 3 & 1 & 3 \\ 0 & 1 & 1 & 0 \\ -3 & 0 & 6 & -1 \\ 3 & 4 & -2 & 1 \\ 2 & 0 & -4 & -2 \end{bmatrix}$$

Solution 14:

(a) The matrix reduces to:
$$\begin{bmatrix} 1 & 0 & -2 & 0 & 1 \\ 0 & 1 & 3 & 0 & -4 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Since the matrix has 3 leading 1's, $\text{rank}(A) = 3$. The rank-nullity theorem now tells us that $\text{nullity}(A) = n - 3 = 5 - 3 = 2$.

(b) A reduces to:
$$\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

This matrix has three leading 1's, which means $\text{rank}(A) = 3$ and $\text{nullity}(A) = 4 - 3 = 1$.

(15) Matrix R is the reduced row echelon form of the matrix A.
$$A = \begin{bmatrix} 0 & 2 & 2 & 4 \\ 1 & 0 & -1 & -3 \\ 2 & 3 & 1 & 1 \\ -2 & 1 & 3 & -2 \end{bmatrix}, R = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

(a) Compute the rank and nullity of A.
(b) Confirm that the rank an nullity satisfy the rank-nullity theorem.

Solution 15:

(a) Since R has three non-zero rows, $\text{rank}(A) = 3$. There is one free variable, so $\text{nullity}(A) = 1$.

(b) The matrix A is 4×4. Thus $n = 4$. In our example we have
$$\text{rank}(A) + \text{nullity}(A) = 3 + 1 \qquad\qquad \text{(from (a))}$$
$$= 4$$

Thus $\text{rank}(A) + \text{nullity}(A) = n$ as predicted by the theorem.

(16) Give two explicit matrices A and B of the same size satisfying $\text{rank}(A) = \text{rank}(B)$, but $\text{rank}(A^2) \neq \text{rank}(B^2)$.
Solution 16:

Let $A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Both clearly have rank 1. Since $A^2 = \mathbf{0}$, we have $\text{rank}(A^2) = 0$. On the other hand $B^2 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$, which has rank 1.

(17) Prove: an $n \times n$ matrix A satisfies $\mathrm{col}(A) \subseteq \mathrm{null}(A)$ if and only if $A^2 = \mathbf{0}$.
Produce an explicit nonzero example of such a matrix.
Solution 17:

Write $A = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}$. Then we have

$$A^2 = \mathbf{0} \iff \begin{bmatrix} | & | & & | \\ A\mathbf{v}_1 & A\mathbf{v}_2 & \cdots & A\mathbf{v}_n \\ | & | & & | \end{bmatrix} = \mathbf{0} \qquad \text{(column method)}$$

$$\iff A\mathbf{v}_j = \mathbf{0} \text{ for all } 1 \leq j \leq n$$
$$\iff \mathbf{v}_j \in \mathrm{null}(A) \text{ for all } 1 \leq j \leq n$$
$$\iff \mathrm{span}\left(\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}\right) \subseteq \mathrm{null}(A)$$
$$\iff \mathrm{col}(A) \subseteq \mathrm{null}(A)$$

The penultimate if and only only is a consequence of the general fact that a subspace W contains $\mathrm{span}(\{\mathbf{w}_1, \ldots, \mathbf{w}_r\})$ if and only if it contains $\mathbf{w}_i$ for all i.

(18) If A is an m by n matrix, what is the largest possible value for its rank and the smallest possible value for its nullity?
Solution 18:

Since the row vectors of A lie in R^n and the column vectors in R^m, the row space of A is at most n-dimensional and the column space is at most m-dimensional. Since the rank of A is the common dimension of its row and column space, it follows that the rank is at most the smaller of m and n. Thus:

$$\mathrm{rank}(A) \leq \min(m, n)$$

Using the rank-nullity theorem we can find a similar bound for the nullity.

$$\mathrm{nullity}(A) = n - \mathrm{rank}(A) \qquad \qquad \text{(Thm. 4.8.2)}$$
$$\geq n - \min(m, n) \qquad \qquad \text{(since } \mathrm{rank}(A) \leq \min(m, n))$$
$$= \max(n - m, 0)$$

The last equality follows from the more general rule that

$$r - \min\{m, n\} = \max\{r - m, r - n\}.$$

(19) Complete the following statements. Justify your answer.
 (a) If A is a 3 by 5 matrix, then the rank of A is at most ...
 (b) If A is a 3 by 5 matrix, then the nullity of A is at most ...

Solution 19:
(a) By the previous exercise $\mathrm{rank}(A)$ is at most $\min(3, 5) = 3$.
(b) The nullity is at most 5. Consider

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The rank of this matrix is 0. Thus the nullity must be 5. Since the rank can never be negative, the largest the nullity can be is 5.

92

(20) Let A be a 5 by 7 matrix with rank 4.
　　(a) What is the dimension of the solution space of $A\mathbf{x} = \mathbf{0}$?
　　(b) Is $A\mathbf{x} = \mathbf{b}$ consistent for all vectors $\mathbf{b}$ in $\mathbb{R}^5$?

Solution 20:
(a) The solution space of $A\mathbf{x} = \mathbf{0}$ is null(A). Thus:
$$\text{nullity}(A) = n - \text{rank}(A) = 7 - 4 = 3$$

(b) No. Since rank$(A) = 4$, dim(col(A)) $= 4$. Since col$(A) \subset \mathbb{R}^5$ and dim$(\mathbb{R}^5) = 5$ it follows that col$(A) \subsetneq \mathbb{R}^5$. Then there is $\mathbf{b} \in \mathbb{R}^5$ such that $\mathbf{b} \notin$ col(A). Since
$$\text{col}(A) = \{\mathbf{b} \in \mathbb{R}^5 \colon A\mathbf{x} = \mathbf{b} \text{ is consistent}\}$$
it follows that we cannot solve $A\mathbf{x} = \mathbf{b}$ for this $\mathbf{b}$.

(21) Prove: If a matrix A is not a square, then either the row vectors or the column vectors of A are linearly dependent.
Solution 21:

Suppose A is a m by n matrix with $n > m$. Then we have n columns in A which we can think of as being vectors in R^m. The columns of A cannot be linearly independent since R^m has dimension m, so any collection of vectors from R^m with more than m vectors must be linearly dependent. So the columns are dependent. Similarly, if $n < m$, then you can think of the rows of A as a collection of m vectors in R^n. Using the same argument we can see that rows must be dependent.

(22) True or false. If true, provide a proof; if false, give an explicit counterexample.
　　(a) Either the row vectors or the column vectors of a square matrix are linearly independent.
　　(b) A matrix with linearly independent row vectors and linearly independent column vectors is square.
　　(c) The nullity of a nonzero m by n matrix is at most m.
　　(d) Adding one additional column to a matrix increases its rank by one.
　　(e) The nullity of a square matrix with linearly dependent rows is at least one.
　　(f) If A is a square and $A\mathbf{x} = \mathbf{b}$ is inconsistent for some vector $\mathbf{b}$, the the nullity of A is zero.
　　(g) If a matrix A has more rows than columns, then the dimension of the row space is greater than the dimension of the column space.
　　(h) rank$(A) = $ rank(A^T).
　　(i) If nullity$(A^T) = $ nullity(A), then A is square.

Solution 22:
(a) False. Consider the matrix:
$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$
Both the row vectors and the column vectors are linearly dependent.
(b) True. Let A be an m by n matrix with linearly independent row vectors and linearly independent column vectors. Assume for the purpose of contradiction that $m \neq n$. Then by Exercise 21 either the row vectors or the column vectors are linearly dependent. This contradicts the fact that A has linearly independent row vectors and linearly independent column vectors. Thus the assumption is false and $m = n$, so A is a square matrix.

(c) False. The $m \times n$ zero matrix has nullity n. So if $n > m$ the nullity of 0_{mn} is greater than n.

(d) False. Consider:

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

Since the matrix has 3 leading 1's, the rank is 3. Now if we add a column of 1's

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

Since the matrix has 3 leading 1's, the rank is 3.

(e) True. Let A be an n by n matrix. Since the rows are linearly dependent, the reduced form of A has a row of zeros for the bottom row. Thus

$$\text{rank}(A) < n$$

And

$$\text{nullity}(A) = n - \text{rank}(A) > 0$$

(f) False. Use the invertibility theorem. Since A is inconsistent for some vector $\mathbf{b}$, part (e) is false. Thus part (o) is false. So A does not have nullity 0.

(g) False. By rank-nullity, the row space and the column space always have the same dimension.

(h) True. We have $\text{rank}(A) = \dim \text{col}(A) = \dim \text{row}(A^T) = \text{rank}(A^T)$.

(i) True. Let A be $m \times n$, and assume $\text{nullity}(A) = \text{nullity}(A^T)$. Then by the rank-nullity theorem

$$
\begin{aligned}
n &= \text{rank}(A) + \text{nullity}(A) && \text{(by rank-nullity theorem, since } A \text{ is } m \times n) \\
&= \text{rank}(A^T) + \text{nullity}(A) && \text{(since } \text{rank}(A) = \text{rank}(A^T)) \\
&= \text{rank}(A^T) + \text{nullity}(A^T) && \text{(by assumption)} \\
&= m && \text{(by rank-nullity theorem, since } A^T \text{ is } n \times m)
\end{aligned}
$$

(23) Suppose $\underset{m \times n}{A}$ reduces to the row echelon matrix $\underset{m \times n}{U}$.

    (a) Prove: the columns $\mathbf{u}_{i_1}, \mathbf{u}_{i_2}, \ldots, \mathbf{u}_{i_r}$ of U form a basis for $\text{col}\, U$ if and only if the corresponding columns $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_r}$ of A form a basis for $\text{col}\, A$.

    (b) Prove: the columns of U with leading 1's form a basis for $\text{col}\, U$.

    (c) Prove: $\dim \text{null}\, U = \#(\text{ free variables })$ in linear system $U\mathbf{x} = \mathbf{0}$.

    (d) Suppose $x_{j_1} = t_{j_1}, x_{j_2} = t_{j_2}, \ldots, x_{j_s} = t_{j_s}$ are the free variables of the system $U\mathbf{x} = \mathbf{0}$. Let $\mathbf{v}_{j_k}$ be the element of $\text{null}\, U$ obtained by setting $t_{j_k} = 1$ and $t_{j_\ell} = 0$ for $\ell \neq k$ in our parametric description of solutions to $U\mathbf{x} = \mathbf{0}$. Prove: $\{\mathbf{v}_{j_1}, \mathbf{v}_{j_2}, \ldots, \mathbf{v}_{j_s}\}$ is a basis for $\text{null}\, U$.

Solution 23:

Recall that A being row equivalent to U is equivalent to there being an invertible matrix Q such that

$(*)$
$$QA = U$$

We will this fact repeatedly.

(a) By the column method of matrix multiplication we have $\mathbf{u}_i = Q\mathbf{a}_i$: i.e., the i-th column of U is obtained by multiplying the i-th column of A on the left by Q. First observe that

$$\mathbf{0} = c_1\mathbf{a}_{i_1} + \cdots + c_r\mathbf{a}_{i_r} \Leftrightarrow Q\mathbf{0} = Q(c_1\mathbf{a}_{i_1} + \cdots + c_r\mathbf{a}_{i_r}) \qquad \text{(since } Q \text{ is invertible)}$$
$$\Leftrightarrow \mathbf{0} = c_1 Q\mathbf{a}_{i_1} + \cdots + c_r Q\mathbf{a}_{i_r} \qquad \text{(matrix arithmetic)}$$
$$\Leftrightarrow \mathbf{0} = c_1\mathbf{u}_{i_1} + \cdots + c_r\mathbf{u}_{i_r}$$

In particular, there a nontrivial linear combination of the $\mathbf{a}_{i_k}$ equal to $\mathbf{0}$ if and only if there is a nontrivial linear combination of the $\mathbf{u}_{i_k}$ equal to $\mathbf{0}$. Thus the $\mathbf{a}_{i_k}$ are independent if and only if the $\mathbf{u}_{i_k}$ are independent.

Next, I claim the $\mathbf{a}_{i_k}$ span $\operatorname{col} A$ if and only if the $\mathbf{u}_{i_k}$ span $\operatorname{col} U$. Indeed, suppose the $\mathbf{a}_{i_k}$ span $\operatorname{col} A$. Take $\mathbf{y} \in \operatorname{col} U$. I will show that $\mathbf{y}$ is a linear combination of the $\mathbf{u}_{i_k}$.

Since $\mathbf{y} \in \operatorname{col} U$, there is an $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{y} = U\mathbf{x}$ (since $\operatorname{col} U = \operatorname{range} U$). We have $U = QA$, and thus $\mathbf{y} = QA(\mathbf{x}) = Q(\mathbf{w})$, where $\mathbf{w} = A\mathbf{x}$. Then $\mathbf{w} \in \operatorname{col} A$, and we can write $\mathbf{w} = c_1\mathbf{a}_{i_1} + \cdots + c_r\mathbf{a}_{i_r}$. It then follows that

$$\mathbf{y} = Q\mathbf{w} = Q(c_1\mathbf{a}_{i_1} + \cdots + c_r\mathbf{a}_{i_r}) = c_1 Q\mathbf{a}_{i_1} + \cdots + c_r Q\mathbf{a}_{i_r} = c_1\mathbf{u}_{i_1} + \cdots + c_r\mathbf{u}_{i_r}.$$

Since $\mathbf{y}$ was any element of $\operatorname{col} U$, we see that the $\mathbf{u}_{i_k}$ span $\operatorname{col} U$, as desired.

To go the other way (i.e., that if the $\mathbf{u}_{i_k}$ span $\operatorname{col} U$, the $\mathbf{a}_{i_k}$ span $\operatorname{col} A$), note that $(*)$ implies $A = Q^{-1}U$, and we can use the same argument above with the roles of $\mathbf{a}_{i_k}$ and $\mathbf{u}_{i_k}$ swapped!

We have shown that the $\mathbf{a}_{i_k}$ are independent if and only if the $\mathbf{u}_{i_k}$, and that they span $\operatorname{col} A$ if and only if the $\mathbf{u}_{i_k}$ span $\operatorname{col} U$. It follows that the $\mathbf{a}_{i_k}$ form a basis for $\operatorname{col} A$ if and only if the $\mathbf{u}_{i_k}$ form a basis for $\operatorname{col} U$.

(b) Now let $\mathbf{u}_{i_1}, \ldots, \mathbf{u}_{i_r}$ be the columns of U with leading 1's, and let $\mathbf{u}_{j_1}, \mathbf{u}_{j_2}, \ldots, \mathbf{u}_{j_s}$ be the columns without leading 1's. To prove the $\mathbf{u}_{i_k}$ form a basis for $\operatorname{col} U$, I will show that given any $\mathbf{y} \in \operatorname{col} U$ there is a *unique* choice of scalars $c_1, c_2, \ldots, c_r$ such that $c_1\mathbf{u}_{i_1} + \cdots + c_r\mathbf{u}_{i_r} = \mathbf{y}$. (Recall that the uniqueness of this choice implies linear independence.)

So assume $\mathbf{y} \in \operatorname{col} U$. Then we can find $\mathbf{x} \in \mathbb{R}^n$ such that $U\mathbf{x} = \mathbf{y}$, which means the linear system with augmented matrix $[\, U \mid \mathbf{y}\,]$ is consistent. Using our Gaussian elimination theory, we know that the solutions $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ to this system are in 1-1 correspondence with choices for the free variables $x_{j_1} = t_{j_1}, x_{j_2} = t_{j_2}, \ldots, x_{j_s} = t_{j_s}$. (Remember that the columns $\mathbf{u}_{j_k}$ without leading 1's correspond to the free variables.) In particular, there is a unique solution to $U\mathbf{x} = \mathbf{y}$ where we set all the free variables equal to 0. By the column method, this gives us a unique linear combination of only the columns $\mathbf{u}_{i_k}$ with leading 1's equal to $\mathbf{y}$. This proves the claim.

(c) Using the notation and result from (b) we see that $\operatorname{rank} U = \dim \operatorname{range} U = \dim \operatorname{col} U = r$, the number of columns of U with leading 1's. By the rank-nullity theorem, $\dim \operatorname{null} U = n - r = s$, the number of columns without leading 1's. This is also equal to the number of free variables in the corresponding system of equations.

(d) The given recipe produces a list of s distinct vectors in $\operatorname{null} U$. Since $s = \dim \operatorname{null} U$, by part (c), it suffice to show the $\mathbf{v}_i$ are linearly independent. This is easy to see. Indeed suppose we have $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_s\mathbf{v}_s = \mathbf{0}$. Since for each $1 \leq i \leq s$, $\mathbf{v}_i$ is the *only* vector with a nonzero entry for the i-th component, we must have $c_i = 0$ for all $1 \leq i \leq s$. Thus we see that the set is linearly independent.

(1) Find the coordinate vector $[\mathbf{v}]_B$ relative to the basis $B = \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ for $\mathbb{R}^3$.
  (a) $\mathbf{v} = (2, -1, 3)$; $\mathbf{v}_1 = (1, 0, 0)$, $\mathbf{v}_2 = (2, 2, 0)$, $\mathbf{v}_3 = (3, 3, 3)$
  (b) $\mathbf{v} = (5, -12, 3)$; $\mathbf{v}_1 = (1, 2, 3)$, $\mathbf{v}_2 = (-4, 5, 6)$, $\mathbf{v}_3 = (7, -8, 9)$

Solution 1:

In both parts, the coordinate vector $[\mathbf{v}]_S$ is computed by finding c_1, c_2, c_3 such that $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. We do so using GE. (a)

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

$$\begin{aligned} c_3 &= 1 \\ c_2 &= -1/2 - 3/2 = -2 \\ c_1 &= 2 - 3 + 4 = 3 \end{aligned}$$

Thus $[\mathbf{v}]_B = (3, -2, 1)$.

(b)

$$\begin{bmatrix} 1 & -4 & 7 & 5 \\ 2 & 5 & -8 & -12 \\ 3 & 6 & 9 & 3 \end{bmatrix} \xrightarrow{2r_1 - r_2} \begin{bmatrix} 1 & -4 & 7 & 5 \\ 0 & -13 & 22 & 22 \\ 3 & 6 & 9 & 3 \end{bmatrix}$$

$$\xrightarrow{3r_1 - r_3} \begin{bmatrix} 1 & -4 & 7 & 5 \\ 0 & -13 & 22 & 22 \\ 0 & -18 & 12 & 12 \end{bmatrix}$$

$$\xrightarrow{-1/13 r_2} \begin{bmatrix} 1 & -4 & 7 & 5 \\ 0 & 1 & -22/13 & -22/13 \\ 0 & -18 & 12 & 12 \end{bmatrix}$$

$$\xrightarrow{18r_2 + r_3} \begin{bmatrix} 1 & -4 & 7 & 5 \\ 0 & 1 & -22/13 & -22/13 \\ 0 & 0 & -240/13 & -240/13 \end{bmatrix}$$

$$\xrightarrow{-13/240 r_3} \begin{bmatrix} 1 & -4 & 7 & 5 \\ 0 & 1 & -22/13 & -22/13 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\begin{aligned} c_3 &= 1 \\ c_2 &= -22/13 + 22/13 = 0 \\ c_1 &= 5 - 7 = -2 \end{aligned}$$

Thus the coordinate vector is $[\mathbf{v}]_B = (-2, 0, 1)$.

(2) Find the coordinate vector $[p]_B$ relative to the basis $B = \{p_1, p_2, p_3\}$ for P_2.
  (a) $p = 4 - 3x + x^2$, $p_1 = 1$, $p_2 = x$, $p_3 = x^2$
  (b) $p = 2 - x + x^2$, $p_1 + x = 1$, $p_2 = 1 + x^2$, $p_3 = x + x^2$

Solution 2:

In both parts, the coordinate vector $[\mathbf{v}]_B$ is computed by finding c_1, c_2, c_3 such that $p = c_1p_1 + c_2p_2 + c_3p_3$. As usual this boils down (after grouping like terms and equating like coefficients) to solving a certain linear system, which we do either by inspection, or by using

GE.

(a) By inspection, we see that $p = 4p_1 + (-3)p_2 + p_3$. Thus $[p]_B = (4, -3, 1)$.

(b) We solve the corresponding linear system in this case using GE:

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{r_2 \leftrightarrow r_3} \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & -1 \end{bmatrix}$$

$$\xrightarrow{r_1 - r_3} \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & -1 & 3 \end{bmatrix}$$

$$\xrightarrow{r_2 - r_3} \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & -2 \end{bmatrix}$$

$$\begin{aligned} c_3 &= -1 \\ c_2 &= 1 + 1 = 2 \\ c_1 &= 2 - 2 = 0 \end{aligned}$$

Thus the coordinate vector is $[p]_B = (0, 2, -1)$.

(3) The set $B = \{A_1, A_2, A_3, A_4\}$ is a basis for M_{22}, where

$$A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A_4 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

(a) Compute $[A]_B$, where

$$A = \begin{bmatrix} 6 & 2 \\ 5 & 3 \end{bmatrix}.$$

(b) Now compute $[A]_B$ for the general matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Solution 3:

(a) Finding c_1, c_2, c_3, c_4 such that $A = c_1 A_1 + c_2 A_2 + c_3 A_3 + c_4 A_4$ amounts to solving the system corresponding to the augmented matrix below:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 6 \\ 0 & 1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 1 & 5 \\ 0 & 0 & 1 & 0 & 3 \end{bmatrix}$$

, or equivalently, the matrix equation

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \\ 5 \\ 3 \end{bmatrix}$$

After GE this reduces to

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 6 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right]$$

Thus

$$\begin{aligned} c_4 &= 4 \\ c_3 &= 3 \\ c_2 &= 2 \\ x_1 &= 1 \end{aligned}$$

Then $A = 1A_1 + 2A_2 + 3A_3 + 4A_4$ and the coordinate vector is:
$[A]_B = (1, 2, 3, 4)$.

(b) For the general matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the same reasoning shows that we need to solve the matrix equation

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

We could proceed exactly as above, using Gaussian elimination. Instead, I observe that A is invertible, and solve (after computing A^{-1})

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = A^{-1} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

$$= \begin{bmatrix} a - b - d \\ b \\ d \\ -a + b + c + d \end{bmatrix}$$

Thus for general $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ we see that

$$[A]_B = (a - b - d, b, d, -a + b + c + d)$$

(4) Suppose $\dim(V) = n$, and let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V. Define

$$T : V \to \mathbb{R}^n$$

$$\mathbf{v} \mapsto [\mathbf{v}]_B.$$

Prove all statements of the coordinate vector map theorem, as listed below.
(a) T is a linear transformation.
(b) $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ if and only if $\mathbf{v}_1 = \mathbf{v}_2$: i.e., T is *one-to-one*.

In particular, $T(\mathbf{v}) = \mathbf{0}$ if and only if $\mathbf{v} = \mathbf{0}_V$.

(c) $\operatorname{range} T = \mathbb{R}^n$: i.e. T is *onto*.

(d) A set $S = \{w_1, w_2, \ldots, w_r\} \subseteq V$ is linear independent if and only if $T(S) = \{T(\mathbf{w}_1), T(\mathbf{w}_2), \ldots, T(\mathbf{w}_r)\} \subseteq \mathbb{R}^n$ is linearly independent.

(e) A set $S = \{w_1, w_2, \ldots, w_r\} \subseteq V$ spans V if and only if $T(S) = \{T(\mathbf{w}_1), T(\mathbf{w}_2), \ldots, T(\mathbf{w}_r)\}$ spans $\mathbb{R}^n$.

Solution 4:

(a) We must show $[c\mathbf{v} + d\mathbf{w}]_B = c[\mathbf{v}]_B + d[\mathbf{w}]_B$ for all $\mathbf{v}, \mathbf{w} \in V$ and all $c, d \in \mathbb{R}$. To this end, given $\mathbf{v}, \mathbf{w}$, write

$$\mathbf{v} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \cdots + a_n \mathbf{v}_n$$
$$\mathbf{w} = b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \cdots + b_n \mathbf{v}_n$$

We can do this since B spans V. It follows that for any $c, d \in \mathbb{R}$

$$c\mathbf{v} + d\mathbf{w} = (ca_1 + db_1)\mathbf{v}_1 + (ca_2 + db_2)\mathbf{v}_2 + \cdots + (ca_n + db_n)\mathbf{v}_n$$

Now, using the definition of coordinate vectors we have

$$[c\mathbf{v} + d\mathbf{w}]_B = \begin{bmatrix} ca_1 + db_1 \\ ca_2 + db_2 \\ \vdots \\ ca_n + db_n \end{bmatrix}$$

$$= \begin{bmatrix} ca_1 \\ ca_2 \\ \vdots \\ ca_n \end{bmatrix} + \begin{bmatrix} db_1 \\ db_2 \\ \vdots \\ db_n \end{bmatrix}$$

$$= c\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + d\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$= c[\mathbf{v}]_B + d[\mathbf{w}]_B.$$

(b) I begin, somewhat ponderously, by first proving that $T(\mathbf{v}) = [\mathbf{v}]_B = (0, 0, \ldots, 0)$ if and only if $\mathbf{v} = \mathbf{0}$. This is fairly obvious. Indeed, given an arbitrary $\mathbf{v} \in V$, we have $\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_n \mathbf{v}_n$ for some $c_i \in \mathbb{R}$, and $[\mathbf{v}]_B = (c_1, c_2, \ldots, c_n)$. Clearly $[\mathbf{v}]_B = (0, 0, \ldots, 0)$ if and only if $c_1 = c_2 = \cdots = c_n = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

We have proved a special case of the general claim: namely $T(\mathbf{v}) = T(\mathbf{0})$ iff $\mathbf{v} = \mathbf{0}$. The general case follows by using the fact that T is linear. We have

$$T(\mathbf{v}_1) = T(\mathbf{v}_2) \Leftrightarrow T(\mathbf{v}_1) - T(\mathbf{v}_2) = (0, 0, \ldots, 0)$$
$$\Leftrightarrow T(\mathbf{v}_1 - \mathbf{v}_2) = (0, 0, \ldots, 0) \qquad (T \text{ is linear})$$
$$\Leftrightarrow \mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0} \qquad (\text{by our special case!})$$
$$\Leftrightarrow \mathbf{v}_1 = \mathbf{v}_2$$

This same trick can be used to show, in general, that a linear transformation T is one-to-one if and only if $\operatorname{null} T = \{\mathbf{0}\}$.

(c) This is obvious. Given any $(c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$, set $\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_n \mathbf{v}_n$. Then by definition $[\mathbf{v}]_B = (c_1, c_2, \ldots, c_n)$.

(d) We can now use our results above to prove the remaining statements. Note that

$$c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_r\mathbf{w}_r = \mathbf{0} \Leftrightarrow T(c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_r\mathbf{w}_r) = (0,0,\ldots,0) \qquad \text{(special case of (b)!)}$$
$$\Leftrightarrow c_1T(\mathbf{w}_1) + c_2T(\mathbf{w}_2) + \cdots + c_rT(\mathbf{w}_r) = (0,0,\ldots,0) \qquad (T \text{ is linear})$$

From this we see that the $\mathbf{w}_i$ are linearly independent if and only if the $T(\mathbf{w}_i)$ are linearly independent.

(e) Suppose the $\mathbf{w}_i$ span. Given any $\mathbf{y} \in \mathbb{R}$, there is a $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{y}$, since range $T = \mathbb{R}^n$ (part (c)). Since the $\mathbf{w}_i$ span, we can write $\mathbf{v} = d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_r\mathbf{w}_r$. Then

$$\mathbf{y} = T(\mathbf{v}) = T(d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_r\mathbf{w}_r) = d_1T(\mathbf{w}_1) + d_2T(\mathbf{w}_2) + \cdots + d_rT(\mathbf{w}_r),$$

showing that the $T(\mathbf{w}_i)$ span $\mathbb{R}^n$.

Conversely, suppose the $T(\mathbf{w}_i)$ span $\mathbb{R}^n$. Given any $\mathbf{v} \in V$, we have $T(\mathbf{v}) = d_1T(\mathbf{w}_1) + d_2T(\mathbf{w}_2) + \cdots + d_rT(\mathbf{w}_r)$, for some $d_i \in \mathbb{R}$, since $T(\mathbf{v}) \in \mathbb{R}^n$ and the $T(\mathbf{w}_i)$ span $\mathbb{R}^n$ by assumption. But then

$$T(\mathbf{v}) = T(d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_r\mathbf{w}_r),$$

since T is linear. It follows that

$$\mathbf{v} = d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_r\mathbf{w}_r,$$

since T is 1-1 by (b) !! This shows that the $\mathbf{w}_i$ span V, as desired.
Note: properties (d) and (e) can be summarized by saying T preserves linear independence and T preserves spanning sets. This property is enjoyed more generally by any isomorphism between vector spaces.

(5) Let $S = \{p_1 = 1 + x + 2x^2, p_2 = 1 - x, p_3 = 1 + x^2, p_4 = 1 + x - x^2\}$ and let $W = \text{span}(S) \subset P$. (Recall P is the space of all polynomials.)

(a) Use "street smarts" to decide whether S is linearly independent.

(b) Use coordinate vectors and an appropriate fundamental space algorithm to choose a basis of $W = \text{span}(S)$ *from among the elements of S*.

(c) Give a satisfying description of W.

Solution 5:

(a) In fact we see that S lies within the smaller subspace P_2, which has dimension 3. Since S has 4 elements and lives in a 3-dimensional space, it is guaranteed to be dependent. Street smarts!

(b) Let B be the standard basis of P_2: i.e., $B = \{1, x, x^2\}$. Then we can translate this whole problem into $\mathbb{R}^3$ using coordinate vectors. Namely we let:

$$\mathbf{v}_1 = (p_1)_B = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \mathbf{v}_2 = (p_2)_B = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v}_3 = (p_3)_B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_4 = (p_4)_B = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix},$$

and set $W' = \text{span}(\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\})$.

Putting the $\mathbf{v}_i$ into the columns of a matrix A and applying our column space algorithm tells us that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ form a basis for W'. (You do the work.)

Translating back to P_3 we conclude that the corresponding polynomials $\{p_1, p_2, p_4\}$ form a basis for W.

(c) Since W has a basis containing three elements, we see that $\dim(W) = 3$. Since $W \subset P_2$ and $\dim(P_2) = \dim(W) = 3$, we see by the dimension theorem compendium (subspace part) that W is in fact all of P_2.

(6) Let $V = P_2$, and let $S = \{x^2 + 2x + 1, 3x^2 + 6x\}$. Extend S to a basis of P_2 by first translating the problem to $\mathbb{R}^3$ using coordinate vectors and applying the relevant algorithm there.

Solution 6:

Let $B = \{1, x, x^2\}$ be the standard basis of P_2. Applying coordinate vectors to S yields the set $\{(1, 2, 1), (0, 6, 3)\}$. The usual "extend to a basis" algorithm tells us that $\{(1, 2, 1), (0, 6, 3), (0, 1, 0)\}$ is a basis for $\mathbb{R}^3$. Now translate back to our original vector space to conclude that $\{x^2 + 2x + 1, 3x^2 + 6x, x\}$ is a basis for P_2 extending S. (Note that the vector $(0, 1, 0)$ corresponds to the polynomial $p(x) = 0 + 1x + 0 = x$ via the coordinate vector map.)

(7) Let $V = P_3$ and consider the polynomials

$$p_1 = x^3 + 1, p_2 = 2x^3 + x + 1, p_3 = 3x^3 + 2x + 1, p_4 = 2x^3 + x^2 + x + 1,$$

Let $S = \{p_1, p_2, p_3, p_4\}$ and define $W = \text{span}(S)$.

Find a subset of S that is a basis of W and compute $\dim(W)$. Use coordinate vectors!

Solution 7:

General technique: translate problem to $\mathbb{R}^4$ using coordinate vectors, answer the problem there using one of the fundamental space algorithms, then translate the results back into the original P_3 context. Let $B = \{1, x, x^2, x^3\}$ be the standard basis of P_3, and translate the whole problem to $\mathbb{R}^4$ using coordinate vectors. That is, we let

$$\mathbf{v}_1 = [p_1]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad\qquad \mathbf{v}_2 = [p_2]_B = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}$$

$$\mathbf{v}_3 = [p_3]_B = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 3 \end{bmatrix} \qquad\qquad \mathbf{v}_4 = [p_4]_B = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

and let $W' = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\} \subseteq \mathbb{R}^4$. We use the *column space* algorithm to determine a basis of W' from among the $\mathbf{v}_i$. That is $W' = \text{col}(A)$ where

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 3 & 2 \end{bmatrix} \xrightarrow{\text{row reduce}} U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the first, second and fourth columns of U form a basis for $\text{col}(U)$, the same is true of A, which means $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is a basis for W'.

Lastly, returning back to P_3 we conclude that $\{p_1, p_2, p_4\}$ is a basis for W in P_3.

(8) Let

$$S = \left\{ A_1 = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}, A_3 = \begin{bmatrix} -1 & 1 \\ -4 & -1 \end{bmatrix}, A_4 = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix} \right\}$$

and define $W = \text{span}(S) \subseteq M_{22}$.

(a) Compute a basis B of W from among the elements of S.

You should first "translate" the problem into $\mathbb{R}^4$ using coordinate vectors and apply an appropriate fundamental space algorithm.

(b) Show that in fact W is equal to the subspace $W' = \left\{ \begin{bmatrix} a & b \\ c & a \end{bmatrix} : a, b, c \in \mathbb{R} \right\}$.

Make your life easier by using a dimension argument, but make sure all your claims are fully justified.

Solution 8:

(a) Translate the problem to $\mathbb{R}^4$ using the coordinate vector map with respect to the standard basis, and identify in this way the given vector space with $\mathrm{col}(A)$, where

$$ A = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 2 & 1 & 1 & 1 \\ 1 & 2 & -4 & 2 \\ 1 & 1 & -1 & 0 \end{bmatrix} $$

The GE procedure for $\mathrm{col}(A)$ tells us that the first, second and fourth columns form a basis for $\mathrm{col}(A)$. Translating back to M_{22}, we conclude that $B = \{A_1, A_2, A_4\}$ is a basis for W.

(b) We see easily that W' has basis

$$ B' = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}. $$

Thus $\dim W' = 3$. Since $B \subseteq W'$, it follows that $W = \mathrm{span}(B) \subseteq W'$. Since $\dim W = \dim W' = 3$, we conclude by the dimension theorem compendium (subspace part) that $W = W'$.

4.4: MATRIX REPRESENTATIONS

(1) Define $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ to be reflection through the plane $\mathcal{P} \colon x + y + z = 0$. In other words, T takes an input (x, y, z) and sends it to its reflection (x', y', z') on the other side of $\mathcal{P}$. You may take it on faith that T is a linear transformation.
  (a) Draw a picture of $\mathcal{P}$, an arbitrary point (x, y, z), and its reflection (x', y', z'). Add some right triangles to your picture to indicate the relation between (x, y, z) and (x', y', z').
  (b) Pick a *nonstandard* basis B' of $\mathbb{R}^3$, for which the behavior of T is easy to determine: i.e., pick a basis that reflects the geometry of this picture. Compute $[T]_{B'}$.

Solution 1:

(a) In general to draw the reflection of a point $P = (x, y, z)$ through a plane $\mathcal{P}$, draw the line ℓ passing through (x, y, z) that is perpendicular to $\mathcal{P}$ (so parallel to its normal vector), and call its intersection with the plane Q. (If you recall, Q is just the projection of P onto $\mathcal{P}$. The reflection of P is then the point P' on ℓ on the other side of $\mathcal{P}$, whose distance to Q is equal to the distance from P to Q. If P is in the plane to begin with, then it is equal

to its own reflection.
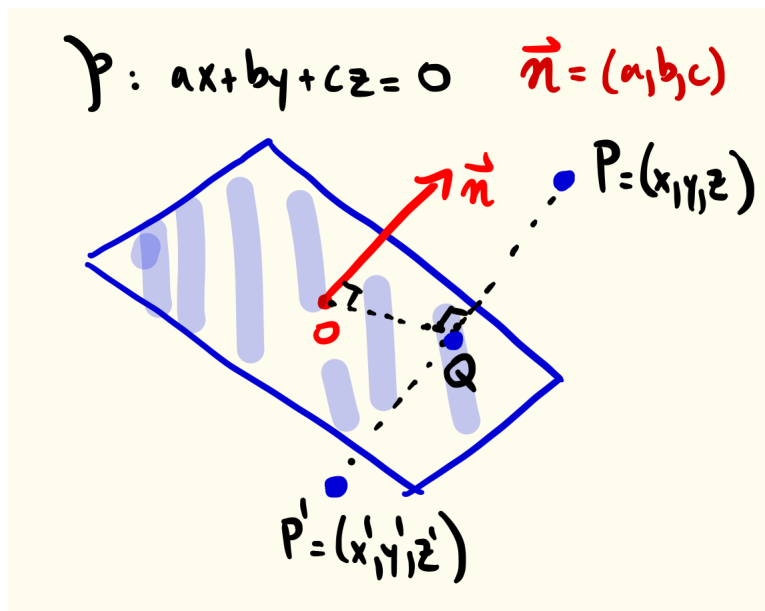
(b) The description above helps us pick a basis B of $\mathbb{R}^3$ that makes $[T]_B$ particularly simple. Start of by choosing a basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for the given plane $\mathcal{P}$: in our case, the vectors $\mathbf{v}_1 = (1, -1, 0), \mathbf{v}_2 = (0, 1, -1)$ will do. To extend this to a full basis of $\mathbb{R}^3$, we add the normal vector $\mathbf{n}$ to the plane. In our case, the basis is thus $B = \{\mathbf{v}_1 = (1, -1, 0), \mathbf{v}_2 = (0, 1, -1), \mathbf{n} = (1, 1, 1)\}$.

Now, we have $T(\mathbf{v}_1) = \mathbf{v}_1$ and $T(\mathbf{v}_2) = \mathbf{v}_2$, since $\mathbf{v}_1$ and $\mathbf{v}_2$ lie in the plane already. Furthermore, from the picture we describe in (a) it is clear that $T(\mathbf{n}) = -\mathbf{n}$. This is because the projection of $\mathbf{n}$ onto the plane is just $\mathbf{0}$; so its reflection is just $-\mathbf{n}$.

Lastly we now easily compute the coordinate vectors of these outputs: $[T(\mathbf{v}_1)]_B = [\mathbf{v}_1]_B = (1, 0, 0), [T(\mathbf{v}_2)]_B = [\mathbf{v}_2]_B = (0, 1, 0)$, and $[T(\mathbf{n})]_B = [-\mathbf{n}]_B = (0, 0, -1)$. We conclude that

$$[T]_B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

(2) Consider the differential equation

(∗)
$$f''(x) + f'(x) = x^2 + x + 1$$

The theory of ODE's tells us that there are infinitely many solutions $f(x)$ to (∗). We will use linear algebra to find infinitely many polynomial solutions.

(a) First define $T\colon P_3 \to P_2$ as $T(f) = f''(x) + f'(x)$. Explain why $T(f)$ indeed lies in P_2 and show that T is linear.

(b) Compute $A = [T]_B^{B'}$ where B and B' are the standard bases of P_3 and P_2, respectively.

(c) Determine $\text{null}(A)$ and $\text{col}(A)$.

(d) Find all solutions in P_3 to the differential equation (∗). (First solve the relevant matrix equation involving A, then "lift up" to P_3.)

Note: though we have found all *polynomial solutions* to the differential equation (∗), we have not found *all* solutions. Observe that $g(x) = e^{-x}$ is a solution to the corresponding homogeneous equation $f'' + f' = 0$, It follows from basic ODE theory that we can add ce^{-x} to any of our solutions in (d) to get a new solution. Our analysis didn't catch these solutions as we restricted our attention to P_3.

Solution 2:

(a) The usual proof shows T is linear. Since taking the derivative knocks down the degree of a polynomial by 1, it's clear that if $p(x)$ has degree at most 3, then $T(p) = p'' + p'$ will have degree at most 2.

(b) Let $B = \{1, x, x^2, x^3\}$ and $B' = \{1, x, x^2\}$. The usual recipe for $[T]_{B',B}$ begins by computing

$$T(1) = 0, T(x) = 1, T(x^2) = 2 + 2x, T(x^3) = 6x + 3x^2.$$

Taking coordinate vectors with respect to B and placing these in columns yields

$$A = \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

(c) The rank of A is clearly 3 (look at rows), and by inspection we see that $\{(1,0,0), (2,2,0), (0,6,3)\}$ is a basis for $\mathrm{col}(A)$. Since $\dim \mathrm{col}(A) = 3$ and $\mathrm{col}(A) \subset \mathbb{R}^3$, it follows from the Subspace Dimension Theorem that $\mathrm{col}(A) = \mathbb{R}^3$.

Since $\mathrm{rank}(A) = 3$, it follows from R-N that $\mathrm{nullity}(A) = 4 - 3 = 1$. To come up with a basis for $\mathrm{null}(A)$ we simply need to find one nonzero element of $\mathrm{null}(A)$. We see $\{(1,0,0)\}$ does the trick.

What does all this mean about T? Since $\mathrm{col}(A) = \mathbb{R}^3$, it follows that $\mathrm{range}(T) = P_2$. Thus for *any* $q(x) \in P_2$, we can find a $p(x) \in P_3$ with $T(p) = p''(x) + p'(x) = q(x)$. This means not only can we solve the given differential equation, we can solve any similar equation of the form $f'' + f' = q$, where q is a polynomial in P_2.

Furthermore, as we will see, the fact that $\mathrm{null}(A)$ is nontrivial will mean that we can always find infinitely many solutions to $(*)$.

(d) The polynomial $x^2 + x + 1$ corresponds via $[\ ]_{B'}$ to the vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. All solutions $\mathbf{x}$

to

$$A\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

will lift up to the solutions p to $T(p) = x^2 + x + 1$.

Using GE, we see the general solution to the matrix equation is

$$\mathbf{x} = c \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -\frac{1}{2} \\ \frac{1}{3} \end{bmatrix}$$

Thus the general *polynomial* solution to the original differential equation is

$$p(x) = c + (1 - \frac{1}{2}x^2 + \frac{1}{3}x^3) = (c+1) - \frac{1}{2}x^2 + \frac{1}{3}x^3.$$

(3) Define $T\colon M_{22} \to M_{22}$ by $T(A) = A^T - A$.
 (a) Compute $A = [T]_B$ where B is the standard basis for M_{22}.
 (b) Compute bases for $\mathrm{null}(A)$ and $\mathrm{col}(A)$.
 (c) Use your result in (b) to give bases for $\mathrm{null}(T)$ and $\mathrm{range}(T)$.
 (d) Identify $\mathrm{null}(T)$ and $\mathrm{range}(T)$ as familiar subspaces of matrices.

Solution 3:

(a) The usual recipe yields

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

We have $\text{null}(A) = \text{span}(\{(1,0,0,0),(0,0,0,1),(0,1,1,0)\})$ and $\text{col}(A) = \text{span}(\{(0,-1,1,0)\})$. By lifting we get that

$$\text{null}(T) = \text{span}\left(\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right\}\right), \quad \text{range}(T) = \text{span}\left(\left\{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right\}\right).$$

From these descriptions it follows easily that $\text{null}(T)$ is the subspace of all symmetric matrices, and $\text{range}(T)$ is the subspace of all skew-symmetric matrices.

(4) Let $T : P_2 \to P_3$ be the linear transformation defined by $T(p) = x \cdot p(x-3)$.
 (a) Find $[T]_B^{B'}$ relative to the bases $B = \{1, x, x^2\}$ and $B' = \{1, x, x^2, x^3\}$.
 (b) Use $[T]_B^{B'}$ to compute bases for $\text{null}\, T$ and $\text{range}\, T$.

Solution 4:

By the formula for T

$$\begin{aligned} T(1) &= x(1) = x \\ T(x) &= x(x-3) = x^2 - 3x \\ T(x^2) &= x^3 - 6x^2 + 9x \end{aligned}$$

Taking these vectors relative the the basis B', we can write

$$[T]_B^{B'} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -3 & 9 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix}$$

(5) Let $T : P_2 \to M_{22}$ be the linear transformation defined by

$$T(p) = \begin{bmatrix} p(0) & p(1) \\ p(-1) & p(0) \end{bmatrix}$$

let B be the standard basis for M_{22}, let $B' = \{1, x, x^2\}$ and $B'' = \{1, 1+x, 1+x^2\}$ be bases for P_2.
 (a) Find $[T]_{B'}^B$ and $[T]_{B''}^B$.
 (b) Use either one of the matrix representations from (a) to compute bases for $\text{null}\, T$ and $\text{range}\, T$.

Solution 5:

Lets work on $[T]_B^{B'}$ first. By the formula for T

$$\begin{aligned} T(1) &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\[4pt] T(x) &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\[4pt] T(x^2) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{aligned}$$

Taking these vectors relative the the basis B, we can write

$$[T]_B^{B'} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Now for $[T]_B^{B''}$. By the formula for T

$$T(1) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$T(1+x) = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$T(1+x^2) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

Taking these vectors relative the the basis B, we can write

$$[T]_B^{B''} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix}$$

(6) Let $T : M_{22} \to R^2$ be the linear transformation given by

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a+b+c \\ d \end{bmatrix}$$

and let B be the standard basis for M_{22}, B' be the standard basis for R^2 and let

$$B'' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\}$$

be another basis for R^2. Find $[T]_B^{B'}$ and $[T]_B^{B''}$.

Solution 6:

By the formula for T, we can find $T(\mathbf{u}_i)$ where $\mathbf{u}_i$ are the vectors in the basis B.

$$T(\mathbf{u}_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$T(\mathbf{u}_2) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$T(\mathbf{u}_3) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$T(\mathbf{u}_4) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Taking these vectors relative the the basis B', we can write

$$[T]_B^{B'} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Taking these vectors relative the the basis B'', we can write

$$[T]_B^{B''} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix}.$$

4.5: CHANGE OF BASIS

(1) Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Compute $\underset{B \to B'}{P}$ for each of the following bases B', and explain what happens to coordinate vectors as we change from B to B'.
   (a) $B' = \{\mathbf{v}_2, \mathbf{v}_1, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$.
   (b) $B' = \{3\mathbf{v}_1, 3\mathbf{v}_2, \ldots, 3\mathbf{v}_n\}$.
   Solution 1:

   In each case we will describe the change of basis matrix column by column, making use $\mathbf{e}_j$, the elements of the standard basis (or equivalently, the columns of the identity matrix).
   The computation of each column is done by inspection, since the new basis is closely related to the old one.

   (a) We have $\underset{B \to B'}{P} = \begin{bmatrix} | & | & & | \\ \mathbf{e}_2 & \mathbf{e}_1 & \cdots & \mathbf{e}_n \\ | & | & & | \end{bmatrix}$.

   When changing from B to B', simply swap the first two entries of the coordinate vector.

   (b) We have $\underset{B \to B'}{P} = \begin{bmatrix} | & | & & | \\ \frac{1}{3}\mathbf{e}_1 & \frac{1}{3}\mathbf{e}_2 & \cdots & \frac{1}{3}\mathbf{e}_n \\ | & | & & | \end{bmatrix}$.

   When changing from B to B', scale each entry of the coordinate vector by $1/3$.

(2) Let $V = P_1$ with basis $B' = \{2x + 1, 3x + 2\}$.
   (a) Compute $[p(x)]_{B'}$, where $p(x) = ax + b$, $a, b \in \mathbb{R}$. Your answer will be in terms of a and b.
   (b) Now do the same computation using a change of basis matrix involving the standard basis B.

   Solution 2:
   (a) We seek the c_1, c_2 such that $c_1(2x + 1) + c_2(3x + 2) = ax + b$. This boils down to solving the matrix equation

   $$\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \iff \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \left(\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \right)^{-1} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2a - 3b \\ -a + 2b \end{bmatrix}$$

   (b) To do the same computation with a change of basis matrix, we let $B = \{x, 1\}$. Then $\underset{B' \to B}{P} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$, and hence

   $$\underset{B \to B'}{P} = \left(\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}.$$

   Thus
   $$[ax + b]_{B'} = \underset{B \to B'}{P}[ax + b]_B = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2a - 3b \\ -a + 2b \end{bmatrix}.$$

(3) Let V be a finite-dimensional vector space, and let B, B', B'' be three different bases of V. Let $P = \underset{B \to B'}{P}$ and $Q = \underset{B \to B''}{P}$.
   (a) Express $\underset{B' \to B''}{P}$ in terms of P, Q, and/or their inverses.
   Verify that your matrix satisfies the defining property of the change of basis matrix.

107

(b) Utilize the formula in (a) to compute $\underset{B'\to B''}{P}$, where $B' = \{2x+1, 3x+2\}$ and $B'' = \{x-1, x+1\}$: two nonstandard bases of P_1.

Solution 3:

(a) I claim $\underset{B'\to B''}{P} = QP^{-1}$ To verify the claim, I need only show that QP^{-1} satisfies the defining property: i.e., $QP^{-1}[\mathbf{v}]_{B'} = [\mathbf{v}]_{B''}$ for all $\mathbf{v} \in V$. We compute

$$QP^{-1}[\mathbf{v}]_{B'} = Q\underset{B'\to B}{P}[\mathbf{v}]_{B'} \qquad\qquad \left(\underset{B'\to B}{P} = (\underset{B\to B'}{P})^{-1}\right)$$
$$= Q[\mathbf{v}]_B$$
$$= [\mathbf{v}]_{B''} \qquad\qquad (Q = \underset{B\to B''}{P})$$

This proves that $QP^{-1} = \underset{B'\to B''}{P}$.

(b) Let $B = \{x, 1\}$. Then we have $\underset{B'\to B}{P} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$, and $\underset{B\to B''}{P} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

We have

$$\underset{B'\to B''}{P} = \underset{B\to B''}{P}\underset{B'\to B}{P} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 3 & 5 \end{bmatrix}$$

(4) Let B, B', and B'' be three different bases for a finite dimensional vector space V. Show that

$$\underset{B\to B''}{P} = \underset{B'\to B''}{P} \cdot \underset{B\to B'}{P}$$

using only the defining property and uniqueness of change of basis matrices.

Solution 4:

Let $C = \underset{B'\to B''}{P} \cdot \underset{B\to B'}{P}$. Then we have

$$C[\mathbf{v}]_B = \underset{B'\to B''}{P} \cdot \underset{B\to B'}{P}[\mathbf{v}]_B$$
$$= \underset{B'\to B''}{P}(\underset{B\to B'}{P}[\mathbf{v}]_B)$$
$$= \underset{B'\to B''}{P}[\mathbf{v}]_{B'} \qquad\qquad (\text{prop. of } \underset{B\to B'}{P})$$
$$= [\mathbf{v}]_{B''} \qquad\qquad (\text{prop. of } \underset{B'\to B''}{P}).$$

Thus $C[\mathbf{v}]_B = [\mathbf{v}]_{B''}$ for all $\mathbf{v} \in V$, which is the defining property of $\underset{B\to B''}{P}$. By *uniqueness* it follows that $C = \underset{B\to B''}{P}$.

(5) True or false. If true, provide a proof; if false, give an explicit counterexample.
   (a) If B is a basis for an n-dimensional space V, then $\underset{B\to B}{P} = I_n$.
   (b) If $\underset{B\to B'}{P}$ is diagonal, then each basis element of B is a scalar multiple of some basis element of B'.
   (c) If each basis element of B' is a scalar multiple of *some* basis element of B, then $\underset{B\to B'}{P}$ is diagonal.
   (d) The coordinate vector of a vector $\mathbf{x}$ in $\mathbb{R}^n$ relative to the standard basis for $\mathbb{R}^n$ is $\mathbf{x}$.

Solution 5:

(a) True. We show this in the lecture notes.
(b) True. Since $\underset{B\to B'}{P}$ is diagonal, so is its inverse $\underset{B\to B'}{P}^{-1} = \underset{B'\to B}{P}$. The j-th column of

this matrix is, using the change of basis formula, just $[\mathbf{v}'_j]_B$, where $\mathbf{v}'_j$ is the j-th element of B'. Since the matrix is diagonal, this column vector is of the form $c_j \mathbf{e}_j$ for some $c_j \in \mathbb{R}$. Unpacking the definition of the coordinate vector, this means simply that $\mathbf{v}_j = c_j \mathbf{e}_j$ for some $c_j \in \mathbb{R}$. So in fact the j-th element of B' is a scalar multiple of the j-th element of B (not just any element).

(c) False. As the color in the previous solution emphasizes, to be diagonal we need the j-th element of B' to be a scalar multiple of the j-th element of B. Consider the bases $B = \{(1,0),(0,1)\}$ and $B' = \{(0,1),(1,0)\}$ of $\mathbb{R}^2$. Then the elements of B' are scalar multiples of elements of B, but $\underset{B \to B'}{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is not diagonal.

(d) True. Let B be the standard basis, and let $\mathbf{e}_i$ be the i-th element B: i.e., $\mathbf{e}_i$ has a 1 in the i-th entry, and 0's elsewhere. Given any $\mathbf{v} = (a_1, a_2, \ldots, a_n)$, we have $\mathbf{v} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \cdots + a_n \mathbf{e}_n$, as we have remarked before. By definition, we then have $[\mathbf{v}]_B = (a_1, a_2, \ldots, a_n) = \mathbf{v}$.

(6) Consider the bases $B = \{p_1 = 6 + 3x, p_2 = 10 + 2x\}$ and $B' = \{q_1 = 2, q_2 = 3 + 2x\}$ for P_1.
    (a) Compute the transition matrix from B' to B.
    (b) Compute the transition matrix from B to B'.
    (c) Compute the coordinate vector $[p]_B$, where $p = -4 + x$ and use the change of basis formula theorem to compute $[p]_{B'}$.
    (d) Now compute $[p]_{B'}$ directly and verify that your answer agrees with part (c).

Solution 6:

To compute the transition matrix from B' to B place the bases in a matrix, with the new basis on the left and the old basis on the right, and use row operations to reduce the left side to the identity.

$$\left[\begin{array}{cc|cc} 6 & 10 & 2 & 3 \\ 3 & 2 & 0 & 2 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & -2/9 & 7/9 \\ 0 & 1 & 1/3 & -1/6 \end{array}\right]$$

Thus transition matrix from B' to B is

$$\begin{bmatrix} -2/9 & 7/9 \\ 1/3 & -1/6 \end{bmatrix}$$

Similarly, we can find the transition matrix from B to B'

$$\left[\begin{array}{cc|cc} 2 & 3 & 6 & 10 \\ 0 & 2 & 3 & 2 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 3/4 & 7/2 \\ 0 & 1 & 3/2 & 1 \end{array}\right]$$

Thus transition matrix from B to B' is

$$\begin{bmatrix} 3/4 & 7/2 \\ 3/2 & 1 \end{bmatrix}$$

Computing $[p]_B$ directly we can see

$$\left[\begin{array}{cc|c} 6 & 10 & -4 \\ 3 & 2 & 1 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -1 \end{array}\right]$$

Thus

$$[p]_B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Using formula (12), we can find $[p]_{B'}$

$$[p]_{B'} = \begin{bmatrix} 3/4 & 7/2 \\ 3/2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -11/4 \\ 1/2 \end{bmatrix}$$

Finally if we compute $[p]_{B'}$ directly we will arrive at the same answer.

$$\left[\begin{array}{cc|c} 2 & 3 & -4 \\ 0 & 2 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & -11/4 \\ 0 & 1 & 1/2 \end{array}\right]$$

Thus

$$[p]_{B'} = \begin{bmatrix} -11/4 \\ 1/2 \end{bmatrix}$$

(7) Let B be the standard basis for $\mathbb{R}^3$, and let $B' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ where $\mathbf{v}_1 = (1, 2, 1)$, $\mathbf{v}_2 = (2, 5, 0)$, and $\mathbf{v}_3 = (3, 3, 8)$.
  (a) Compute $\underset{B' \to B}{P}$.
  (b) Now compute $\underset{B \to B'}{P}$.
  (c) Confirm these two matrices are inverses of one another.
  (d) Let $\mathbf{w} = (5, -3, 1)$. Compute $[\mathbf{w}]_{B'}$ directly, then compute $[\mathbf{w}]_{B'}$ using $[\mathbf{w}]_B$ and the relevant change of basis matrix.
  (e) Let $\mathbf{w} = (3, -5, 0)$. Compute $[\mathbf{w}]_B$ by inspection, then compute $[\mathbf{w}]_{B'}$ using $[\mathbf{w}]_B$ and the relevant transition matrix.

Solution 7:

(a) By inspection

$$\underset{B' \to B}{P} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$$

(b) We now compute $\underset{B \to B'}{P} = (\underset{B' \to B}{P})^{-1}$. Using formula (14)

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

we can see that

$$\underset{B \to B'}{P} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

(c) It is easy to see that these are inverses

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix} \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(d) Calculating $[\mathbf{w}]_{B'}$ directly

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 5 \\ 2 & 5 & 3 & -3 \\ 1 & 0 & 8 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -239 \\ 0 & 1 & 0 & 77 \\ 0 & 0 & 1 & 30 \end{array}\right]$$

Thus

$$[\mathbf{w}]_{B'} = \begin{bmatrix} -239 \\ 77 \\ 30 \end{bmatrix}$$

Now using change of basis

$$[\mathbf{w}]_{B'} = \underset{B \to B'}{P}[\mathbf{w}]_B = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -239 \\ 77 \\ 30 \end{bmatrix}$$

(e) By inspection we have $[(-3, 5, 0)]_B = (-3, 5, 0)$. Then, using our change of basis matrix, we have

$$[\mathbf{w}]_{B'} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ -5 \\ 0 \end{bmatrix} = \begin{bmatrix} -200 \\ 64 \\ 25 \end{bmatrix}$$

(8) The matrix
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$
is the change of basis matrix from what basis B to the basis $\{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ in $\mathbb{R}^3$?

Solution 8:

Let $B = \{(a, b, c), (d, e, f), (h, g, i)\}$ Then
$$\begin{aligned} (a, b, c) &= 1(1,1,1) + 0(1,1,0) + 0(1,0,0) = (1,1,1) \\ (d, e, f) &= 0(1,1,1) + 3(1,1,0) + 1(1,0,0) = (4,3,0) \\ (g, h, i) &= 0(1,1,1) + 2(1,1,0) + 1(1,0,0) = (3,2,0) \end{aligned}$$
Thus $B = \{(1,1,1), (4,3,0), (3,2,0)\}$. We can even double check this by using formula (14)

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 4 & 3 \\ 1 & 1 & 0 & 1 & 3 & 2 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 3 & 2 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right]$$

(9) Define $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ as

$$T(x_1, x_2, x_3) = (x_1 + 2x_2 - x_3, -x_2, x_1 + 7x_3)$$

Let B be the standard basis of $\mathbb{R}^3$, and let B' be the nonstandard basis
$$B' = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}.$$
First compute $[T]_B$, then compute $[T]_{B'}$ using the change of basis formula for transformations.

Solution 9:

Start by using T on each vector in the standard basis
$$\begin{aligned} T((1,0,0)) &= (1,0,1) \\ T((0,1,0)) &= (2,-1,0) \\ T((0,0,1)) &= (-1,0,7) \end{aligned}$$
Then we can take these as the columns of $[T]_B$.

$$[T]_B = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 7 \end{bmatrix}$$

According to the change of basis formula, we have $[T]_{B'} = P_{B \to B'}[T]_B P_{B' \to B}$. Recall that in this very specific context ($V = \mathbb{R}^n$, B the standard basis) the change of basis matrix $P_{B' \to B}$ is obtained by simply placing the elements of B' as the columns of a vector:

$$P_{B' \to B} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Then we have

$$\underset{B\to B'}{P} = (\underset{B'\to B}{P})^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

Now we can use the change of basis formula to compute

$$[T]_{B'} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 7 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 3 \\ -1 & -2 & -9 \\ 1 & 1 & 8 \end{bmatrix}$$

(10) Define $T\colon P_1 \to P_1$ as

$$T(a_0 + a_1 x) = a_0 + a_1(x+1) = (a_0 + a_1) + a_1 x.$$

The following are two different bases for P_1: $B = \{6 + 3x, 10 + 2x\}$; $B' = \{2, 3 + 2x\}$.
First compute $[T]_B$, then use the change of basis formula for transformations to compute $[T]_{B'}$.

Solution 10:

In a similar fashion to the last problem

$$\begin{aligned}
[T(6+3x)]_B &= [9+3x]_B = (2/3, 1/2) \\
[T(10+2x)]_B &= [12+2x]_B = (-2/9, 4/3)
\end{aligned}$$

Thus

$$[T]_B = \begin{bmatrix} 2/3 & -2/9 \\ 1/2 & 4/3 \end{bmatrix}$$

We compute $P_{B\to B'}$ using the recipe

$$P_{B\to B'} = \begin{bmatrix} | & | \\ [6+3x]_{B'} & [10+2x]_{B'} \\ | & | \end{bmatrix} = \begin{bmatrix} 3/4 & 7/2 \\ 3/2 & 1 \end{bmatrix}$$

Then

$$P_{B'\to B} = (P_{B\to B'})^{-1} = \begin{bmatrix} -2/9 & 7/9 \\ 1/3 & -1/6 \end{bmatrix}$$

Then we have

$$[T]_{B'} = \begin{bmatrix} 3/4 & 7/2 \\ 3/2 & 1 \end{bmatrix}\begin{bmatrix} 2/3 & -2/9 \\ 1/2 & 4/3 \end{bmatrix}\begin{bmatrix} -2/9 & 7/9 \\ 1/3 & -1/6 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

(11) Let $T\colon V \to W$ be a linear transformation between two finite-dimensional vector spaces. Suppose B_1, B_2 are two different bases of V, and B'_1, B'_2 are two different bases of W.

State and prove a matrix equation relating $[T]^{B'_2}_{B_2}$ in terms of $[T]^{B'_1}_{B_1}$, $\underset{B_1\to B_2}{P}$, $\underset{B'_1\to B'_2}{P}$, and/or the inverses of these change of basis matrices.

To prove your equality, you should show that the LHS and RHS matrices both satisfy the defining property of $[T]^{B'_2}_{B_2}$.

Solution 11:

I claim

$$[T]^{B'_2}_{B_2} = \underset{B'_1\to B'_2}{P}\, [T]^{B'_1}_{B_1}\, \underset{B_2\to B_1}{P} = \underset{B'_1\to B'_2}{P}\, [T]^{B'_1}_{B_1}\, (\underset{B_1\to B_2}{P})^{-1}$$

112

To prove the claim, I need only show the matrix on the RHS, call it C, satisfies the defining property of $[T]^{B_2'}_{B_2}$: namely,

$$C[\mathbf{v}]_{B_2} = [T(\mathbf{v})]_{B_2'} \text{ for all } \mathbf{v} \in V.$$

To this end we compute:

$$\underset{B_1'\to B_2'}{P}\ [T]^{B_1'}_{B_1}\ \underset{B_2\to B_1}{P}[\mathbf{v}]_{B_2} = \underset{B_1'\to B_2'}{P}\ [T]^{B_1'}_{B_1}[\mathbf{v}]_{B_1} \qquad \text{(convert to } B_1 \text{ coordinates)}$$

$$= \underset{B_1'\to B_2'}{P}[T(\mathbf{v})]_{B_1'} \qquad \text{(defining property of } [T]^{B_1'}_{B_1})$$

$$= [T(\mathbf{v})]_{B_2'} \qquad \text{(convert to } B_2'\text{-coordinates)}$$

(12) Suppose $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation, and that for any vector $\mathbf{v}$ of the form $\mathbf{v} = c(1,1) + d(-1,1)$, we have

$$T(\mathbf{v}) = T\left(c(1,1) + d(-1,1)\right) = (c+d)(1,1) + d(-1,1).$$

Note: this is an example of a *shearing transformation* along the vector $(1,1)$.
(a) Let $B' = \{(1,1), (-1,1)\}$. Compute $[T]_{B'}$.
(b) Use the change of basis formula to compute $A = [T]_B$, where B is the standard basis.
(c) Compute $T(1,2)$.

Solution 12:
(a) We compute $T(1,1) = T\left(1(1,1) + 0(-1,1)\right) = (1+0)(1,1) + 0(-1,1) = (1,1)$. Thus $[T(1,1)]_{B'} = (1,0)$.
 Similarly, $T(-1,1) = T\left(0(1,1) + 1(-1,1)\right) = 1(1,1) + 1(-1,1)$. Thus $[T(-1,1)]_{B'} = (1,1)$.

It follows that $[T]_{B'} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

(b) As usual we see by inspection that $\underset{B'\to B}{P} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

Using inverses we then compute $\underset{B\to B'}{P} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

Thus

$$[T]_B = \underset{B'\to B}{P}[T]_{B'}\underset{B\to B'}{P} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}.$$

(c) We have

$$T(1,2) = A\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

(13) For fixed direction vector $\mathbf{d} = (a,b) \neq \mathbf{0}$, let ℓ be the line in $\mathbb{R}^2$ passing through the origin that it defines.
Define $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ to be reflection through ℓ, as defined in Exercise 3 of Section 3.2. Recall that T is a linear transformation.
(a) Construct a basis B' of $\mathbb{R}^2$ which pays attention to the geometry involved in the definition of T, and compute $[T]_{B'}$.
 (If you picked your basis appropriately, $[T]_{B'}$ should be diagonal.)
(b) Now use the change of basis formula to compute the A such that $T = T_A$.

Solution 13:

(a) I pick my basis to be $B' = \{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1 = (a, b) = \mathbf{d}$, the direction vector of ℓ, and $\mathbf{v}_2 = (-b, a)$, a vector pointing orthogonally to ℓ.

First observe that $T(\mathbf{v}_1) = \mathbf{v}_1$. This is because $\mathbf{v}_1$ lies on ℓ, and the reflection of a point that lies on ℓ is equal to itself.

Next, observe that $T(\mathbf{v}_2) = -\mathbf{v}_2$. This is because $\mathbf{v}_2$ points orthogonally to ℓ, so its orthogonal projection onto ℓ is just $\mathbf{0}$, and its reflection is thus $-\mathbf{v}_2$. (As usual, I conflate vectors with points in this discussion.) See picture below.
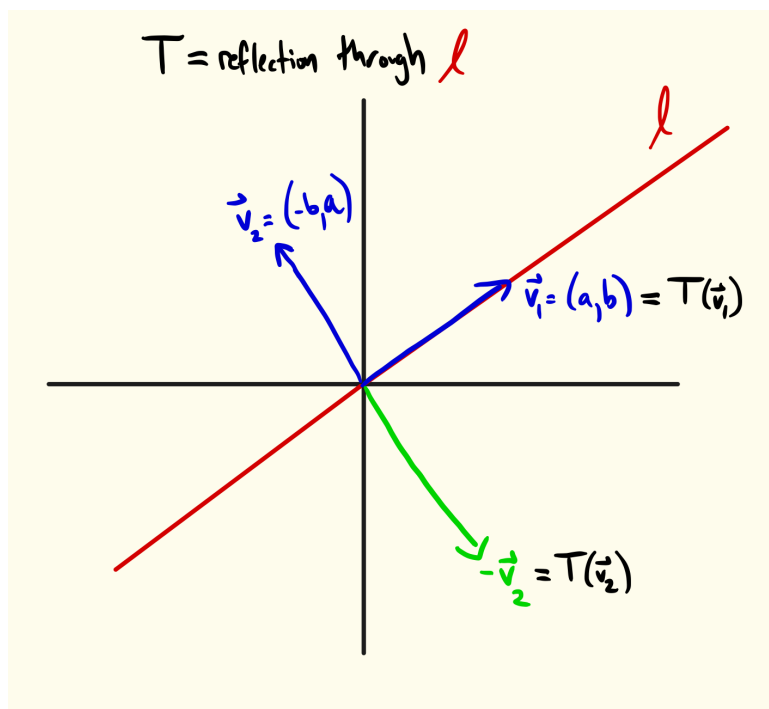
Then we have

$$[T]_{B'} = \begin{bmatrix} | & | \\ [T(\mathbf{v}_1)]_{B'} & [T(\mathbf{v}_2)]_{B'} \\ | & | \end{bmatrix} = \begin{bmatrix} | & | \\ [\mathbf{v}_1]_{B'} & [-\mathbf{v}_2]_{B'} \\ | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

(b) To compute the A such that $T = T_A$, note that $A = [T]_B$ where B is the standard basis of $\mathbb{R}^2$. We can then use the change of basis formula:

$$A = [T]_B = \underset{B' \to B}{P} [T]_{B'} \underset{B \to B'}{P}$$

We compute

$$\underset{B' \to B}{P} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad \underset{B \to B'}{P} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}^{-1} = \frac{1}{a^2 + b^2} \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

Thus

$$A = [T]_B = \frac{1}{a^2 + b^2} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a & b \\ -b & a \end{bmatrix} = \frac{1}{a^2 + b^2} \begin{bmatrix} a^2 - b^2 & 2ab \\ 2ab & b^2 - a^2 \end{bmatrix}.$$

Recall that earlier we derived the formula

$$A = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix},$$

where θ is the angle that the line ℓ makes with the x-axis. Is our new formula consistent with this? Observe that the vector (a, b) has polar coordinates (r, θ) where

$$r = \sqrt{a^2 + b^2}, \quad \cos\theta = \frac{a}{\sqrt{a^2 + b^2}}, \quad \sin\theta = \frac{b}{\sqrt{a^2 + b^2}}.$$

It follows that

$$\frac{1}{a^2 + b^2}\begin{bmatrix} a^2 - b^2 & 2ab \\ 2ab & b^2 - a^2 \end{bmatrix} = \begin{bmatrix} \cos^2\theta - \sin^2\theta & 2\cos\theta\sin\theta \\ 2\cos\theta\sin\theta & \sin^2\theta - \cos^2\theta \end{bmatrix} = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} \checkmark!$$

(14) *Reflection through the plane $\mathcal{P}$: $x + y + z = 0$.* Define $T: \mathbb{R}^3 \to \mathbb{R}^3$ to be reflection through the plane $\mathcal{P}$: $x + y + z = 0$. See the next exercise for a precise definition of this operation.
   (a) Construct an explicit basis of the form $B' = \{\mathbf{v}_1, \mathbf{v}_2, (1, 1, 1)\}$ where $\mathbf{v}_1, \mathbf{v}_2$ are elements of $\mathcal{P}$.
   (b) Compute $[T]_{B'}$.
   (c) Find the A such that $T = T_A$ using your result in (b) and a change of basis formula.
   (d) Compute $T(1, 2, 3)$.

Solution 14:
This exercise is really just a special case of the next one. I give more details in the solution there.
(a) Set $B' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{n}\}$, where $\mathbf{v}_1 = (1, -1, 0), \mathbf{v}_2 = (0, 1, -1), \mathbf{n} = (1, 1, 1)$. As is explained in more detail in the next exercise, we have $T(\mathbf{v}_1) = \mathbf{v}_1, T(\mathbf{v}_2) = \mathbf{v}_2$, and $T(\mathbf{n}) = -\mathbf{n}$. It follows that

$$[T]_{B'} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

(b) To compute $A = [T]_B$, we use the change of basis formula

$$A = [T]_B = \underset{B' \to B}{P}[T]_{B'}\underset{B \to B'}{P}$$

$$= \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}^{-1}$$

$$= \frac{1}{3}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}.$$

(c) We have

$$T(1, 2, 3) = A\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \\ -1 \end{bmatrix}.$$

(15) *Reflection through a general plane in $\mathbb{R}^3$.* Fix $\mathbf{n} = (a, b, c) \neq \mathbf{0}$. Let W be the plane passing through the origin with normal vector $\mathbf{n} = (a, b, c)$: i.e., $W : ax + by + cz = 0$. Reflection through W is the map $T: \mathbb{R}^3 \to \mathbb{R}^3$ defined as follows:
   given $P = (x, y, z)$ let ℓ be the line passing through P and perpendicular to W; let Q be the intersection of this line with W; then the reflection $T(P)$ is the unique point $P' = (x', y', z')$ on ℓ such that the distance from P to P' is twice the distance from P to Q.

Intuitively, $T(P) = P'$ is the point "on the other side of the plane" from P and an equal distance away.

You may take for granted that T is a linear transformation. Also, observe that if P lies in W, then the definition implies $T(P) = P$: P is its own reflection.

Compute the matrix A such that $T = T_A$ as follows:

(a) First compute $A' = [T]_{B'}$, where $B' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for $\mathbb{R}^3$ satisfying $\mathbf{v}_1, \mathbf{v}_2 \in W$, $\mathbf{v}_3 = (a, b, c)$.

(b) Now use the change of basis formula to compute $A = [T]_B$, where B is the standard basis.

Solution 15:

(a) If we have a basis B' as described, then $[T]_{B'} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. This is because T fixes any vector in W (reflection of point in plane is itself) and maps $\mathbf{n}$ to $-\mathbf{n}$. Once we have found the basis B', letting B be the standard basis, we have

$$A = [T]_B = \underset{B' \to B}{P} [T]_{B'} \underset{B \to B'}{P}.$$

Furthermore, we know

$$\underset{B' \to B}{P} = \begin{bmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{bmatrix}, \quad \text{and} \quad \underset{B \to B'}{P} = (\underset{B' \to B}{P})^{-1}.$$

So we are essentially done, except for some slightly unpleasant matrix arithmetic, after we find the basis B'. Let's do this now.

To make our life easier, I assume that *at least two* of the elements a, b, c are nonzero. The excluded case, where exactly one of the elements a, b, and c is nonzero, would just be reflection through one of the coordinate planes, which is very easy to compute. Furthermore, the matrix formula we derive is valid even in these excluded cases, as you can check below!

So assume that at least two of the elements a, b, c are nonzero. Then $\mathbf{v}_1 = (b, -a, 0)$ and $\mathbf{v}_2 = (0, c, -b)$ make up a basis of W, as is easily checked. Setting $\mathbf{v}_3 = \mathbf{n}$, we then have

$$\underset{B' \to B}{P} = \begin{bmatrix} b & 0 & a \\ -a & c & b \\ 0 & -b & c \end{bmatrix}$$

and hence

$$A = \begin{bmatrix} b & 0 & a \\ -a & c & b \\ 0 & -b & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \left(\begin{bmatrix} b & 0 & a \\ -a & c & b \\ 0 & -b & c \end{bmatrix} \right)^{-1}$$

$$= \frac{1}{a^2 + b^2 + c^2} \begin{bmatrix} -a^2 + b^2 + c^2 & -2ab & -2ac \\ -2ab & a^2 - b^2 + c^2 & -2bc \\ -2ac & -2bc & a^2 + b^2 - c^2 \end{bmatrix}$$

Admittedly, I used technology to compute the inverse and the matrix multiplication in the last line.

116

(1) Compute the characteristic polynomial, eigenvalues, and eigenspace bases for each of the following matrices.

(a) $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & -3 \\ 0 & 2 \end{bmatrix}$

(c) $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

(d) $A = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}$

Solution 1:

(a) The characteristic polynomial is

$$\det(tI - A) = \begin{vmatrix} t-2 & -1 \\ -1 & t-2 \end{vmatrix} = (t-2)^2 - 1 = t^2 - 4t + 3 = (t-3)(t-1) = 0$$

Thus the eigenvalues are $\lambda = 3$ and $\lambda = 1$.

For W_3 we see that $3I - A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ has rank 1, and hence nullity 1. By inspection we see that $\mathbf{v}_1 = (1,1)$ comprises a basis for W_3.

For $\lambda = 1$, we have $I - A = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$. Again, by inspection we see that $I - A$ has rank 1, hence nullity 1, and that $\mathbf{v}_2 = (1, -1)$ forms a basis.

(b)

$$\det(tI - A) = \begin{vmatrix} t-2 & 3 \\ 0 & t-2 \end{vmatrix} = t^2 - 4t + 4 = 0$$

Which factors as $(t-2)^2 = 0$. Thus we have only one eigenvalue, $\lambda = 2$.

We have $2I - A = \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$, a matrix of rank 1, and hence nullity 1. We see that $\mathbf{v}_1 = (1, 0)$ forms a basis for W_2.

(c)

$$\det(tI - A) = \begin{vmatrix} t-2 & 0 \\ 0 & t-2 \end{vmatrix} = t^2 - 4t + 4 = 0$$

Which factors as $(t-2)^2 = 0$. Thus we have only one eigenvalue, $t = 2$.

We have $2I - A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ Then clearly $W_2 = \text{null}(2I - A) = \mathbb{R}^2$, and $\{(1, 0), (0, 1)\}$ is a basis of W_2: i.e., all nonzero vectors of $\mathbb{R}^2$ are eigenvectors of A!

(d) The characteristic polynomial is

$$\det(tI - A) = \begin{vmatrix} t-1 & - \\ 2 & t+1 \end{vmatrix} = t^2 + 3 = 0$$

Which does not have any real eigenvalues.

(2) Compute the characteristic polynomial, eigenvalues, and eigenspace bases for

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{bmatrix}$$

Solution 2:

The characteristic polynomial is

$$\det(tI - A) = \begin{vmatrix} t-1 & 0 & 2 \\ 0 & t & 0 \\ 2 & 0 & t-4 \end{vmatrix} = t^3 - 5t = 0$$

This factors as $t^2(t-5) = 0$. So the eigenvalues are $\lambda = 0$ and $\lambda = 5$. To find the basis for $\lambda = 0$ we use

$$\begin{bmatrix} -1 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus $x_1 = 2x_3$, with the free variables $x_2 = t$ and $x_3 = r$.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2r \\ t \\ r \end{bmatrix} = r \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Thus the basis when $\lambda = 0$ is

$$\left\{ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Now for $\lambda = 5$ we use

$$\begin{bmatrix} 4 & 0 & 2 \\ 0 & 5 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus $x_1 = -1/2x_3$, $x_2 = 0$ and the free variable $x_3 = r$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1/2r \\ 0 \\ r \end{bmatrix} = r \begin{bmatrix} -1/2 \\ 0 \\ 1 \end{bmatrix}$$

Thus the basis when $\lambda = 5$ is

$$\left\{ \begin{bmatrix} -1/2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

(3) Compute the characteristic polynomial, eigenvalues, and eigenspace bases for

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Solution 3:

The characteristic polynomial is

$$\det(tI - A) = \begin{vmatrix} t & -1 & -1 \\ -1 & t & -1 \\ -1 & -1 & t \end{vmatrix} = t^3 - 3t - 2 = 0$$

This factors as $(t+1)^2(t-2) = 0$. So the eigenvalues are $\lambda = -1$ and $\lambda = 2$. To find the basis for $\lambda = -1$ we use

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus $x_1 = -x_2 - x_3$, with the free variables $x_2 = t$ and $x_3 = r$.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -t - r \\ t \\ r \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + r \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Thus the basis when $\lambda = -1$ is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} \right\}$$

Now for $\lambda = 2$ we use

$$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus $x_1 = x_3$, $x_2 = x_3$ and the free variable $x_3 = r$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} r \\ r \\ r \end{bmatrix} = r \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Thus the basis when $\lambda = 2$ is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

(4) Compute the characteristic polynomial, eigenvalues, and eigenspace bases for

$$A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$$

Solution 4:

The characteristic polynomial is

$$\det(tI - A) = \begin{vmatrix} t - 1 & 3 & -3 \\ -3 & t + 5 & -3 \\ -6 & 6 & t - 4 \end{vmatrix} = t^3 - 12t - 16 = 0$$

This factors as $(t + 2)^2(t - 4) = 0$. So the eigenvalues are $\lambda = -2$ and $\lambda = 4$. To find the basis for $\lambda = -2$ we use

$$\begin{bmatrix} -3 & 3 & -3 \\ -3 & 3 & -3 \\ -6 & 6 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus $x_1 = x_2 - x_3$, with the free variables $x_2 = t$ and $x_3 = r$.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t - r \\ t \\ r \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + r \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Thus the basis when $\lambda = -2$ is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Now for $\lambda = 4$ we use
$$\begin{bmatrix} 3 & 3 & -3 \\ -3 & 9 & -3 \\ -6 & 6 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus $x_1 = 1/2x_3$, $x_2 = 1/2x_3$ and the free variable $x_3 = r$
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1/2r \\ 1/2r \\ r \end{bmatrix} = r \begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

Thus the basis when $\lambda = 4$ is
$$\left\{ \begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix} \right\}$$

(5) Find the characteristic polynomial by inspection.
$$\begin{bmatrix} 9 & -8 & 6 & 3 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 7 \end{bmatrix}$$

Solution 5:

Since the given matrix is upper triangular, so is $tI - A$. Since determinants of triangular matrices are just the product of their diagonal entries, we have
$$p(t) = \det(tI - A) = (t - 9)(t + 1)(t - 3)(t - 7).$$

Similar reasoning shows, in general, that the eigenvalues of a triangular matrix (upper or lower or diagonal) are simply the entries along the diagonal.

(6) Let $A = \begin{bmatrix} -5 & 0 & 3 \\ -6 & 1 & 3 \\ -6 & 0 & 4 \end{bmatrix}$. Compute $p(t)$ and find bases for all eigenspaces of A.

Solution 6:

The characteristic polynomial of A is $p(t) = t^3 - 3t + 2 = (t - 1)^2(t + 2)$.
We compute:
$$W_1 = \text{null}(I - A) = \text{span}(\{(1, 0, 2), (0, 1, 0)\})$$
$$W_{-2} = \text{null}(-2I - A) = \text{span}(\{(1, 1, 1)\})$$

(7) The matrix $B = \begin{bmatrix} -2 & -3 & 3 \\ -3 & -3 & 4 \\ -3 & -4 & 5 \end{bmatrix}$ has the same characteristic polynomial of A in Exercise 6.
Find bases for all eigenspaces of B.

Solution 7:

Since B also has characteristic polynomial $p(t) = (t - 1)^2(t + 2)$, it has only two eigenspaces.

120

We compute:

$$W_1 = \text{null}(I - B) = \text{span}(\{(0, 1, 1)\})$$
$$W_{-2} = \text{null}(-2I - B) = \text{span}(\{(1, 1, 1)\})$$

The two problems taken together show that the characteristic polynomial is not the whole story! The matrix A has more linearly independent eigenvectors than B.

(8) Use induction to prove that if λ is an eigenvalue of A, then λ^k is an eigenvalue of A^k for all positive integers k. Make sure to invoke the definition of eigenvalue in your proof.
Solution 8:

Take $\mathbf{v} \neq \mathbf{0}$ with $A\mathbf{v} = \lambda\mathbf{v}$. We show by induction that $A^k\mathbf{v} = \lambda^k\mathbf{v}$, and thus that λ^k is an eigenvalue of A^k.
The base case, $k = 1$, is just the statement that $\mathbf{v}$ is an eigenvector.
Induction step: assume $A^k\mathbf{v} = \lambda^k\mathbf{v}$. Then

$$
\begin{aligned}
A^{k+1}\mathbf{v} &= AA^k\mathbf{v} \\
&= A(\lambda^k\mathbf{v}) && \text{(induction hypo.)} \\
&= \lambda^k A\mathbf{v} && (\lambda \text{ is a scalar}) \\
&= \lambda^k \lambda\mathbf{v} && (\mathbf{v} \text{ is an eigenvector}) \\
&= \lambda^{k+1}\mathbf{v}.
\end{aligned}
$$

If we didn't know proof by induction, our "proof" would look something like

$$
\begin{aligned}
A^k\mathbf{v} &= \underbrace{A \cdot A \cdots A}_{k \text{ times}}\mathbf{v} \\
&= \underbrace{A \cdot A \cdots A}_{k-1 \text{ times}}(A\mathbf{v}) \\
&= \underbrace{A \cdot A \cdots A}_{k-1 \text{ times}}(\lambda\mathbf{v}) && (\text{since } A\mathbf{v} = \lambda\mathbf{v}) \\
&= \lambda \underbrace{A \cdot A \cdots A}_{k-1 \text{ times}}(\mathbf{v}) && (\text{since } \lambda \text{ is a scalar}) \\
&= \lambda^2 \underbrace{A \cdot A \cdots A}_{k-2 \text{ times}}(\mathbf{v}) && (\text{by same logic}) \\
&\ \ \vdots \\
&= \lambda^{k-1} A\mathbf{v} \\
&= \lambda^k\mathbf{v}
\end{aligned}
$$

You can think of proof by induction as a way of replacing all these ... and ⋮ with a rigorous argument!

(9) Prove: A is invertible if and only if $\lambda = 0$ is *not* an eigenvalue of A.
Solution 9:

There are many different ways of proving this.
Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A. Recall that $\det(A) = \lambda_1\lambda_2 \cdots \lambda_n$.

Then we have

$$A \text{ invertible} \Leftrightarrow \det(A) \neq 0$$
$$\Leftrightarrow \lambda_1 \lambda_2 \cdots \lambda_n \neq 0$$
$$\Leftrightarrow \lambda_i \neq 0 \text{ for all } i$$
$$\Leftrightarrow 0 \text{ is not an eigenvalue of } A.$$

(10) Let $B = \{\cos(x), \sin(x), \cosh(x), \sinh(x)\}$, let $V = \mathrm{span}(B)$, and define $T \colon V \to V$ by $T(f) = f'$.

You may take as a fact that the set B is linearly independent, hence a basis.

Recall: $\cosh(x) = \dfrac{e^x + e^{-x}}{2}$, $\sinh(x) = \dfrac{e^x - e^{-x}}{2}$.

(a) Compute $A = [T]_B$. What does A tell you about the invertibility of T?
(b) Find all eigenvalues of T.
(c) For each eigenvalue λ of T, find a basis for W_λ.
(d) Use your work in (c) to find a new basis B' of V that contains two eigenvectors of T.
(e) Compute $[T]_{B'}$.

Solution 10:

We find that $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$.

Since A is invertible (its rank is clearly 4), it follows that T is invertible.

The characteristic polynomial of A (in factored form) is $p(t) = (t-1)(t+1)(t^2+1)$. Thus the eigenvalues of A (and hence also T) are $\lambda = 1$ and $\lambda = -1$.

We compute bases for the eigenspaces of A:

$$W_1 = \mathrm{null}(I - A) = \mathrm{span}(\{(0,0,1,1)\}$$
$$W_{-1} = \mathrm{null}(-I - A) = \mathrm{span}(\{(0,0,1,-1)\}$$

Lifting these vectors back to V via the $[\ ]_B$ map yields bases for the same eigenspaces of T:

$$W_1 = \mathrm{span}(\{\cosh x + \sinh x\} = \mathrm{span}(\{e^x\})$$
$$W_{-1} = \mathrm{span}(\{\cosh x - \sinh x\} = \mathrm{span}(\{e^{-x}\})$$

We construct the new basis $B' = \{\cos(x), \sin(x), e^x, e^{-x}\}$ (it is easy to see that these four functions are independent), and compute

$$A' = [T]_{B'} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

(11) Suppose A is a 2×2 matrix satisfying $\mathrm{tr}\, A = 2$ and $\det A = -15$. Find the eigenvalues of A.

Solution 11:

Let λ, λ' be the eigenvalues of A (possibly complex, possibly repeated). We have

$$\lambda + \lambda' = \operatorname{tr} A = 2 \qquad\qquad \text{(trace is the sum of the eigenvalues)}$$
$$\lambda\lambda' = \det A = -15 \qquad\qquad \text{(det. is the product of the eigenvalues)}$$

The system implies $\lambda' = 2 - \lambda$, and hence that $\lambda^2 - 2\lambda - 15 = 0$. Thus $\lambda = 5$ or $\lambda = -3$. Thus the eigenvalues of A are 5 and -3.

Alternatively, the characteristic polynomial of A is

$$p(x) = x^2 - \operatorname{tr} Ax + \det A = x^2 - 2x - 15 = (x - 5)(x + 3).$$

Thus the eigenvalues of A are 5 and -3.

(12) Let $T\colon V \to V$ be a linear transformation, and suppose $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.
Prove: $[T]_B$ is a diagonal matrix if and only if $\mathbf{v}_j$ is an eigenvector of T for all $1 \le i \le n$.
Moral. We can represent T by a diagonal matrix if and only if T has "enough" linearly independent eigenvectors to build a basis.
Solution 12:

We prove the two implications separately.

($\Rightarrow$) Suppose the vectors $\mathbf{v}_i$ are all eigenvectors with $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$. Using the recipe for $[T]_B$ we compute

$$[T]_B = \begin{bmatrix} | & | & & | \\ [T(\mathbf{v}_1)]_B & [T(\mathbf{v}_2)] & \cdots & [T(\mathbf{v}_n)]_B \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | \\ [\lambda_1 \mathbf{v}_1]_B & [\lambda_2 \mathbf{v}_2] & \cdots & [\lambda_n \mathbf{v}_n]_B \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

where the last equality holds since $\lambda_1 \mathbf{v}_1 = \lambda_1 \mathbf{v}_1 + 0\mathbf{v}_2 + \ldots, \lambda_2 \mathbf{v}_2 = 0\mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + 0\mathbf{v}_3 + \ldots,$ etc.

($\Leftarrow$) Suppose

$$[T]_B = D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & d_n \end{bmatrix}.$$

I claim $T(\mathbf{v}_i) = d_i \mathbf{v}_i$ for all i, which proves B consists of eigenvectors of T. To do so, first observe that

$$[T(\mathbf{v}_i)]_B = [T]_B [\mathbf{v}_i]_B \qquad \text{(defining property of } [T]_B)$$
$$= D\mathbf{e}_i \qquad \text{(since } [\mathbf{v}_i]_B = (0, 0, \ldots, 1, 0, \ldots, 0) = \mathbf{e}_i)$$
$$= \begin{bmatrix} 0 \\ \vdots \\ d_i \\ \vdots \\ 0 \end{bmatrix}.$$

But we have $[T(\mathbf{v}_i)]_B = (0, 0, \ldots, d_i, 0, \ldots, 0)$ if and only if $T(\mathbf{v}_i) = d_i \mathbf{v}_i$, by definition of the coordinate vector map. This proves each $\mathbf{v}_i$ is an eigenvector with eigenvalue d_i.

4.7: DIAGONALIZATION

(1) Find the eigenvalues of A. For each eigenvalue λ, find the rank of the matrix $\lambda I - A$. Is A diagonalizable? Justify.

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

Solution 1:

We compute $p(t)$:

$$\det(tI - A) = \begin{vmatrix} t-3 & 0 & 0 \\ 0 & t-2 & 0 \\ 0 & -1 & t-2 \end{vmatrix} = (t-3)(t-2)^2 = 0$$

Thus the eigenvalues are $\lambda = 3$ and $\lambda = 2$.
We compute $\dim W_3 = \text{null}(3I - A)$:

$$3I - A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Since the rank of this matrix is 2, its nullity is 3-2=1. This means $\dim(W_3) = 1$.
Next, we have

$$2I - A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

Once again the nullity is 1, which means $\dim(W_2) = 1$. Since the sum of the dimensions of the eigenspaces is $1 + 1 \neq 3$, we conclude that A is not diagonalizable.

(2) For each matrix A below, find the geometric and algebraic multiplicity of each eigenvalue of the matrix A, and determine whether A is diagonalizable. If A is diagonalizable, then find the matrix P that diagonalizes A, and find $P^{-1}AP$.

(a)
$$A = \begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}$$

(b)
$$A = \begin{bmatrix} 19 & -9 & -6 \\ 25 & -11 & -9 \\ 17 & -9 & -4 \end{bmatrix}$$

(c)
$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

(d)
$$A = \begin{bmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{bmatrix}$$

Solution 2:

(a) Start by finding the eigenvalues of matrix A.

$$\det(tI - A) = \begin{vmatrix} t+1 & -4 & 2 \\ 3 & t-4 & 0 \\ 3 & -1 & t-3 \end{vmatrix} = (t-1)(t-2)(t-3) = 0$$

We have a theorem that says a matrix $A \in M_{nn}$ with n distinct eigenvalues is diagonalizable. Thus A is diagonalizable. To construct the matrix P, we need to come up with bases for each space $W_\lambda = \text{null}(\lambda I - A)$. Each of these has dimension 1, and we easily come up with bases by inspection. This yields the basis $\mathbf{v}_1 = (1,1,1)$, $\mathbf{v}_2 = (2,3,3)$ and $\mathbf{v}_3 = (1,3,4)$.

According to the recipe, we then have

$$P = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 4 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \text{ and } D = P^{-1}AP.$$

(b) Start by finding the eigenvectors in the usual way

$$\det(tI - A) = \begin{vmatrix} t-19 & 9 & 6 \\ -25 & t+11 & 9 \\ -17 & 9 & t+4 \end{vmatrix} = t^3 - 4t^2 + 5t - 2 = (t-1)^2(t-2) = 0$$

For $\lambda = 1$, we can see that the algebraic multiplicity is 2. Consider the matrix

$$\begin{bmatrix} -18 & 9 & 6 \\ -25 & 12 & 9 \\ -17 & 9 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -4/3 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus the rank is 2 and the geometric multiplicity is 1. Since the geometric multiplicity is less than the algebraic multiplicity for one of the eigenspaces, we know immediately that A is not diagonalizable.

However, we continue on nonetheless. For $\lambda = 2$, we can see that the algebraic multiplicity is 1. We know that $1 \leq \dim W_2 \leq 1$, hence we must have $\dim W_2 = 1$. Let's verify this:

$$2I - A = \begin{bmatrix} -17 & 9 & 6 \\ -25 & 13 & 9 \\ -17 & 9 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -3/4 \\ 0 & 1 & -3/4 \\ 0 & 0 & 0 \end{bmatrix},$$

a matrix of rank 2, and hence nullity 1.

Lastly, we see that $\dim W_1 + \dim W_2 = 1 + 1 = 2 \neq 3$. Hence A is not diagonalizable, as we claimed above.

(c)

$$\det(tI - A) = \begin{vmatrix} t & 0 & 0 \\ 0 & t & 0 \\ -3 & 0 & t-1 \end{vmatrix} = t^2(t-1) = 0$$

Thus the eigenvalues are $\lambda = 0$ and $\lambda = 1$. For $\lambda = 0$, notice that the algebraic multiplicity is 2. Now consider the matrix

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -3 & 0 & -1 \end{bmatrix}$$

The rank is 1, thus the geometric multiplicity is 2. For $\lambda = 1$, notice that the algebraic multiplicity is 1. Now consider the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 0 \end{bmatrix}$$

The rank is 2, thus the geometric multiplicity is 1. Since for all λ the geometric multiplicity is equal to the algebraic multiplicity, the initial matrix is diagonalizable. As usual, we build P by first computing bases for W_0 and W_1, and then putting these in as the columns of P. One possible answer is

$$P = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

with corresponding

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(d)

$$\det(tI - A) = \begin{vmatrix} t-5 & 0 & 0 \\ -1 & t-5 & 0 \\ 0 & -1 & t-5 \end{vmatrix} = (t-5)^3 = 0$$

Thus $\lambda = 5$ is the only eigenvalue and its algebraic multiplicity is 3. Now considering when λ is 5, we look at the matrix

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

The rank of this matrix is 2, thus the geometric multiplicity is 1. Since algebraic multiplicity is not equal to the geometric multiplicity, the initial matrix is not diagonalizable.

(3) Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Derive necessary and sufficient conditions for A to be diagonalizable expressed in terms of equalities and inequalities involving a, b, c, d.

Solution 3:

First analyze the diagonalizability of A based on the number of distinct real eigenvalues it has: 0, 1, or 2.

If A has 0 eigenvalues, then it is clearly not diagonalizable, as it has no eigenvectors.

If A has 2 distinct eigenvalues, then it is automatically diagonalizable, by the corollary to the linear independence of eigenvectors theorem.

If A has exactly one eigenvalue λ, then A is diagonalizable if and only if $\dim W_\lambda = 2$ (by the diagonalization theorem) if and only if $\dim \operatorname{null} \lambda I - A = 2$ if and only if $\operatorname{rank} \lambda I - A = 0$ if and only if $\lambda I - A = \underset{2 \times 2}{\mathbf{0}}$ if and only if $A = \lambda I$ is diagonal !!

We conclude that A is diagonalizable if and only if (a) it has two distinct eigenvalues or (b) it has one eigenvalue and is diagonal. We can relate these conditions to a, b, c, d by looking at the characteristic polynomial $p(x)$.

The characteristic polynomial of $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $p(x) = x^2 - (a+d)x + (ad - bc)$. Its roots are given by

$$x = \frac{1}{2}(a + d \pm \sqrt{a^2 + 2ad + d^2 - 4ad + 4bc}) = \frac{1}{2}(a + d \pm \sqrt{(a-d)^2 + 4bc})$$

We see that (a) A has two real eigenvalues iff $(a - d)^2 > -4bc$, and that (b) A is diagonal with one distinct eigenvalue iff $a = d$ and $b = c = 0$.

We conclude:

$$A \text{ is diagonalizable} \iff (a - d)^2 > -4bc \text{ or } a - d = b = c = 0.$$

(4) Prove all statements of the properties of similarity theorem, stated below.

Properties of similarity theorem. Suppose A is similar to B: i.e., there is an invertible matrix P such that $B = P^{-1}AP$. Then:
 (a) B is similar to A. (*Similarity is symmetric.*)
 (b) A and B have the same trace and determinant.
 (c) A and B have the same rank nullity and rank.
 (d) A and B have the same characteristic polynomial.
 (e) A and B have the same eigenvalues.
 (f) Given any $\lambda \in \mathbb{R}$, let W_λ be the corresponding eigenspace for A, and W'_λ the corresponding eigenspace for B. Then $\dim W_\lambda = \dim W'_\lambda$.

Solution 4:

(a) If $B = P^{-1}AP$, then $A = PBP^{-1} = Q^{-1}BQ$, where $Q = P^{-1}$.

(d) This was proven in the lecture notes.

(b)& (e) Follow directly from (d) and our theorem about characteristic polynomials. Indeed, let $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be the characteristic polynomial of A and B. Then $\operatorname{tr} A = \operatorname{tr} B = -a_{n-1}$, $\det A = \det B = (-1)^n a_0$, and the eigenvalues of both A and B are precisely the roots of $p(x)$. (c) I claim $\mathbf{x}' \in \operatorname{null} B$ if and only if $\mathbf{x}' = P^{-1}\mathbf{x}$ for some $\mathbf{x} \in \operatorname{null} B$. To see this, first observe that if $\mathbf{x} \in \operatorname{null} A$, then

$$B\mathbf{x}' = B(P^{-1}\mathbf{x}) = P^{-1}AP(P^{-1}\mathbf{x}) = P^{-1}A\mathbf{x} = \mathbf{0}.$$

Now suppose $\mathbf{x}' \in \text{null}\, B$. A similar argument (using $A = PBP^{-1}$) shows that $\mathbf{x} = P\mathbf{x}' \in \text{null}\, A$; and we have $\mathbf{x}' = P\mathbf{x}$.

From our claim it follows that P^{-1} defines an invertible linear transformation (with inverse P) from null A to null B. In fancy terms this means null A and null B are isomorphic, and hence have the same dimension. More directly, if $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is a basis for null A, then you can show that $\{P^{-1}\mathbf{v}_1, P^{-1}\mathbf{v}_2, \ldots, P^{-1}\mathbf{v}_r\}$ is a basis for null B.

Either way, we see that nullity $A = \dim \text{null}\, A = \dim \text{null}\, B = \text{nullity}\, B$. The rank-nullity theorem then implies that rank $A = \text{rank}\, B$.

(f) An argument exactly like the the one above shows that the map $\mathbf{x} \mapsto P^{-1}\mathbf{x}$ defines an invertible linear transformation from W_λ to W_λ', and that P^{-1} maps any basis of W_λ to a basis of W_λ'. Thus $\dim W_\lambda = \dim W_\lambda'$.

(5) Similar matrices have the same rank. Show the converse is false by showing the matrices
$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$
have the same rank, but are not similar.
Solution 5:

Notice that $\text{rank}(A) = \text{rank}(B) = 1$. Also notice that $\text{tr}(A) = 1 \neq 0 = \text{tr}(B)$. Since the traces are different, the matrices are not similar.

(6) Similar matrices have the same eigenvalues. Show the converse is false by showing the matrices
$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
have the same eigenvalues, but are not similar.
Solution 6:

Notice both A and B are triangular, thus $\det(tI - A) = (t - 1)^2 = \det(tI - B)$. Thus $\lambda = 1$ is the only eigenvalue for both A and B.

However, as one can easily compute, for A the eigenspace W_1 has dimension 1, whereas for B the eigenspace W_1 has dimension 2. Thus the two cannot be similar!

(7) Prove that if $B = P^{-1}AP$, and $\mathbf{v}$ is an eigenvector of B corresponding to the eigenvalue λ, the $P\mathbf{v}$ is the eigenvector of A corresponding to λ.
Solution 7:

We assume that $\mathbf{v}$ is an eigenvector of B, and must show that $P\mathbf{v}$ is an eigenvector of A. Thus by definition of eigenvector we assume
(i) $B\mathbf{v} = \lambda\mathbf{v}$, and
(ii) $\mathbf{v} \neq \mathbf{0}$.
But then

$$
\begin{aligned}
A(P\mathbf{v}) &= AP\mathbf{v} & \\
&= PB\mathbf{v} & \text{(since } B = P^{-1}AP) \\
&= P(\lambda\mathbf{v}) & \text{(since } B\mathbf{v} = \lambda\mathbf{v}) \\
&= \lambda P\mathbf{v}) & \text{(since } \lambda \text{ is a scalar).}
\end{aligned}
$$

Thus $A(P\mathbf{v}) = \lambda P(\mathbf{v})$.

To show $P\mathbf{v}$ is an eigenvector, it remains only to show that $P\mathbf{v} \neq \mathbf{0}$. But since P is invertible, and since $\mathbf{v} \neq \mathbf{0}$, it follows that $P\mathbf{v} \neq \mathbf{0}$.

(8) Let A be an n by n matrix, and let $q(A)$ be the matrix

$$q(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I_n$$

Prove that if $B = P^{-1}AP$, then $q(B) = P^{-1}q(A)P$.
Solution 8:

First we need to show (by induction) that $B^n = P^{-1}A^nP$ for all n. The base case, when $n = 1$ is give by the hypothesis. Assume that equality holds when $n \leq k$. We need to show that equality holds when $n = k + 1$.

$$
\begin{aligned}
B^{k+1} &= BB^k \\
&= (P^{-1}AP)B^k \\
&= (P^{-1}AP)(P^{-1}A^kP) \\
&= P^{-1}A(PP^{-1})A^kP \\
&= P^{-1}A^{k+1}P
\end{aligned}
$$

Thus $B^n = P^{-1}A^nP$ holds for all n. Now we can show the desired result.

$$
\begin{aligned}
q(B) &= a_n B^n + a_{n-1} B^{n-1} + \cdots + a_1 B + a_0 I_n \\
&= a_n(P^{-1}A^nP) + a_{n-1}(P^{-1}A^{n-1}P) + \cdots + a_1(P^{-1}AP) + a_0 I_n \\
&= P^{-1}(a_n A^n P) + (a_{n-1}A^{n-1}P) + \cdots + (a_1 AP) + a_0 I_n) \\
&= P^{-1}(a_n A^n) + (a_{n-1}A^{n-1}) + \cdots + (a_1 A) + a_0 I_n)P \\
&= P^{-1}(q(A))P
\end{aligned}
$$

(9) Let A be a 3×3 matrix with eigenvalues $1, -1, 0$.
 (a) Prove that A is diagonalizable. What is D in this case?
 (b) Show that $A^n = A$ for any **odd** $n \geq 0$: e.g., $A^{13} = A$, $A^{77} = A$, etc.

Solution 9:
The matrix is diagonalizable as it has three distinct eigenvalues. We have $D = P^{-1}AP$
where $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.
It is easy to see that

$$
D^n = \begin{bmatrix} 1^n & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & 0^n \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & 0 \end{bmatrix}
$$

Thus we see clearly that $D^n = D$ if and only if n is odd.
Next, write $A = PDP^{-1}$. Then $A^n = (PDP^{-1})^n = PD^nP^{-1}$, where this last step is the extremely useful fact that "conjugation respects powers of matrices". If n is odd, we thus have $A^n = PD^nP^{-1} = PDP^{-1} = A$, as claimed.

(10) The *Fibonacci sequence* $F_0, F_1, F_2, \ldots$ is defined recursively as follows:

$$F_0 = 0$$
$$F_1 = 1$$
$$F_n = F_{n-2} + F_{n-1} \qquad\qquad \text{for } n \geq 2$$

Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$.

(a) Show that

$$A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix}$$

for all $n \geq 0$.

(b) Diagonalize A. Write its two eigenvalues as λ_1 and λ_2.

(c) Derive a closed formula for A^n.

(d) Derive a closed formula for F_n.

Solution 10:

(a) Proof by induction. Base case $(n = 0)$ is trivial. For induction step, suppose $A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix}$, then

$$
\begin{aligned}
A^{n+1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} &= A A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= A \begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix} && \text{(induction hypo.)} \\
&= \begin{bmatrix} F_{n+1} \\ F_n + F_{n+1} \end{bmatrix} \\
&= \begin{bmatrix} F_{n+1} \\ F_{n+2}. \end{bmatrix}
\end{aligned}
$$

as desired.

(b) The eigenvalues of A are $\lambda_1 = \frac{1}{2}(1 + \sqrt{5})$ and $\lambda_2 = \frac{1}{2}(1 - \sqrt{5})$.
Note that the λ_i satisfy
$\lambda_1 + \lambda_2 = 1$
$\lambda_1 \lambda_2 = -1$
$\lambda_1^2 = \lambda_1 + 1$, $\lambda_2^2 + \lambda_2 = 1$.
(These follow from the fact that they are roots of $p(x) = x^2 - x - 1 = (x - \lambda_1)(x - \lambda_2)$.)

From these relations it is easy to verify that $\mathbf{v}_1 = (1, \lambda_1)$ and $\mathbf{v}_2 = (1, \lambda_2)$ form bases of W_{λ_1} and W_{λ_2}, respectively.

Then we have $D = P^{-1}AP$ where

$$
D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}, \quad P^{-1} = -\frac{1}{\sqrt{5}} \begin{bmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{bmatrix}.
$$

(c) Now we compute

$$
A^n = (PDP^{-1})^n = PD^nP^{-1} = -\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} \begin{bmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \lambda_1^{n-1} - \lambda_2^{n-1} & \lambda_1^n - \lambda_2^n \\ \lambda_1^n - \lambda_2^n & \lambda_1^{n+1} - \lambda_2^{n+1} \end{bmatrix}
$$

(d) Lastly, using our formula above for A^n we see that

$$\begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix} = A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \lambda_1^n - \lambda_2^n \\ \lambda_1^{n+1} - \lambda_2^{n+1} \end{bmatrix}.$$

It follows that

$$F_n = \frac{1}{\sqrt{5}}(\lambda_1^n - \lambda_2^n),$$

where $\lambda_1 = \frac{1}{2}(1 + \sqrt{5}), \lambda_2 = \frac{1}{2}(1 - \sqrt{5})$.

(11) Let $A \in M_{nn}$ be defined as the matrix of all 1's: i.e., $a_{ij} = 1$ for all $1 \le i, j \le n$.
Show that A has exactly 2 eigenvalues and that A is diagonalizable.
Hint: do NOT compute the characteristic polynomial of A. Instead, begin by first considering null(A).
Solution 11:

Since rank$(A) = 1$, we have nullity$(A) = \dim W_0 = n - 1$. Thus 0 is an eigenvalue of A with geometric multiplicity $n - 1$. Since geometric multiplicity is bounded by algebraic multiplicity, we must have the characteristic polynomial $p(t)$ factoring as $p(t) = t^{n-1}(t - c)$ for some $c \in \mathbb{R}$. (Remember that $\deg p(t) = n$.) Now use the fact that

$$n = \operatorname{tr} A = (\text{sum of eigenvalues}) = 0 + 0 + \cdots + 0 + c.$$

It follows that $c = n$. Thus the distinct eigenvalues of A are 0 and c. Since $\dim W_0 = n - 1$ and $1 \le \dim W_n \le 1$, we have

$$\dim W_0 + \dim W_n = n - 1 + 1 = n.$$

Thus A is diagonalizable.
By the way, try diagonalizing A yourself: that is, produce a basis of eigenvectors.

(12) The *Cayley-Hamilton theorem* states that any $A \in M_{nn}$ satisfies its own characteristic polynomial $p(t)$: i.e., $p(A) = \mathbf{0}_{n \times n}$.
Prove the Cayley-Hamilton theorem for *diagonalizable* matrices.
Solution 12:

First, since A is diagonalizable, we write $D = P^{-1}AP$ where

$$D = \begin{bmatrix} d_1 & 0 & \cdots & & \\ 0 & d_2 & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 0 & d_n \end{bmatrix}$$

Next, observe that for any polynomial $q(t)$, we easily compute $q(D)$ as follows:

$$q(D) = \begin{bmatrix} q(d_1) & 0 & \cdots & & \\ 0 & q(d_2) & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 0 & q(d_n) \end{bmatrix}$$

That is, when evaluating $q(D)$ we simply get the diagonal matrix whose j-th diagonal entry is $q(d_j)$.

131

Now let $p(t)$ be the characteristic polynomial of A. Since A and D are similar, $p(t)$ is also the characteristic polynomial of D, which is clearly $p(t) = (t - d_1)(t - d_2) \cdots (t - d_n)$. From the observation above, we see that

$$p(D) = \begin{bmatrix} p(d_1) & 0 & \ldots & & \\ 0 & p(d_2) & 0 & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & \ldots & 0 & p(d_n) \end{bmatrix} = 0_{n \times n},$$

since clearly $p(d_j) = 0$ for all j.

Lastly, we compute

$$p(A) = p(PDP^{-1}) = Pp(D)P^{-1} = P0_{n \times n}P^{-1} = 0_{n \times n}.$$

(Note that the equality $p(PDP^{-1}) = Pp(D)P^{-1}$ follows from Exercise 8.)
We have thus shown that A satisfies its own characteristic polynomial, as claimed.

(13) Information about three matrices is given below. Find all eigenvalues of each matrix and decide whether it is diagonalizable (you are given enough information in each case).
  (a) $A_1 \in M_{33}$, $p(t) = \det(tI - A_1) = t^3 - t^2$, $\text{null}(A_1) =$the plane perpendicular to $(1, 0, 1)$.
  (b) $A_2 \in M_{33}$, $p(t) = \det(tI - A_2) = t^3 + t^2 - t$.
  (c) $A_3 \in M_{22}$, $\text{tr } A = 4$, $\det A = 3$.

Solution 13:
(a) The eigenvalues of A are $0, 1$. We are given that $\text{null}(A_1) = W_0$ is a plane, hence 2-dimensional. Since W_1 must be 1-dimensional (looking at multiplicity), we have $\dim W_0 + \dim W_1 = 3$. Hence A is diagonalizable.
(b) The eigenvalues of A are $0, (-1 + \sqrt{5})/2, (-1 - \sqrt{5})/2$. Since A has three distinct eigenvalues, it is diagonalizable.
(c) Let λ_1, λ_2 be the roots of $p(t)$. Then we have

$$\text{tr } A = \lambda_1 + \lambda_2 = 4$$
$$\det A = \lambda_1 \lambda_2 3$$

We easily solve this system of two equations in the two unknown λ_i to conclude $\lambda_1 = 1$, $\lambda_2 = 3$. Since A has two distinct eigenvalues, it is diagonalizable.

(14) A matrix $A \in M_{nn}$ is called **nilpotent** if $A^k = 0$ for some $k \geq 1$.
   In what follows, assume A is nilpotent.
   (a) Prove that 0 is the only *possible* eigenvalue of A.

   (b) Prove that 0 is indeed an eigenvalue of A.
      Thus 0 is the only eigenvalue of A.

Solution 14:
(a) Suppose λ is an eigenvalue of A. Then we have $A\mathbf{v} = \lambda\mathbf{v}$ for some $\mathbf{v} \neq 0$. Then $A^k\mathbf{v} = \lambda^k\mathbf{v}$, as shown earlier. Since $A^k = 0_{n \times n}$, we thus have $\mathbf{0} = \lambda^k\mathbf{v}$. Since $\mathbf{v}$ is *nonzero* (!!), we conclude that $\lambda^k = 0$, and hence that $\lambda = 0$.
(b) Now we show that 0 actually is an eigenvalue. Since $A^k = 0_{n \times n}$, we have $\det(A^k) = (\det A)^k = 0$. Hence A is not invertible. From the invertibility theorem we conclude that 0 is an eigenvalue of A.

(15) A matrix $A \in M_{nn}$ is called **idempotent** if $A^2 = A$.
   In what follows, assume A is idempotent.

(a) Prove that the only *possible* eigenvalues of A are 0 and 1.
(b) Prove that either 0 is an eigenvalue of A or else $A = I_n$.
(c) Give an example of an idempotent matrix for each of the following three possibilities: only 0 is an eigenvalue; only 1 is an eigenvalue; 0 and 1 are both eigenvalues.

Solution 15:
(a) Suppose λ is an eigenvalue of A. Then we have $A\mathbf{v} = \lambda\mathbf{v}$ for some $\mathbf{v} \neq 0$. Then $A^2\mathbf{v} = \lambda^2\mathbf{v}$ as shown earlier. Since $A^2 = A$, we have $A\mathbf{v} = \lambda^2\mathbf{v}$, and hence that $\lambda^2\mathbf{v} = \lambda\mathbf{v}$. Since $\mathbf{v}$ is *nonzero (!!)*, we conclude that $\lambda^2 = \lambda$, and hence that $\lambda = 0$ or $\lambda = 1$.
(b) Suppose 0 is not an eigenvalue. Then A is invertible.
Then $A^2 = A \Rightarrow A^{-1}A^2 = A^{-1}A \Rightarrow A = I$.
(c) Easy. You do it.

(16) Let $A \in M_{nn}$ and suppose the characteristic polynomial of A factors as
$$p(t) = (t - \lambda)^m g(t),$$
where the $\lambda \in \mathbb{R}$ and λ is not a root of $g(t)$.
Recall that we define the algebraic multiplicity of λ to be m; and we define the geometric multiplicity of λ to be $\dim W_\lambda$.
In this exercise we will show $\dim W_\lambda \leq m$. To do so, set $\dim W_\lambda = r$.
Show that A is similar to a matrix B whose characteristic polynomial is divisible by $(t-\lambda)^r$.
Conclude that $(t - \lambda)^r$ divides $p(t)$, and hence that $r \leq m$.
Solution 16:

Set $T = T_A$. Suppose $\dim W_\lambda = r$. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ be a basis for W_λ and extend to a full basis B' of $\mathbb{R}^n$. Representing T with respect to B' yields
$$A' = [T]_{B'} = \begin{bmatrix} \lambda I_r & C \\ 0_{(n-r)\times r} & D \end{bmatrix}$$

Here I've represented A' in terms of matrix blocks. Now, since A and A' are just two matrix representations of T, they are similar, hence have the same characteristic polynomial. From the block matrix description above it is easy to see that $p(t) = \det(tI - A')$ factors as $p(t) = (t - \lambda)^r q(t)$. Thus the algebraic multiplicity of λ is *at least* $r = \dim W_\lambda$, as claimed.

(17) Each of the matrices below has characteristic polynomial $p(x) = (x - 1)^2(x + 2)$.
For each decide whether it is diagonalizable. If yes, provide explicit P and D witnessing this fact.

(a) $A = \begin{bmatrix} -5 & 0 & 3 \\ -6 & 1 & 3 \\ -6 & 0 & 4 \end{bmatrix}$.

(b) $B = \begin{bmatrix} -2 & -3 & 3 \\ -3 & -3 & 4 \\ -3 & -4 & 5 \end{bmatrix}$.

Solution 17:
    These are the matrices that showed up in a previous exercise, where we did essentially all the necessary work.
    For A we have $\dim W_1 + \dim W_{-2} = 2 + 1 = 3$, so A is diagonalizable.
    For B we have $\dim W_1 + \dim W_{-2} = 1 + 1 = 2 < 3$ so B is not diagonalizable.

(18) Let $A = \begin{bmatrix} -5 & 0 & 3 \\ -6 & 1 & 3 \\ -6 & 0 & 4 \end{bmatrix}$. Use your work above to do the following to find a matrix C such that $C^3 = A$.

Solution 18:

From a previous exercise, we know that A has eigenspaces W_1 and W_{-2} with corresponding bases $\{(1,0,2),(0,1,0)\}$ and $\{(1,1,1)\}$. It follows that $D = P^{-1}AP$, where

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix}$$

We first find a cube-root of D. This is easy! Take

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\sqrt[3]{2} \end{bmatrix}$$

for example. Then $C = PEP^{-1}$ satsifies $C^3 = (PEP^{-1})^3 = PE^3P^{-1} = PDP^{-1} = A$. Thus C is a a cube-root of A. Numerically, we have

$$C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\sqrt[3]{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1-2\sqrt[3]{2} & 0 & 1+\sqrt[3]{2} \\ -2-2\sqrt[3]{2} & 1 & 1+\sqrt[3]{2} \\ -2-2\sqrt[3]{2} & 0 & 2+\sqrt[3]{2} \end{bmatrix}.$$

(19) Suppose A is 3×3, has determinant 0, and has among its eigenvalues $\lambda = 2$ with eigenspace W_2 a plane in $\mathbb{R}^3$.
  (a) Find all eigenvalues of A.
  (b) Find $\text{tr}(A)$.
  (c) Decide whether A is diagonalizable.

Solution 19:
Consider the characteristic polynomial $p(x)$ of A. Observe that $p(x)$ has degree 3.
Since W_2 is has dimension 2, we must have $p(x) = (x-2)^2(x-a)$, where a is possibly a second distinct eigenvalue, though it is also possible that $a = 2$. However $\det(A) = 0$ implies 0 is an eigenvalue of A (see previous classwork). Thus we must have $a = 0$ and $p(x) = x(x-2)^2$.
The trace of A is the sum of the eigenvalues. Thus $\text{tr}(A) = 2 + 2 + 0 = 4$.
We have $\dim W_2 = 2$. We have automatically that $\dim W_0 \geq 1$ (if λ is an eigenvalue, then by definition the eigenspace W_λ is nonzero). Since the algebraic multiplicity of 0 is 1, it follows that $\dim W_0 = 1$, and $\dim W_2 + \dim W_0 = 3$. This shows A is diagonalizable.

(20) Give an example of a 3×3 matrix A satisfying the following conditions:
  (i) $\lambda = 2$ is an eigenvalue and $(1,0,1),(1,1,1)$ are 2-eigenvectors.
  (ii) $\lambda = -1$ is an eigenvalue and $(1,0,-1)$ is a -1-eigenvector.

Solution 20:
Such an A will be diagonalizable (since we have a basis of eigenvectors). Using this fact to reverse engineer A, we see that the corresponding P and D would be

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix}, D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Since $D = P^{-1}AP$, it follows that

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

(21) True or false. If true, provide a proof; if false, give an explicit counterexample.
  (a) If $A \in M_{nn}$ has less than n distinct eigenvalues, then A is not diagonalizable.
  (b) If $A \in M_{nn}$ has fewer than n linearly independent eigenvectors, then A is not diagonalizable.
  (c) If $A \in M_{nn}$ is diagonalizable, then there is a *unique* matrix P such that $P^{-1}AP$ is diagonal.
  (d) If $A \in M_{nn}$ is diagonalizable, then A^{-1} is diagonalizable.
  (e) If $A \in M_{nn}$ is diagonalizable, then A^T is diagonalizable. s

Solution 21:
(a) False. Take the identity matrix I_2. It is diagonal, hence diagonalizable!
(b) True. This is the content of our main diagonalization theorem.
(c) False. Take $A = I_2$. Then we have $D = I_2$, and $D = P^{-1}AP$ for *any* invertible matrix P whatsoever!
(d) True. Suppose $D = P^{-1}AP$ is diagonal. Then so is D^{-1}, and we have $D^{-1} = P^{-1}A^{-1}(P^{-1})^{-1} = P^{-1}A^{-1}P$.
(e) True. Suppose $D = P^{-1}AP$ is diagonal. Then so is D^T and we have

$$D^T = P^T A^T (P^{-1})^T = P^T A^T (P^T)^{-1} = Q^{-1}AQ,$$

where $Q = (P^T)^{-1}$.

<center>5.1: INNER PRODUCTS</center>

In what follows V is always a vector space along with an inner product $\langle \, , \rangle$. Notions of norm and orthogonality always refer to this fixed choice of inner product.

(1) Show that $\frac{\mathbf{v}}{\|\mathbf{v}\|} := \frac{1}{\|\mathbf{v}\|} \cdot \mathbf{v}$ is a unit vector for any $\mathbf{v} \in V$.
  Solution 1:

  In any inner product space the norm operation satisfies $\|c\mathbf{v}\| = |c| \, \|\mathbf{v}\|$. (Observe the subtle notational difference: $|c|$ is absolute value of the real number c; $\|\mathbf{v}\|$ is norm of the vector $\mathbf{v}$.)

  Now let $\mathbf{w} = \dfrac{1}{\|\mathbf{v}\|}\mathbf{v}$. Then

$$\begin{aligned} \|\mathbf{w}\| &= \left\| \frac{1}{\|\mathbf{v}\|} \mathbf{v} \right\| \\ &= \left| \frac{1}{\|\mathbf{v}\|} \right| \|\mathbf{v}\| \\ &= \frac{1}{\|\mathbf{v}\|} \|\mathbf{v}\| \qquad\qquad \left(\text{since } \frac{1}{\|\mathbf{v}\|} \geq 0\right) \\ &= 1 \end{aligned}$$

<center>135</center>

(2) Let $(V, \langle\, ,\rangle)$ be an inner product space. Prove: $\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|$ for all $c \in \mathbb{R}$.
   (Here $|c|$ is the absolute value of c.)
   Your proof must be valid for a general inner product space; i.e., you can only invoke the definition of $\|\_\|$ in terms of $\langle\, ,\rangle$, the axioms of an inner product, and properties of real number arithmetic.
   Solution 2:

$$\|c\mathbf{v}\|^2 = \langle c\mathbf{v}, c\mathbf{v}\rangle$$
$$= c^2\langle \mathbf{v}, \mathbf{v}\rangle$$
$$= c^2\,\|\mathbf{v}\|^2$$

This implies
$$\|c\mathbf{v}\| = \sqrt{c^2}\,\|\mathbf{v}\| = |c|\,\|\mathbf{v}\|.$$

(3) Compute the angle between $\mathbf{v} = (1,1,1,1)$ and $\mathbf{w} = (1,-1,1,1)$. Do NOT use a calculator.
   Solution 3:

   The equation for the angle θ ends up being $\cos(\theta) = \frac{1}{2}$. The unique θ in the range $[0, \pi]$ satisfying this is $\theta = \pi/3 = 60°$.

(4) Let $V = C([0,1])$ with inner product $\langle f, g\rangle = \displaystyle\int_0^1 f(x)g(x)\,dx$. Let $f(x) = 1$ and $g(x) = x$.
   Compute the angle θ between f and g with respect to this inner product.
   Your answer cannot be expressed in terms of inverse trig functions: i.e., the angle is a familiar one that you can solve for by hand.
   Solution 4:

   The angle θ is the unique angle between 0 and π satisfying

$$\cos(\theta) = \frac{\langle f, g\rangle}{\|f\|\,\|g\|} = \frac{\int_0^1 x\,dx}{\sqrt{\int_0^1 1\,dx}\sqrt{\int_0^1 x^2\,dx}} = \frac{1/2}{1/\sqrt{3}} = \frac{\sqrt{3}}{2}.$$

   Thus $\theta = \pi/6$.

(5) Prove that if $\mathbf{w}$ is orthogonal to $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, then it is orthogonal to any linear combination of the $\mathbf{v}_i$. In other words prove the implication:

$$(\mathbf{w} \perp \mathbf{v}_i \text{ for all } i) \Rightarrow (\mathbf{w} \perp (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) \text{ for any } c_i)$$

   Solution 5:

   We prove both directions of the if and only if separately.
   ($\Leftarrow$): this is the easy direction. If $\mathbf{w}$ is orthogonal to *any* linear combination of the $\mathbf{v}_i$, then in particular it is orthogonal to each of the $\mathbf{v}_i$.
   ($\Rightarrow$): assume $\mathbf{w} \perp \mathbf{v}_i$ for all i. Then $\mathbf{w} \cdot \mathbf{v}_i = 0$ for all i. But then

$$\mathbf{w} \cdot (c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = \mathbf{w} \cdot (c_1\mathbf{v}_1) + \cdots + \mathbf{w} \cdot (c_n\mathbf{v}_n) \qquad \text{(add. prop. of dot)}$$
$$= c_1(\mathbf{w} \cdot \mathbf{v}_1) + \cdots + c_n\mathbf{w} \cdot \mathbf{v}_n) \qquad \text{(scalar mult. prop. of dot)}$$
$$= 0 + 0 + \cdots + 0 = 0.$$

   Thus $\mathbf{w} \perp (c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n)$ for any c_i, as desired.

(6) For each of the following operations on $\mathbb{R}^2$, determine whether it defines an inner product on $\mathbb{R}^2$. If it fails to be an inner product, identify which of the three inner product axioms (if any) it does satisfy, and provide explicit counterexamples for any axiom that fails.
   (a) $\langle (x_1, x_2), \ (y_1, y_2) \rangle = x_1 y_2 + x_2 y_1$.
   (b) $\langle (x_1, x_2), \ (y_1, y_2) \rangle = 2x_1 y_1 + x_1 y_2 + x_2 y_1 + 3x_2 y_2$.
   (c) $\langle (x_1, x_2), \ (y_1, y_2) \rangle = x_1^2 y_1^2 + x_2^2 y_2^2$.

Solution 6:
In what follows we let $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$, $\mathbf{z} = (z_1, z_2)$.
   (a) The proposed operation satisfies all axioms except positivity. Let's see why.
Symmetry. We have

$$\langle \mathbf{y}, \ \mathbf{x} \rangle = y_1 x_2 + y_2 x_1$$
$$= x_1 y_2 + x_2 y_1$$
$$= \langle \mathbf{x}, \mathbf{y} \rangle$$

Linearity in first variable. We have

$$\langle c\mathbf{x} + d\mathbf{y}, \mathbf{z} \rangle = \langle (cx_1 + dy_1, cx_2 + dy_2), \ (z_1, z_2) \rangle$$
$$= (cx_1 + dy_1)z_2 + (cx_2 + dy_2)z_1$$
$$= c(x_1 z_2 + x_2 z_1) + d(y_1 z_2 + y_2 z_1)$$
$$= c\langle \mathbf{x}, \ \mathbf{y} \rangle + d\langle \mathbf{x}, \ \mathbf{z} \rangle$$

Positivity. This fails. Take $\mathbf{x} = (1, -1)$. Then

$$\langle \mathbf{x}, \mathbf{x} \rangle = 1 \cdot (-1) + (-1) \cdot 1 = -2 < 0.$$

Furthermore, taking $\mathbf{x} = (0, 1)$, we see that $\langle \mathbf{x}, \mathbf{x} \rangle = 0$. Thus both parts of positivity fail.

(b) This is in fact an inner product! The proofs that properties (i) and (ii) hold proceed exactly as the argument above. Let's focus on positivity. We have

$$\langle \mathbf{x}, \mathbf{x} \rangle = 2x_1^2 + 2x_1 x_2 + 3x_2^2$$
$$= 2(x_1 + \frac{1}{2}x_2)^2 + \frac{11}{4}x_2^2 \qquad \text{(complete the square)}$$
$$= 2u^2 + \frac{11}{4}v^2 \qquad (u = x_1 + \frac{1}{2}x_2, \ v = x_2)$$

Clearly the expression $2u^2 + \frac{11}{4}v^2$ is nonnegative. Furthermore this will be equal to zero if and only if $u = v = 0$ if and only if $x_1 + \frac{1}{2}x_2 = x_2 = 0$ if and only if $x_1 = x_2 = 0$. This proves $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

(c) It is clear that the operation is symmetric: that is $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$. Thus axiom (i) is satsified.
The operation does not satisfy axiom (ii). For example, Let $\mathbf{x} = (1, 1)$, $\mathbf{y} = (0, 1)$. Then

$$\langle 2\mathbf{x}, \mathbf{y} \rangle = \langle (2, 2), (0, 1) \rangle = 4 \neq 2\langle \mathbf{x}, \mathbf{y} \rangle = 1.$$

The operation does satisfy positivity, as is easily shown.

(7) Equip P_2 with the *evaluation inner product* $\langle p, q \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$. Find all polynomials $p(x)$ orthogonal to $q(x) = x$ with respect to this inner product.
Note: this is an infinite set of polynomials. Describe it by giving a parameter description of its elements.
Solution 7:

A polynomial p is orthogonal to q if and only if $p(-1)q(-1)+p(0)q(0)+p(1)q(1) = 0$ if and only if $-p(-1)+p(1) = 0$. If $p(x) = ax^2 + bx + c$, this means $-(a-b+c)+(a+b+c) = 0$, or $2b = 0$. Thus $p(x) = ax^2 + c$, for $a, c \in \mathbb{R}$ any real numbers.

(8) Equip $V = C([-\pi, \pi])$ with the *integral inner product* $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)\ dx$.
    (a) Show that $f(x) = \cos(x)$ and $g(x) = \sin(x)$ are orthogonal with respect to this inner product.
    (b) Compute $\|\cos(x)\|$ with respect to this inner product.
    (c) Show that if $f(x)$ is an odd function (i.e., $f(x) = -f(-x)$ for all x) and $g(x)$ is an even function ($g(-x) = g(x)$ for all x), then f and g are orthogonal with respect to this inner product.
    Hint: use the area interpretation of the integral, as well as graphical properties of even/odd functions.

Solution 8:
Orthogonality follows from $\int_{-\pi}^{\pi} \cos x \sin x\ dx = \int_{\pi}^{\pi} \frac{1}{2}\sin(2x)\ dx = 0$, as one easily shows.
    We have
$$\|\cos x\| = \sqrt{\langle \cos x, \cos x \rangle} = \sqrt{\int_{-\pi}^{\pi} \cos^2(x)\ dx} = \sqrt{\pi}.$$

If f is odd and g is even, then their product fg is odd. The integral of an odd function over a symmetric interval like $[-\pi, \pi]$ is always 0. You can see this using the area interpretation (any region to the right of the y-axis has a flipped region to the left that cancels out its contribution to the integral), or by using integral properties:

$$\int_{-\pi}^{\pi} f(x)g(x)\ dx = \int_{\pi}^{-\pi} -f(-u)g(-u)\ du \qquad \text{(change of variables } u = -x, du = -dx)$$

$$= \int_{\pi}^{-\pi} f(u)g(u)\ du \qquad\qquad \text{(since } fg \text{ is odd)}$$

$$= -\int_{-\pi}^{\pi} f(x)g(x)\ dx \qquad \text{(flip the enpoints, change variable name)}$$

Since $\int_{-\pi}^{\pi} f(x)g(x)\ dx = -\int_{-\pi}^{\pi} f(x)g(x)\ dx$, we conclude $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)\ dx = 0$.

(9) Let $\mathbf{v}, \mathbf{w} \in V$, and let θ be the angle between them. Prove the following equivalence:
$$\|\mathbf{v} + \mathbf{w}\| = \|\mathbf{v}\| + \|\mathbf{w}\| \text{ if and only if } \theta = 0.$$
Your proof should be a *chain of equivalences* with each step justified.
Solution 9:

$$\|\mathbf{v} + \mathbf{w}\| = \|\mathbf{v}\| + \|\mathbf{w}\| \Leftrightarrow \|\mathbf{v} + \mathbf{w}\|^2 = (\|\mathbf{v}\| + \|\mathbf{w}\|)^2 \qquad\qquad (\text{ square both sides})$$

$$\Leftrightarrow (\mathbf{v} + \mathbf{w}) \cdot (\mathbf{v} + \mathbf{w}) = \|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2$$

$$\Leftrightarrow \mathbf{v} \cdot \mathbf{v} + 2\mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{v} + 2\|\mathbf{v}\|\|\mathbf{w}\| + \mathbf{w} \cdot \mathbf{w}$$

$$\Leftrightarrow \mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\|$$

$$\Leftrightarrow \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} = 1$$

$$\Leftrightarrow \cos(\theta) = 1$$

$$\Leftrightarrow \theta = 0.$$

(10) Take $\mathbf{v}, \mathbf{w} \in V$. Suppose $\|\mathbf{v}\| = 2$ and $\|\mathbf{w}\| = 3$. Find the maximum and minimum possible values of $\|\mathbf{v} - \mathbf{w}\|$, and give explicit examples where those values occur.

Solution 10:

We we have
$$\|\mathbf{v} - \mathbf{w}\|^2 = (\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w})$$
$$= \|\mathbf{v}\|^2 - 2\mathbf{v} \cdot \mathbf{w} + \|\mathbf{w}\|^2 .$$

By Cauchy-Schwarz
$$- \|\mathbf{v}\| \, \|\mathbf{w}\| \leq \mathbf{v} \cdot \mathbf{w} \leq \|\mathbf{v}\| \, \|\mathbf{w}\| .$$

Thus
$$\|\mathbf{v}\|^2 - 2 \|\mathbf{v}\| \, \|\mathbf{w}\| + \|\mathbf{w}\|^2 \leq \|\mathbf{v} - \mathbf{w}\|^2 \leq \|\mathbf{v}\|^2 + 2 \|\mathbf{v}\| \, \|\mathbf{w}\| + \|\mathbf{w}\|^2$$

which implies
$$(\|\mathbf{v}\| - \|\mathbf{w}\|)^2 \leq \|\mathbf{v} - \mathbf{w}\|^2 \leq (\|\mathbf{v}\| + \|\mathbf{w}\|)^2$$

and thus
$$\|\mathbf{v}\| - \|\mathbf{w}\| \leq \|\mathbf{v} - \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$$

So in our example
$$1 \leq \|\mathbf{v} - \mathbf{w}\| \leq 5.$$

In the special case when $V = \mathbb{R}^2$ with the dot product, the picture of this situation consists of two circles centered about the origin: one of radius 2, the other of radius 3. The vector $\mathbf{v}$ represents a point on the small circle; the vector $\mathbf{w}$ represents a point on the large one. These are closest when they lie along the same radius vector from the origin; they are furthest when lying on diametrically opposed radii.

(11) Prove the *Pythagorean theorem for general inner product spaces*: if $\mathbf{v} \perp \mathbf{w}$, then
$$\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 .$$

Solution 11:

We have
$$\|\mathbf{v} + \mathbf{w}\|^2 = \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle$$
$$= \langle \mathbf{v}, \mathbf{v} \rangle + 2\langle \mathbf{v}, \mathbf{w} \rangle^{0} + \langle \mathbf{w}, \mathbf{w} \rangle \qquad \text{(since } \mathbf{v} \perp \mathbf{w} \text{)}$$
$$= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$$

(12) Prove each inequality below using a judicious choice of inner product in conjunction with the Cauchy-Schwarz inequality (and possibly a judicious choice of one of the vectors in the Cauchy-Schwarz inequality).

(a) For all $f, g \in C([a, b])$
$$\left(\int_a^b f(x) g(x) \ dx \right)^2 \leq \int_a^b f^2(x) \ dx \int_a^b g^2(x) \ dx.$$

(b) For all $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$,
$$(x_1 + x_2 + \cdots + x_n) \leq \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \sqrt{n}.$$

(c) For all $a, b, \theta \in \mathbb{R}$
$$(a \cos(\theta) + b \sin(\theta))^2 \leq a^2 + b^2.$$

Solution 12:
(a) Once we translate the integrals as inner products, this inequality is obtained by squaring both sides of the Cauchy-Schwarz (C-S) inequality.
(b) This is an instance of C-S, where $V = \mathbb{R}^n$ with dot product, $\mathbf{x}$ is arbitrary, and $\mathbf{y} = (1, 1, \ldots, 1)$.
(c) This is an instance of C-S, where $V = \mathbb{R}^2$ with dot product, $\mathbf{x} = (a, b)$ and $\mathbf{y} = (\cos(\theta), \sin(\theta))$. Note that $\|\mathbf{y}\| = 1$ for this choice of $\mathbf{y}$.

(13) Let $(V, \langle\,,\,\rangle)$ be an inner product space. Recall in this context that we define $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$, and $d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|$.
An **isometry** of V is a function $f \colon V \to V$ that preserves distance: i.e.,

$$d(f(\mathbf{v}), f(\mathbf{w})) = d(\mathbf{v}, \mathbf{w}) \text{ for all } \mathbf{v}, \mathbf{w} \in V.$$

In this exercise we will show that any isometry that maps $\mathbf{0}$ to $\mathbf{0}$ is a linear transformation. (This is a handy fact to know. For example, reflections through lines in $\mathbb{R}^2$ and planes in $\mathbb{R}^3$ clearly preserve distances and map $\mathbf{0}$ to itself; it follows immediately that these operations are linear.)
In what follows, assume that f is an isometry of V satisfying $f(\mathbf{0}) = \mathbf{0}$.
(a) Prove that $\|f(\mathbf{v})\| = \|\mathbf{v}\|$: i.e., f preserves norms.
(b) Prove $\langle f(\mathbf{v}), f(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$: i.e., f preserves inner products.
   Hint: first prove that $\langle \mathbf{v}, \mathbf{w}\rangle = \frac{1}{2}(\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - \|\mathbf{v} - \mathbf{w}\|^2)$.
(c) To prove f is linear it is enough to show $f(\mathbf{v} + c\mathbf{w}) = f(\mathbf{v}) + cf(\mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in V$, $c \in \mathbb{R}$.
   To do so, use the above parts to show that

$$\|f(\mathbf{v} + c\mathbf{w}) - (f(\mathbf{v}) + cf(\mathbf{w}))\| = 0.$$

   Hint: regroup this difference in a suitable manner so that you can use parts (a)-(b). You may also want to use the identity

$$\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 - 2\langle \mathbf{v}, \mathbf{w}\rangle + \|\mathbf{w}\|^2.$$

Solution 13:
(a) We have

$$
\begin{aligned}
\|f(\mathbf{v})\| &= \|f(\mathbf{v}) - \mathbf{0}\| \\
&= \|f(\mathbf{v}) - f(\mathbf{0})\| && \text{(since } f(\mathbf{0}) = \mathbf{0}) \\
&= d(f(\mathbf{v}), f(\mathbf{0}) && \text{(by def.)} \\
&= d(\mathbf{v}, \mathbf{0}) && (f \text{ is an isometry}) \\
&= \|\mathbf{v}\|
\end{aligned}
$$

(b) The hint here is proved by starting with the RHS and substituting $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v}\rangle, \|\mathbf{w}\|^2 = \langle \mathbf{w}, \mathbf{w}\rangle, \|\mathbf{v} - \mathbf{w}\|^2 = \langle \mathbf{v} - \mathbf{w}, \mathbf{v} - \mathbf{w}\rangle$, and then using algebraic properties of the inner product to simplify the resulting expression until you get the LHS.

Once we have shown the equality mentioned in the hint, the rest is easy:

$$\langle f(\mathbf{v}), f(\mathbf{w})\rangle = \frac{1}{2}(\|f(\mathbf{v})\|^2 + \|f(\mathbf{w})\|^2 - \|f(\mathbf{v}) - f(\mathbf{w})\|^2) \qquad \text{(hint equality)}$$

$$= \frac{1}{2}(\|f(\mathbf{v})\|^2 + \|f(\mathbf{w})\|^2 - d(f(\mathbf{v}), f\mathbf{w}))^2) \qquad \text{(def. of dist.)}$$

$$= \frac{1}{2}(\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - d(\mathbf{v}, \mathbf{w})^2) \qquad (f \text{ preserves norm and dist.})$$

$$= \frac{1}{2}(\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - \|\mathbf{v} - \mathbf{w}\|^2) \qquad \text{(def. of dist.)}$$

$$= \langle \mathbf{v}, \mathbf{w}\rangle \qquad \text{(hint equality)}$$

(c) Note that proving

$$\|f(\mathbf{v} + c\mathbf{w}) - (f(\mathbf{v}) + cf(\mathbf{w}))\| = 0$$

implies that $f(\mathbf{v} + c\mathbf{w}) - (f(\mathbf{v}) + cf(\mathbf{w})) = \mathbf{0}$, and hence that $f(\mathbf{v} + c\mathbf{w}) = f(\mathbf{v}) + cf(\mathbf{w})$, as desired. We will prove the square of this norm is 0, and we group the expression as $(f(\mathbf{v} + c\mathbf{w}) - f(\mathbf{v})) - cf(\mathbf{w})$.
We have

$$\|(f(\mathbf{v} + c\mathbf{w}) - f(\mathbf{v})) - cf(\mathbf{w})\|^2$$

$$= \|f(\mathbf{v} + c\mathbf{w}) - f(\mathbf{v})\|^2 - 2\langle f(\mathbf{v} + c\mathbf{w}) - f(\mathbf{v}), cf(\mathbf{w})\rangle + \|cf(\mathbf{w})\|^2 \qquad \text{(suggested identity)}$$

$$= \|f(\mathbf{v} + c\mathbf{w}) - f(\mathbf{v})\|^2 - 2c\langle f(\mathbf{v} + c\mathbf{w}) - f(\mathbf{v}), f(\mathbf{w})\rangle + c^2 \|f(\mathbf{w})\|^2$$

$$= \|\mathbf{v} + c\mathbf{w} - \mathbf{v}\|^2 - 2c\langle f(\mathbf{v} + c\mathbf{w}) - f(\mathbf{v}), f(\mathbf{w})\rangle + c^2 \|\mathbf{w}\|^2 \qquad (f \text{ preserves distance and norms})$$

$$= c^2 \|\mathbf{w}\|^2 - 2c\langle f(\mathbf{v} + c\mathbf{w}), f(\mathbf{w})\rangle + 2c\langle f(\mathbf{v}), f(\mathbf{w})\rangle + c^2 \|\mathbf{w}\|^2$$

$$= c^2 \|\mathbf{w}\|^2 - 2c\langle \mathbf{v} + c\mathbf{w}, \mathbf{w}\rangle + 2c\langle \mathbf{v}, \mathbf{w}\rangle + c^2 \|\mathbf{w}\|^2 \qquad f \text{ preserves inner product)}$$

$$\vdots \qquad \text{(inner product simplification)}$$

$$= 0$$

5.2: ORTHOGONAL BASES AND PROJECTION

All questions related to orthogonality, norm, distance, etc. in the following exercises are always understood to be with regard to the given inner product.

(1) Let $(V, \langle, \rangle)$ be an inner product space, and let $W \subseteq V$ be a subspace. Show that $W^\perp$ is a subspace.
Solution 1:

(i) Since $\langle \mathbf{v}, \mathbf{0}\rangle = 0$ for *any* $\mathbf{v}$ *whatsoever*, we have $\mathbf{0} \in W^\perp$.
(ii)+(iii) Suppose $\mathbf{v}_1, \mathbf{v}_2 \in W^\perp$. We must show $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \in W^\perp$ for any $c_1, c_2 \in \mathbb{R}$. Given any $\mathbf{w} \in W$, we compute

$$\langle \mathbf{w}, c_1\mathbf{v}_1 + c_2\mathbf{v}_2\rangle = c_1\langle \mathbf{w}, \mathbf{v}_1\rangle + c_2\langle \mathbf{w}, \mathbf{v}_2\rangle \qquad \text{(bilinearity}$$

$$= c_1 \cdot 0 + c_2 \cdot 0 \qquad (\text{since } \mathbf{v}_i \in W^\perp)$$

$$= 0.$$

This proves $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \in W^\perp$, as desired.

(2) Let $(V, \langle,\rangle)$ be an inner product space, and let $W \subseteq V$ be a subspace. Show that $W \cap W^\perp = \{\mathbf{0}\}$.
Solution 2:

First observe that $\mathbf{0}_V \in W$ and $\mathbf{0}_V \in W^\perp$, since both are subspaces. It follows that $\mathbf{0}_V \in W \cap W^\perp$ (by def. of $\cap$), and thus that $\{\mathbf{0}_V\} \subseteq W \cap W^\perp$.
To show the other subset relation, suppose $\mathbf{v} \in W \cap W^\perp$. Recall that by definition of $\cap$ this means $\mathbf{v} \in W$ and $\mathbf{v} \in W^\perp$. Since an element of $W^\perp$ is orthogonal to everything in W we must have $\mathbf{v}$ orthogonal to itself. This means $\langle \mathbf{v}, \mathbf{v} \rangle = 0$. By the positivity axiom it now follows that $\mathbf{v} = \mathbf{0}_V$. Thus the only element of $W \cap W^\perp$ is $\mathbf{0}_V$, and hence $W \cap W^\perp = \{\mathbf{0}_V\}$.

(3) Consider the vector space $V = C([0,1])$ with standard *integral inner product*.
Apply Gram-Schmidt to the basis $B = \{1, 2^x, 3^x\}$ of $W = \operatorname{span}(B)$ to obtain an orthogonal basis of W.
Solution 3:

The resulting orthogonal basis is $B' = \{f_1, f_2, f_3\}$, where

$$f_1 = 1$$
$$f_2 = 2^x - (\langle 2^x, 1 \rangle / \langle 1, 1 \rangle)1$$
$$= 2^x - \left(\int_0^1 2^x\ dx\right)/\left(\int_0^1 1\ dx\right) = 2^x - \frac{1}{\ln 2}$$
$$f_3 = 3^x - (\langle 3^x, 2^x - \frac{1}{\ln 2}\rangle / \langle 2^x - \frac{1}{\ln 2}, 2^x - \frac{1}{\ln 2}\rangle)(2^x - \frac{1}{\ln 2}) - (\langle 3^x, 1 \rangle / \langle 1, 1 \rangle)1$$
$$= 3^x - \frac{\frac{2}{\ln 2 \ln 3} + \frac{5}{\ln 6}}{\frac{1}{(\ln 2)^2} + \frac{3}{\ln 4}}(2^x - \frac{1}{\ln 2}) - \frac{1}{\ln 3}$$

OK, I admit, I used technology to compute those integrals.

(4) Consider the vector space $V = P_2$ with the *evaluation inner product*

$$\langle p(x), q(x) \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1).$$

Apply Gram-Schmidt to the standard basis of P_2 to obtain an orthogonal basis of P_2.
Solution 4:

Applying Gram-Schmidt to the basis $B = \{1, x, x^2\}$ yields the orthogonal basis

$$B' = \left\{p_1 = 1, p_2 = x, p_3 = x^2 - \frac{2}{3}\right\}.$$

Here are the steps:

$$p_1 = 1$$

$$p_2 = x - \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle} 1$$

$$= x \qquad\qquad\qquad (\text{since } \langle x, 1 \rangle = -1 + 0 + 1 = 0)$$

$$p_3 = x^2 - \frac{\langle x^2, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2 - \frac{\langle x^2, 1 \rangle}{\langle 1, 1 \rangle}$$

$$= x^2 - \frac{0}{2} x - \frac{2}{3} 1 \qquad\qquad \left(\begin{array}{c} \langle x^2, x \rangle = -1 + 0 + 1 = 0 \\ \langle x^2, 1 \rangle = 1 + 0 + 1 = 2 \\ \langle 1, 1 \rangle = 1 + 1 + 1 = 3 \end{array} \right)$$

$$= x^2 - \frac{2}{3}$$

(5) Let $V = M_{22}$ with inner product $\langle A, B \rangle = \mathrm{tr}(A^T B)$, and let $W \subseteq V$ be the subspace of matrices whose trace is 0.
   (a) Compute an orthogonal basis for W.
   (b) Compute $\mathrm{proj}_W(A)$, where A is the matrix $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$

Solution 5:
(a) The general element of W looks like $A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$. From this description we see easily that

$$B = \left\{ A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$

is a basis for W. Furthermore, a simple computation shows this basis is already orthogonal. What luck!
(b) With the orthogonal basis B in hand, we use the orthogonal projection formula to compute

$$\mathrm{proj}_W(A) = \langle A, A_1 \rangle / \langle A_1, A_1 \rangle A_1 + \langle A, A_2 \rangle / \langle A_2, A_2 \rangle A_2 + \langle A, A_3 \rangle / \langle A_3, A_3 \rangle A_3$$
$$= 0 A_1 + 2 A_2 + A_3$$
$$= \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$$

(6) Let $V = M_{nn}$ with inner product $\langle A, B \rangle = \mathrm{tr}(A^T B)$.
Let W_1 be the subspace of all symmetric matrices $(A^T = A)$, and let W_2 be the subspace of all skew-symmetric matrices $(A^T = -A)$.
Recall (from a previous exercise) that $\dim W_1 = \frac{n(n+1)}{2}$ and $\dim W_2 = \frac{n(n-1)}{2}$.
   (a) Show that $W_2 \subseteq (W_1)^\perp$: i.e., skew-symmetric matrices are orthogonal to symmetric matrices!
   (b) Use the result of Exercise 12 and a dimension argument to show that in fact $W_2 = W_1^\perp$.
   (c) Use parts (a)-(b) and the orthogonal projection theorem to show that any matrix $A \in M_{nn}$ can be written uniquely in the form $A = A_1 + A_2$, where A_1 is a symmetric matrix and A_2 is a skew-symmetric matrix.

Solution 6:

First I reveal a dirty little secret. This "exotic" inner product is really nothing more than the dot product in disguise, once we take these rectangular arrays and stretch them out into column vectors! More rigorously, given $A = (a_{ij})$ and $B = (b_{ij})$, we have

$$
\begin{aligned}
\langle A, B \rangle &= \operatorname{tr}(A^T B) \\
&= \sum_{k=1}^{n} \left(A^T B\right)_{kk} \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{n} (A^T)_{k\ell}(B)_{\ell k} \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{n} (A)_{\ell k}(B)_{\ell k} \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{n} a_{\ell k} b_{\ell k} \\
&= \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} a_{ij} b_{ij}
\end{aligned}
$$

In other words, the inner product of two matrices is just the sum of the products of their corresponding entries: aka, the dot product!

Let's use this last formula to compute $\langle A, B \rangle$ when $A^T = -A$ and $B^T = B$. First, since skew-symmetric matrices must have 0's along the diagonal, we have $a_{ii} = 0$ for all i. For $i \neq j$, we have $a_{ji} = -a_{ij}$ and $b_{ji} = b_{ij}$. Then

$$
\begin{aligned}
\langle A, B \rangle &= \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} a_{ij} b_{ij} \\
&= \sum_{1 \leq i \leq n} a_{ii} b_{ii} + \sum_{1 \leq i < j \leq n} a_{ij} b_{ij} + \sum_{1 \leq i < j \leq n} a_{ji} b_{ji} \quad \text{(grouping the terms of the sum any way we like)} \\
&= 0 + \sum_{1 \leq i < j \leq n} a_{ij} b_{ij} + \sum_{1 \leq i < j \leq n} -a_{ij} b_{ij} \quad\quad\quad\quad\quad (a_{ii} = 0,\ a_{ji} = -a_{ij},\ b_{ij} = b_{ji}) \\
&= \sum_{1 \leq i < j \leq n} a_{ij} b_{ij} - \sum_{1 \leq i < j \leq n} a_{ij} b_{ij} \\
&= 0.
\end{aligned}
$$

Here's an alternative, slightly slicker approach. First observe that in general $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. This is a somewhat surprising property, but follows easily from an argument similar to above: namely,

$$
\operatorname{tr}(AB) = \sum_{1 \leq i, j \leq n} a_{ji} b_{ij} = \sum_{1 \leq i \leq j \leq n} b_{ij} a_{ji} = \operatorname{tr}(BA).
$$

Now assume $A^T = -A$ and $B^T = B$. Then

$$
\begin{aligned}
\operatorname{tr}(A^T B) &= \operatorname{tr}\left((A^T B)^T\right) && \text{(since transpose doesn't change diagonal)} \\
&= \operatorname{tr}(B^T A) && \text{(prop. of transpose)} \\
&= \operatorname{tr}(AB^T) && \text{(prop. of trace mentioned above)} \\
&= \operatorname{tr}(-A^T B) && \text{(since } A^T = -A,\ B^T = B) \\
&= -\operatorname{tr}(A^T B) && \text{(prop. of trace)}
\end{aligned}
$$

We've shown $\operatorname{tr}(A^T B) = -\operatorname{tr}(A^T B)$; it follows that $\operatorname{tr}(A^T B) = 0$.

(b) We have $W_2 \subseteq W_1^{\perp}$. By Exercise 12, we know $\dim W_1^{\perp} = \dim M_{nn} - \dim W_1 = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$. But we've seen in a previous exercise that $\dim W_2 = \frac{n(n-1)}{2}$. It follows from the dimension theorem compendium (subspace part) that $W_2 = W_1^{\perp}$.

(c) By the orthogonal projection theorem, any $A \in M_{nn}$ can be expressed as $A = B + B^{\perp}$ where $B \in W_1$ and $B^{\perp} \in W_1^{\perp}$. We have W_1 the space of symmetric matrices, and we just showed that $W_1^{\perp}$ is the space of skew-symmetric matrices. Thus B is symmetric and $B^{\perp}$ is skew-symmetric, as desired.

(7) Let $V = M_{22}$ with inner product $\langle A, B \rangle = \operatorname{tr}(A^T B)$, let W_1 be the subspace of symmetric matrices, and let W_2 be the subspace of skew-symmetric matrices.

(a) Verify that $B_1 = \left\{ A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is an orthogonal basis of W_1.

Verify that $B_2 = \left\{ A_3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$ is an orthogonal basis of W_2.

(b) Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$.

  (i) Use the previous exercise and the orthogonal projection formula to write $A = A_1 + A_2$, where A_1 is symmetric and A_2 is skew-symmetric.
  (ii) Compute $d(A, W_1)$, the distance from A to the subspace W_1.

Solution 7:
(a) It is easy to see that the given bases are orthogonal.
(b) From the previous exercise, we know that the orthogonal decomposition of $A = B + B^{\perp}$ yields the desired decomposition of A into a sum of a symmetric matrix and skew-symmetric matrix. We use the projection formulas to compute the projections B and $B^{\perp}$:

$$
\begin{aligned}
B &= \operatorname{proj}_{W_1}(A) \\
&= \langle A, A_1 \rangle / \langle A_1, A_1 \rangle A_1 + \langle A, A_2 \rangle / \langle A_2, A_2 \rangle A_2 + \langle A, A_3 \rangle / \langle A_3, A_3 \rangle A_3 \\
&= \frac{1}{1} A_1 + \frac{3}{2} A_2 + \frac{1}{1} A_3 \\
&= \begin{bmatrix} 1 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
B^{\perp} &= A - \operatorname{proj}_{W_1}(A) \\
&= \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}
\end{aligned}
$$

Observe that B and $B^{\perp}$ are indeed symmetric and skew-symmetric, resp., and that $A = B + B^{\perp}$.

By definition $d(A, W_1) = \|A - B\| = \|B^{\perp}\| = \sqrt{1/2} = \frac{\sqrt{2}}{2}$.

(8) Let $V = C([0,1])$ with the standard *integral inner product*, and let $f(x) = x$. Find the function of the form $g(x) = a + b\cos(2\pi x) + c\sin(2\pi x)$ that best approximates $f(x)$ (with respect to this inner product).
Solution 8:

The given inner product is $\langle f, g \rangle = \int_0^1 f(x)g(x)\ dx$. The desired function g is simply $\text{proj}_W(f)$, where $W = \text{span}\{f_1 = 1, f_2 = \cos(2\pi x), f_3 = \sin(2\pi x)\}$. The given functions form an orthogonal basis for W with respect to the given inner product, allowing us to use the orthogonal projection formula:

$$g = \text{proj}_W(f)$$
$$= \frac{\langle f, f_1 \rangle}{\langle f_1, f_1 \rangle} f_1 + \frac{\langle f, f_2 \rangle}{\langle f_2, f_2 \rangle} f_2 + \frac{\langle f, f_3 \rangle}{\langle f_3, f_3 \rangle} f_3$$
$$= \frac{1}{2} + \frac{0}{1/2}\cos(2\pi x) + \frac{1/2\pi}{1/2}\sin(2\pi x) \qquad \text{(I've omitted all the integral details)}$$
$$= \frac{1}{2} + \frac{1}{\pi}\sin(2\pi x).$$

(9) Compute an orthogonal basis of W, where W is the plane $x + 2y - z = 0$. Then extend this orthogonal basis to an orthogonal basis of $\mathbb{R}^3$.
Solution 9:

The plane W is the null space of the matrix $[1\ 2\ -1]$. The usual nulls space algorithm produces the basis $\{(1,0,1),(-2,1,0)\}$ for W. Use Gram-Schmidt to convert this to an orthogonal basis $B' = \{(1,0,1),(-1,1,1)\}$.
To extend B' to a basis of $\mathbb{R}^3$, we need only add another nonzero vector that is orthogonal to both $(1,0,1),(-1,1,1)$. That's easy! Since both vectors are elements of W, they are both orthogonal to its defining normal vector $\mathbf{n} = (1,2,-1)$. Thus $B'' = \{(1,0,1),(-1,1,1),(1,2,-1)\}$ is an orthogonal basis of $\mathbb{R}^3$.

(10) Let A be an $n \times n$ matrix.
    Prove: $A^T = A^{-1}$ if and only if the columns of A are an orthonormal basis of $\mathbb{R}^n$.
    Matrices satisfying $A^T = A^{-1}$ are called **orthogonal matrices**.
Solution 10:

(11) Let V an inner product space, and let $W \subseteq V$ be a finite-dimensional subspace.
    Recall that $\text{proj}_W(\mathbf{v})$ is defined as the unique $\mathbf{w} \in W$ satisfying $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$, where $\mathbf{w}^\perp \in W^\perp$.
    Use this definition (including the uniqueness claim) to prove the following properties:
    (a) $\text{proj}_W(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\text{proj}_W(\mathbf{v}_1) + c_2\text{proj}_W(\mathbf{v}_2)$ for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and all $c_1, c_2 \in \mathbb{R}$;
    (b) $\mathbf{v} \in W \Longleftrightarrow \text{proj}_W(\mathbf{v}) = \mathbf{v}$;
    (c) $\mathbf{v} \in W^\perp \Longleftrightarrow \text{proj}_W(\mathbf{v}) = \mathbf{0}$.

Solution 11:
(a) By the orthogonal projection theorem we can write
$$\mathbf{v}_1 = \mathbf{w}_1 + \mathbf{w}_1^\perp \quad \text{where } \mathbf{w}_1 \in W, \mathbf{w}_1^\perp \in W^\perp$$
$$\mathbf{v}_2 = \mathbf{w}_2 + \mathbf{w}_2^\perp \quad \text{where } \mathbf{w}_2 \in W, \mathbf{w}_2^\perp \in W^\perp.$$

By definition, this means $\text{proj}_W(\mathbf{v}_i) = \mathbf{w}_i$ for $i = 1, 2$. Then we have

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 = c_1(\mathbf{w}_1 + \mathbf{w}_1^\perp) + c_2(\mathbf{w}_2 + \mathbf{w}_2^\perp)$$
$$= (c_1\mathbf{w}_1 + c_2\mathbf{w}_2) + (c_1\mathbf{w}_1^\perp + c_2\mathbf{w}^\perp)$$

Since $c_1\mathbf{w}_1 + c_2\mathbf{w}_2 \in W$ and $c_1\mathbf{w}_1^\perp + c_2\mathbf{w}^\perp \in W^\perp$, we recognize

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 = (c_1\mathbf{w}_1 + c_2\mathbf{w}_2) + (c_1\mathbf{w}_1^\perp + c_2\mathbf{w}^\perp)$$

as an orthogonal decomposition. It follows, again by definition, that

$$\text{proj}_W(\mathbf{v}) = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 = c_1\text{proj}_W(\mathbf{v}_1) + c_2\text{proj}_W(\mathbf{v}_2).$$

(b) We have $\mathbf{v} \in W$ if and only if $\mathbf{v} = \mathbf{v} + \mathbf{0}$ ($\mathbf{v} \in W, \mathbf{0} \in W^\perp$) is the orthogonal decomposition of $\mathbf{v}$ with respect to W if and only if $\text{proj}_W(\mathbf{v}) = \mathbf{v}$.

(c) We have $\mathbf{v} \in W^\perp$ if and only if $\mathbf{v} = \mathbf{0} + \mathbf{v}$ ($\mathbf{0} \in W, \mathbf{v} \in W^\perp$) is the orthogonal decomposition of $\mathbf{v}$ with respect to W if and only if $\text{proj}_W(\mathbf{v}) = \mathbf{0}$.

(12) Let $(V, \langle \, , \, \rangle)$ be an inner product space of dimension n, and suppose $W \subseteq V$ is a subspace of dimension r. Prove: $\dim W^\perp = n - r$.
Hint: begin by picking an *orthogonal* basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ of W and extend to an *orthogonal* basis $B' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{u}_1, \ldots, \mathbf{u}_{n-r}\}$ of all of V. Show the $\mathbf{u}_i$ form a basis for $W^\perp$.
Solution 12:

Observe first that if we can show all the steps in the hint, then $W^\perp$ has as a basis $B'' = \{\mathbf{u}_1, \ldots, \mathbf{u}_{n-r}\}$, and hence $\dim(W^\perp) = n - r$, as desired.

To show B'' is a basis of $W^\perp$ we must show that (1) it is independent, and (2) it spans $W^\perp$. Independence is easy: this follows from the fact that the big set B' is independent (in fact orthogonal); any subset of a linearly independent set is itself linearly independent.

Now we show that B'' spans $W^\perp$. First note that each $\mathbf{u}_i$ is indeed an element of $W^\perp$ since by assumption the set B' is orthogonal. This means that $\langle \mathbf{u}_i, \mathbf{v}_j \rangle = 0$ for all j; since the $\mathbf{v}_j$ are a basis for W, it follows that $\langle \mathbf{u}_i, \mathbf{w} \rangle = 0$ for all $\mathbf{w} \in W$.

Next take any element $\mathbf{u} \in W^\perp$. Since B' is a basis of the whole space V, we can write $\mathbf{u}$ as a linear combination of the $\mathbf{v}_i$ and $\mathbf{u}_i$:

$$\mathbf{u} = c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r + d_1\mathbf{u}_1 + \cdots d_{n-r}\mathbf{u}_{n-r}.$$

We will show that each $c_i = 0$, and hence that in fact we have

$$\mathbf{u} = d_1\mathbf{u}_1 + \cdots d_{n-r}\mathbf{u}_{n-r} \in \text{span}(\{\mathbf{u}_1, \ldots, \mathbf{u}_{n-r}\}).$$

To show $c_i = 0$ we simply take the inner product of $\mathbf{u}$ with $\mathbf{v}_i$:

$$0 = \langle \mathbf{u}, \mathbf{v}_i \rangle \qquad \text{(since } \mathbf{u} \in W^\perp)$$
$$= \langle c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r + d_1\mathbf{u}_1 + \cdots d_{n-r}\mathbf{u}_{n-r}, \mathbf{v}_i \rangle$$
$$= c_1\cancel{\langle \mathbf{v}_1, \mathbf{v}_i \rangle}^{0} + \cdots c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle + \cdots c_r\cancel{\langle \mathbf{v}_r, \mathbf{v}_i \rangle}^{0} + d_1\cancel{\langle \mathbf{u}_1, \mathbf{v}_i \rangle}^{0} + \cdots + d_{n-r}\cancel{\langle \mathbf{u}_{n-r}, \mathbf{v}_i \rangle}^{0}$$
$$= c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle.$$

Since $0 = c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle$ and $\langle \mathbf{v}_i, \mathbf{v}_i \rangle \neq 0$ (by positivity), it follows that $c_i = 0$, as desired.

(13) Let $(V, \langle \, , \, \rangle)$ be an inner product space of dimension n. Let $T: V \to V$ be defined as $T(\mathbf{v}) = \text{proj}_W(\mathbf{v})$: i.e., T is orthogonal projection onto W.
Let B_1 be a basis for W and B_2 be a basis for $W^\perp$.
(a) Using Exercises 11 and 12, as well as the orthogonal projection theorem, show that $B = B_1 \cup B_2$ is a basis for V consisting of eigenvectors of T.

(b) Compute the corresponding diagonal matrix $D = [T]_B$. Make sure to indicate how many times a given eigenvalue is repeated.

Note: this shows that any orthogonal projection of a finite-dimensional inner product space is a diagonalizable.

Solution 13:

(a) Let $B_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ be a basis of W. By Exercise **??** dim $W^\perp = n - r$, and hence we can find a basis $B_2 = \{\mathbf{u}_1, \ldots, \mathbf{u}_{n-r}\}$.

By Exercise 11.b, $T(\mathbf{v}) = \mathbf{v}$ if and only if $\mathbf{v} \in W$. Thus $W = W_1$, the $(\lambda = 1)$-eigenspace of T.

By Exercise 11.c, we have have $T(\mathbf{v}) = \mathbf{0} = 0\mathbf{v}$ if and only if $\mathbf{v} \in W^\perp$. Thus $W^\perp = W_0$, the $(\lambda = 0)$-eigenspace of T.

Since the $\mathbf{v}_i$ are linearly independent elements of W_1, and the $\mathbf{u}_j$ are linearly independent vectors of W_0, it follows that $B = B_1 \cup B_2 = \{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{u}_1, \ldots, \mathbf{u}_{n-r}\}$ is linearly independent. (This is a consequence of the fact that eigenvectors with distinct eigenvalues are linearly independent.)

Since $\#B = n + (n - r) = n$ we see that B is a basis of V consisting of eigenvectors of T.

(b) Having shown that our B is a basis of eigenvectors, we know that $[T]_B$, the matrix representing T with respect to B, is diagonal. More specifically, following the recipe for computing $[T]_B$, we see that for $1 \leq j \leq r$, the j-th column is

$$[T(\mathbf{v}_j)]_B = [\mathbf{v}_j]_B = \underbrace{(0, 0, \ldots, 1, 0, \ldots, 0)}_{j\text{th-entry}} = \mathbf{e}_j;$$

and for $r + 1 \leq j \leq n$ the j-th column is

$$[T(\mathbf{u}_j)]_B = [\mathbf{0}]_B = (0, 0, \ldots, 0).$$

It follows that $[T]_B$ is a diagonal matrix whose first r diagonal entries are 1, and whose remaining $n - r$ diagonal entries are 0.

(14) We consider the problem of fitting a collection of data points (x, y) with a quadratic curve of the form $y = f(x) = ax^2 + bx + c$.

Thus we are *given* some collection of points (x, y), and we *seek* parameters a, b, c for which the graph of $f(x) = ax^2 + bx + c$ "best fits" the points in some way.

(a) Show, using linear algebra, that if we are given any three points $(x, y) = (r_1, s_1), (r_2, s_2), (r_3, s_3)$, where the x-coordinates r_i are all distinct, then there is a *unique* choice of a, b, c such that the corresponding quadratic function agrees *precisely* with the data.

In other words, given just about any three points in the plane, there is a unique quadratic curve connecting them.

(b) Now suppose we are given the four data points

$$P_1 = (0, 2), P_2 = (1, 0), P_3 = (2, 2), P_4 = (3, 6).$$

   (i) Use the least-squares method described in the lecture notes to come up with a quadratic function $y = f(x)$ that "best fits" the data.

  (ii) Graph the function f you found, along with the points P_i. (You may want to use technology.)

Use your graph to explain precisely in what sense f "best fits" the data.

Solution 14:

(a) Each data point (r, s) yields a *linear equation* in the unknowns a, b, c: namely, $r^2 a + rb + c = s$. Thus given three data points we get a linear system of three equations in the

three unknowns a, b, c that is represented as

$$\begin{bmatrix} r_1^2 & r_1 & 1 \\ r_2^2 & r_2 & 1 \\ r_3^2 & r_3 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}$$

The matrix on the left has determinant $(r_1 - r_2)(r_1 - r_3)(r_2 - r_3)$. Since the r_i are distinct, this is nonzero! It follows that the matrix is invertible, and hence there is a unique solution (a, b, c).

(b) Using the reasoning above, the matrix equation we would like to solve for a, b, c is

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 6 \end{bmatrix}$$

To find a least-squares solution, we solve the corresponding equation $A^T A \mathbf{x} = A^T \mathbf{y}$, which boils down to

$$\begin{bmatrix} 98 & 36 & 14 \\ 36 & 14 & 6 \\ 14 & 6 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 62 \\ 22 \\ 10 \end{bmatrix}$$

The least-squares solution is $(a, b, c) = (3/2, -31/10, 19, 10)$. Thus the quadratic curve that "best fits" the data is

$$f(x) = \frac{3}{2}x^2 - \frac{31}{10}x + \frac{19}{10}.$$

It is a "best fit" in the sense that it minimizes the following sum of squared errors:

$$(f(0) - 2)^2 + (f(1) - 0)^2 + (f(2) - 2)^2 + (f(3) - 6)^2.$$

You can think of this sum as a cumulative measure of how far off the function is from the desired outputs for the inputs $x = 0, 1, 2, 3$. In terms of the graph below, f is the quadratic function that minimizes the sum of squares of the vertical distances between the green points and the blue curve.