Mathematics 106 Matrices & Linear Algebra: Solutions

Professor Peter M. Higgins

March 9, 2018

Solutions and Comments

Problem Set 1 Matrices and determinants

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

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

2. $|AB| = (2 \times 2) - (4(-2)) = 4 + 8 = 12$. On the other hand

$$|BA| = 3\begin{vmatrix} 0 & 3 \\ -2 & 1 \end{vmatrix} - 2\begin{vmatrix} -3 & 3 \\ -1 & 1 \end{vmatrix} - 3\begin{vmatrix} -3 & 0 \\ -1 & -2 \end{vmatrix}$$

$$= 3(0 - (-6)) - 2(-3 + 3) - 3(6 - 0) = 18 - 0 - 18 = 0.$$

Comment If we use the fact that rank(AB) =rank(BA) we see that rank$(BA) =$ 2 and, since BA is then not of full rank, $|BA| = 0$. We may also note that for non-square matrices we have an example where $|AB| \neq |BA|$ although for products of square matrices we always have equality.

3. Using Gaussian elimination to solve the equations $a(3, 2, -3)+b(-3, 0, 3) = (-1, -2, 1)$ we have the matrix

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

which gives the equations $2b = -\frac{4}{3} \Rightarrow b = -\frac{2}{3}$ and $3a = 3b+1 = 3(-\frac{2}{3})+1 = -1$. Hence we conclude that

$$-(3, 2, -3) - \frac{2}{3}(-3, 0, 3) = (-1, -2, 1).$$

Comment This also shows that rank$(BA) = 2$ as clearly rank$(BA) > 1$ as no row is a multiple of either of the other rows.

4. By the first row expansion of the determinant we obtain:

$$1(2\times 0-\frac{1}{2}\times 4)-1(3\times 0-\frac{1}{2}\times(-1))+1(3\times 4-2\times(-1)) = -2-\frac{1}{2}+14 = \frac{23}{2}.$$

By the second column expansion we obtain:

$$-1(3\times0-\frac{1}{2}(-1))+2(1\times0-1\times(-1))-4(1\times\frac{1}{2}-1\times3) = -\frac{1}{2}+2-4(-\frac{5}{2}) = \frac{3}{2}+10 = \frac{23}{2}.$$

5. Let

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

Then take the transpose A^T and we find that $AA^T = I_2$, the 2×2 identity matrix. In contrast $A^T A$ is a singular 3×3 matrix with I_2 in the top left hand corner.

6.

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

7.

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

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

$$\to \begin{bmatrix} 1 & 0 & 0 & \frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & -\frac{1}{4} & \frac{3}{4} & \frac{5}{4} \end{bmatrix}.$$

$$\therefore A^{-1} = \frac{1}{4} \begin{bmatrix} 1 & -3 & -1 \\ 1 & 1 & -1 \\ -1 & 3 & 5 \end{bmatrix}.$$

8. We have $|A| = \cos^2\theta + \sin^2\theta = 1$. Hence we may write the inverse down as:

$$A^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

Comment Or we can observe that A induces a rotation about the origin through the angle θ and so A^{-1} is found by replacing θ by $-\theta$ throughout.

9. We use the obvious row operations to reduce the matrix to echelon form:

$$\begin{bmatrix} -1 & -2 & 2 \\ 2 & 1 & 4 \\ 0 & 1 & -1 \\ -3 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} -1 & -2 & 2 \\ 0 & -3 & 9 \\ 0 & 1 & -1 \\ 0 & 8 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & -3 \\ 0 & 1 & -1 \\ 0 & 8 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 2 \\ 0 & 0 & 21 \end{bmatrix} \to$$

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

since there are 3 non-zero rows in the echelon form the rank of the matrix is 3.

Comment Note that since row and column rank of a matrix are always equal, the rank of every 4×3 matrix is no more than 3.

10. Directly we get

$$(a+bi)(a-ib)-(c+id)(id-c)=(a^2+b^2)-(-c^2-d^2)=a^2+b^2+c^2+d^2.$$

It follows that if m and n are both the sum of four squares then mn can be written in the form

$$mn=\begin{vmatrix} z_1 & w_1 \\ -\overline{w_1} & \overline{z_1} \end{vmatrix}\cdot\begin{vmatrix} z_2 & w_2 \\ -\overline{w_2} & \overline{z_2} \end{vmatrix}=\begin{vmatrix} z_1z_2-w_1\overline{w_2} & z_1w_2+w_1\overline{z_2} \\ -\overline{w_1}z_2-\overline{w_2 z_1} & -\overline{w_1}w_2+\overline{z_1 z_2} \end{vmatrix}$$

which has the same form, and so is also a sum of four squares.

Comment Any positive integer m is the sum of four squares. The previous result represents a lemma that reduces this general problem to the case where m is a prime.

Problem Set 2 Systems of Linear Equations

1.

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

$$\begin{bmatrix} 1 & 1 & 1 & -1 & 0 & 2 \\ 0 & 1 & -\frac{1}{2} & -2 & -\frac{1}{2} & -1 \\ 0 & 0 & 1 & 3 & 0 & 1 \end{bmatrix}\to\begin{bmatrix} 1 & 0 & \frac{3}{2} & 1 & \frac{1}{2} & 3 \\ 0 & 1 & -\frac{1}{2} & -2 & -\frac{1}{2} & -1 \\ 0 & 0 & 1 & 3 & 0 & 1 \end{bmatrix}\to\begin{bmatrix} 1 & 0 & 0 & -\frac{7}{2} & \frac{1}{2} & \frac{3}{2} \\ 0 & 1 & 0 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 3 & 0 & 1 \end{bmatrix}$$

We put $x_4=c_1$ and $x_5=c_2$ $(c_1,c_2\in\mathbb{R})$ to yield the solution set described by

$$\begin{aligned}\mathbf{x}=&(\frac{3}{2}+\frac{7}{2}c_1-\frac{1}{2}c_2,\ -\frac{1}{2}+\frac{1}{2}c_1+\frac{1}{2}c_2,\ 1-3c_1,\ c_1,\ c_2)\\ =&(\frac{3}{2},\frac{1}{2},1,0,0)+c_1(\frac{7}{2},\frac{1}{2},-3,1,0)+c_2(-\frac{1}{2},\frac{1}{2},0,0,1).\end{aligned}$$

2.

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

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

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

which yields $x_3 = -1$, $x_4 = 1$,$x_2 = c$, $x_1 = 1 - 2c$ or in vector form the solution set is given by:

$$\mathbf{x} = (1 - 2c,\ c,\ -1, 1) = (1, 0, -1, 1) + c(-2, 1, 0, 0).$$

3.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 3 & 4 & 5 & 0 & 1 & 0 \\ 3 & 6 & 10 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -3 & 1 & 0 \\ 0 & 3 & 7 & -3 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & 4 & -1 & 0 \\ 0 & 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 1 & 6 & -3 & 1 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & 0 & 10 & -4 & 1 \\ 0 & 1 & 0 & -15 & 7 & -2 \\ 0 & 0 & 1 & 6 & -3 & 1 \end{bmatrix}$$

hence

$$A^{-1} = \begin{bmatrix} 10 & -4 & 1 \\ -15 & 7 & -2 \\ 6 & -3 & 1 \end{bmatrix}.$$

4. $|A| = (40 - 30) - (30 - 15) + (18 - 12) = 10 - 15 + 6 = 1.$

$$|A_1| = \begin{vmatrix} 6 & 1 & 1 \\ 22 & 4 & 5 \\ 31 & 6 & 10 \end{vmatrix} = 6(40 - 30) - (220 - 155) + (132 - 124) = 60 - 65 + 8 = 3;$$

$$|A_2| = \begin{vmatrix} 1 & 6 & 1 \\ 3 & 22 & 5 \\ 3 & 31 & 10 \end{vmatrix} = (220 - 155) - 6(30 - 15) + (93 - 66) = 65 - 90 + 27 = 2.$$

$$|A_3| = \begin{vmatrix} 1 & 1 & 6 \\ 3 & 4 & 22 \\ 3 & 6 & 31 \end{vmatrix} = (124 - 132) - (93 - 66) + 6(18 - 12) = -8 - 27 + 36 = 1.$$

Hence $(x, y, z) = (3, 2, 1)$.

5. We have by Question 4 that $|A| = 1$. The entries of the columns of A^{-1} are therefore given by

$$\begin{vmatrix} 1 & 1 & 1 \\ 0 & 4 & 5 \\ 0 & 6 & 10 \end{vmatrix} = 40{-}30 = 10, \begin{vmatrix} 1 & 1 & 1 \\ 3 & 0 & 5 \\ 3 & 0 & 10 \end{vmatrix} = -(30{-}15) = -15, \begin{vmatrix} 1 & 1 & 1 \\ 3 & 4 & 0 \\ 3 & 6 & 0 \end{vmatrix} = 18{-}12 = 6;$$

$$\begin{vmatrix} 0 & 1 & 1 \\ 1 & 4 & 5 \\ 0 & 6 & 10 \end{vmatrix} = -(10{-}6) = -4, \begin{vmatrix} 1 & 0 & 1 \\ 3 & 1 & 5 \\ 3 & 0 & 10 \end{vmatrix} = 10{-}3 = 7, \begin{vmatrix} 1 & 1 & 0 \\ 3 & 4 & 1 \\ 3 & 6 & 0 \end{vmatrix} = -(6{-}3) = -3;$$

$$\begin{vmatrix} 0 & 1 & 1 \\ 0 & 4 & 5 \\ 1 & 6 & 10 \end{vmatrix} = 5 - 4 = 1, \begin{vmatrix} 1 & 0 & 1 \\ 3 & 0 & 5 \\ 3 & 1 & 10 \end{vmatrix} = -(5 - 3) = -2, \begin{vmatrix} 1 & 1 & 0 \\ 3 & 4 & 0 \\ 3 & 6 & 1 \end{vmatrix} = 4 - 3 = 1.$$

$$\therefore A^{-1} = \begin{bmatrix} 10 & -4 & 1 \\ -15 & 7 & -2 \\ 6 & -3 & 1 \end{bmatrix}.$$

6. Observe that the determinant of an integer matrix is an integer. Hence if $|A| = \pm 1$ it follows from Cramer's rule that all the entries of A^{-1} are also integers. Conversely, if both$A$ and A^{-1} are integer matrices then both $|A|$ and $|A^{-1}| = |A|^{-1}$ are both integers and hence $|A| = \pm 1$. Therefore an integer matrix A has an inverse that is also an integer matrix if and only if $|A| = \pm 1$.

7. If $|A| \neq 0$ then, by Cramer's rule, A^{-1} may be calculated. Hence if A has no inverse then $|A| = 0$. Conversely, since $|A^{-1}| = |A|^{-1}$whence it follows that if A possesses an inverse then $|A| \neq 0$. Therefore A^{-1} exists if and only if $|A| \neq 0$, that is to say A is *non-singular*.

8. Note first that $S \neq \emptyset$ as $\mathbf{0} \in S$. Now let $\mathbf{a} \in S$ so that $A\mathbf{a} = \mathbf{b}$ and let $\mathbf{y} \in S$. Then

$$A(\mathbf{a} + \mathbf{y}) = A\mathbf{a} + A\mathbf{y} = \mathbf{b} + \mathbf{0} = \mathbf{b},$$

from which it follows that all members of $\mathbf{a} + S$ are indeed solutions to $A\mathbf{x} = \mathbf{b}$.

Conversely suppose that $\mathbf{x}$ is a solution to our system. Then $\mathbf{x} = \mathbf{a} + (\mathbf{x} - \mathbf{a})$ and so to complete the proof we need only check that $\mathbf{x} - \mathbf{a} \in S$. To this end:

$$A(\mathbf{x} - \mathbf{a}) = A\mathbf{x} - A\mathbf{a} = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

9. Suppose that the system has two distinct solutions $\mathbf{a}$ and $\mathbf{c}$ say. Then let $\mathbf{y} = \lambda(\mathbf{a} - \mathbf{c})$. $(\lambda \neq 0)$. We see that

$$A\mathbf{y} = A\lambda(\mathbf{a} - \mathbf{c}) = \lambda A\mathbf{a} - \lambda A\mathbf{c} = \lambda\mathbf{b} - \lambda\mathbf{b} = \mathbf{0}.$$

Since $\mathbf{a} - \mathbf{c} \neq \mathbf{0}$ it follows that the solution set S of the homogeneous system $A\mathbf{x} = \mathbf{0}$ is infinite. Therefore so is the solution set $\mathbf{a} + S$ of the system $A\mathbf{x} = \mathbf{b}$.

10. By Question 8, A is non-singular if and only if A^{-1} exists, in which case we have

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

and in particular $A^{-1}\mathbf{b}$ is the unique solution of our system.

Conversely, since the solution set of $A\mathbf{x} = \mathbf{b}$ has the form $S + \mathbf{a}$ where S is the solution set of $A\mathbf{x} = \mathbf{0}$. Hence if one system has a unique solution then so does the other, in which case $\mathbf{0}$ is the unique solution of $A\mathbf{x} = \mathbf{0}$. It follows that $\lambda = 0$ is not an eigenvalue of A, which is equivalent to the statement that $|A| \neq 0$.

Problem Set 3 Eigenvalues and Eigenvectors

1. Our characteristic equation is

$$0 = (-17-\lambda)(18-\lambda) + 300 = \lambda^2 - \lambda - 6 = (\lambda-3)(\lambda+2)$$

and so, in ascending order, the eigenvalues are -2 and 3.
2. For $\lambda = 2$, applying row reduction we obtain:

$$A - \lambda I = \begin{bmatrix} -15 & 30 \\ -10 & 20 \end{bmatrix} \to \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}$$

giving the eigenvector $(2,1)^T$ while for $\lambda = 3$ we get:

$$A - \lambda I = \begin{bmatrix} -20 & 30 \\ -10 & 20 \end{bmatrix} \to \begin{bmatrix} 1 & -\frac{3}{2} \\ 0 & 0 \end{bmatrix}$$

which gives the eigenvector $(3,2)^T$.
3. From Question 2, our diagonalizing matrix is:

$$P = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \Rightarrow P^{-1} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$$

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

4.

$$A^6 = (PDP^{-1})^6 = PD^6P^{-1} =$$

$$\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 64 & 0 \\ 0 & 729 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -6305 & 12738 \\ -12994 & 26052 \end{bmatrix}.$$

5. Suppose that $A\mathbf{x} = \lambda\mathbf{x}$ for some non-zero vector $\mathbf{x}$. Let $B = PAP^{-1}$ say by similar to A. Then $P\mathbf{x}$ is an eigenvector of B with eigenvalue λ:

$$B(P\mathbf{x}) = (PAP^{-1})P\mathbf{x} = PA\mathbf{x} = P\lambda\mathbf{x}$$

$$= \lambda P\mathbf{x}.$$

Since P is not singular it follows that $P\mathbf{x}$ is not the zero vector. Also since $A = P^{-1}B(P^{-1})^{-1}$, it follows that A is similar to B and so, by the same argument, each eigenvalue of B is also an eigenvalue of A. Therefore similar matrices share the same eigenvalue set.
6. We note that

$$BA = A^{-1}(AB)A, \; AB = A(BA)A^{-1}$$

and so by symmetry, if either of A and B is non-singular, then AB are similar matrices.

Next take

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

Then

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = A, \; BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = Z;$$

Now for all invertible 2×2 matrices P , since $P^{-1}ZP = Z$ and $A \neq Z$ it follows that AB and BA are not only unequal but are not similar.

7. Write $A = [a_{ij}]$ and $B = [b_{ij}]$ say. Then

$$\text{tr}(AB) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ji} = \sum_{j=1}^{n} \sum_{i=1}^{n} b_{ji} a_{ij} = \text{tr}(BA).$$

8. Let A and B be similar with $B = PAP^{-1}$ say. Then, using Question 7 we have

$$\text{tr}(B) = \text{tr}((PA)P^{-1})) = \text{tr}(AP^{-1}P) = \text{tr}(A).$$

9. Write A_{ij} for the matrix obtained from A by deleting the ith row and jth column of A. Taking the first row expansion of A we have

$$|A| = a_{11}|A_{11}| = a_{11}\Pi_{k=2}^{n} a_{kk} = \Pi_{k=1}^{n} a_{kk},$$

where the second equality above follows by induction (the $n = 1$ case being trivially true).

10. Consider $B = A - a_{kk}I$. Then B is also an upper triangular matrix with $b_{kk} = 0$. We show by induction on n that for such a matrix B, $|B| = 0$ from which it follows that each a_{kk} is an eigenvalue of A. The claim is clear for $n = 1$ so suppose that $n \geq 2$. If $k = 1$ so that $b_{11} = 0$ then the first column of B is a zero column and so $|B| = 0$. Otherwise $k \geq 2$ and we have

$$|B| = b_{11}|B_{11}| = b_{11} \cdot 0 = 0$$

where the second equality follows by induction as B_{11} is an $(n-1) \times (n-1)$ upper triangular matrix with diagonal entry $b_{kk} = 0$.

Conversely let $\mu \in \mathbb{R}$ be such that $\mu \neq a_{kk}$ for all $k = 1, 2, \cdots, n$. Then $A - \mu I$ is an upper triangular matrix and so by Question 9 we have

$$|A - \mu I| = \Pi_{k=1}^{n}(a_{kk} - \mu) \neq 0$$

as all terms in the product are non-zero. Hence μ is not an eigenvalue of A and we conclude that the eigenvalues of A are exactly the entries on its main diagonal.

Problem Set 4 Eigenvalues and eigenvectors application

1. From the definition we find that $\mathbf{x_0} = (1, 0, 0)^T$, $\mathbf{x_1} = (2, 1, 1)^T$.

2. Since $3 \equiv 9 \equiv 0 \pmod 3$ and $7 \equiv 1 \pmod 3$, $2 \equiv 2 \pmod 3$ we get the recurrence equations:

$$x_n = 2x_{n-1}+y_n+z_n,\ y_n = x_{n-1}+2y_{n-1}+z_{n-1},\ z_n = x_{n-1}+y_{n-1}+2z_{n-1}\ (n \geq 1).$$

3. $\mathbf{x_n} = A\mathbf{x_{n-1}}$ where

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

and n-fold repetition gives the equation $\mathbf{x_n} = A^n\mathbf{x_0}$.

4. Put $|A - \lambda I| = 0$ to give

$$(2-\lambda)((2-\lambda)^2 - 1)) - ((2-\lambda) - 1) + (1 - (2-\lambda)) = 0$$

$$\Rightarrow (2-\lambda)^3 - 2 + \lambda - 2 + \lambda + 1 - 1 + \lambda = 8 - 12\lambda + 6\lambda^2 - \lambda^3 + 3\lambda - 4 = 0$$

$$\Rightarrow \lambda^3 - 6\lambda^2 + 9\lambda - 4 = (\lambda - 1)^2(\lambda - 4) = 0.$$

Hence the eigenvalues of A are $\lambda = 1$ and $\lambda = 4$.

5. For $\lambda = 1$ we obtain:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \mathbf{x} = \mathbf{0} \Rightarrow (1, 0, -1)^T,\ (0, 1, -1)^T \text{ independent eigenvectors,}$$

for $\lambda = 4$ we obtain:

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \mathbf{x} = \mathbf{0} \Rightarrow (1, 1, 1)^T \text{ is an eigenvector.}$$

6. $A = PDP^{-1}$ where

$$P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{bmatrix},\ D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \text{ and by a standard calculation we also get}$$

$$P^{-1} = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

7.

$$A^n = PD^nP^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4^n \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} =$$

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

8. We require $A^n\mathbf{x_0}$ which is:

$$\mathbf{x_n} = \frac{1}{3}\begin{bmatrix} 4^n+2 \\ 4^n-1 \\ 4^n-1 \end{bmatrix}.$$

9. Question 8 gives the correct result $(x_0, y_0, z_0) = (1,0,0)$ (and $(x_1, y_1, z_1) = (2,1,1)$). By induction we then get

$$x_n = 2x_{n-1} + y_{n-1} + z_{n-1} = \frac{1}{3}((2\cdot 4^{n-1} + 2\cdot 2) + (4^{n-1}-1) + (4^{n-1}-1))$$

$$= \frac{1}{3}\left(4\cdot 4^{n-1}+2\right) = \frac{1}{3}\left(4^n+2\right) \text{ and so the induction continues.}$$

Similar inductive calculations for y_n and z_n give the result.

10.

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

Hence

$$\frac{(tA)^n}{n!} = \frac{t^n}{n!}I_2 \text{ if } n \equiv 0 (\text{mod } 4),\ \frac{(tA)^n}{n!} = \frac{t^n}{n!}A \text{ if } n \equiv 1 \ (\text{mod } 4)$$

$$\frac{(tA)^n}{n!} = \frac{t^n}{n!}A^2 \text{ if } n \equiv 2 (\text{mod } 4),\ \frac{(tA)^n}{n!} = \frac{t^n}{n!}A^3 \text{ if } n \equiv 3 \ (\text{mod } 4).$$

Hence the entries at $(1,1)$, $(1,2)$,$(2,1)$ and $(2,2)$ in e^{tA} respectively are:

$$\sum_{n=0}^{\infty}\frac{t^{4n}}{(4n)!} - \sum_{n=0}^{\infty}\frac{t^{4n+2}}{(4n+2)!} = \sum_{n=0}^{\infty}\frac{(-1)^n t^{2n}}{(2n)!} = \cos t,$$

$$\sum_{n=0}^{\infty}\frac{-t^{4n+1}}{(4n+1)!} + \sum_{n=0}^{\infty}\frac{t^{4n+3}}{(4n+3)!} = -\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n+1}}{(2n+1)!} = -\sin t,$$

$$\sum_{n=0}^{\infty}\frac{t^{4n+1}}{(4n+1)!} - \sum_{n=0}^{\infty}\frac{t^{4n+3}}{(4n+3)!} = \sum_{n=0}^{\infty}\frac{(-1)^n t^{2n+1}}{(2n+1)!} = \sin t,$$

$$\sum_{n=0}^{\infty}\frac{t^{4n}}{(4n)!} - \sum_{n=0}^{\infty}\frac{t^{4n+2}}{(4n+2)!} = \sum_{n=0}^{\infty}\frac{(-1)^n t^{2n}}{(2n)!} = \cos t.$$

$$\therefore e^{tA} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}.$$

Problem Set 5 Positive definite matrices and quadratic forms

1 & 2. Clearly I_n is symmetric and equal to its conjugate transpose. Moreover for any $\mathbf{z} = (z_1, \cdots, z_n) \in \mathbb{C}^n$ we have

$$z^{*T} I_n \mathbf{z} = \mathbf{z}^{*T}\mathbf{z} = \overline{z_1} z_1 + \cdots + \overline{z}_n z = |z_1|^2 + \cdots + |z_n|^2 > 0,$$

as $\mathbf{z} \neq \mathbf{0}$, as required.

3. Let $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$. Then

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x - y & -x + 2y - z & -y + 2z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} =$$

$$= (2x^2 - xy) + (-xy + 2y^2 - yz) + (-yz + 2z^2) = 2(x^2 + y^2 + z^2) - (2xy + 2yz)$$
$$= x^2 + z^2 + (x - y)^2 + (y - z)^2 > 0.$$

4. Let $\mathbf{x} = (x, y)$. Then

$$\mathbf{x}^T A\mathbf{x} = x^2 + 4xy + y^2 = (x + y)^2 + 2xy.$$

Hence if we take $y = -x \neq 0$ then the outcome will be negative and so A is not positive definite. In particular, taking $\mathbf{x} = [1\ -1]^T$ gives $\mathbf{x}^T A\mathbf{x} = -2 < 0$.

5. Suppose that $\mathbf{x}$ is an eigenvector of the positive definite matrix A with eigenvalue λ. Then

$$0 < \mathbf{x}^T A\mathbf{x} = \mathbf{x}^T(\lambda\mathbf{x}) = \lambda\mathbf{x}^T\mathbf{x}.$$

Since $\mathbf{x}^T\mathbf{x} > 0$ it follows that $\lambda > 0$.

6&7. Let the entries of A be a, b, c and d in the usual way. Equating $\mathbf{x}^T A\mathbf{x}$ with $Q(x, y)$ then gives

$$ax^2 + (b + c)xy + dy^2 = 5x^2 - 10xy + y^2.$$

Hence we need $a = 5$, $d = 1$ and b and c can be any numbers subject to the constraint that $b + c = -10$. In Question 7 we demand that A is symmetric so that $b = c = -5$. Hence the unique symmetric solution to our equation is:

$$A = \begin{bmatrix} 5 & -5 \\ -5 & 1 \end{bmatrix}.$$

8. In general, taking $M = (a_{ij})$ and writing the coefficient of $x_i x_j$ in $q(x_1, x_2, \cdots, x_n)$ to be b_{ij} we see that $b_{ij} = a_{ij} + a_{ji}$. Hence we can take M to be symmetric by putting $a_{ij} = a_{ji} = \frac{b_{ij}}{2}$.

9. Using the fact that A is symmetric and that the transpose of a constant is itself we have the sequence of equalities:

$$\mathbf{v}_1^T A\mathbf{v}_2 = \mathbf{v}_1^T \lambda_2 \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \bullet \mathbf{v}_2 = \lambda_2 \mathbf{v}_2 \bullet \mathbf{v}_1$$

$$= (A\mathbf{v}_2)^T\mathbf{v}_1 = \mathbf{v}_2^T A^T\mathbf{v}_1 = \mathbf{v}_2^T A\mathbf{v}_1 = \mathbf{v}_2^T\lambda_1\mathbf{v}_1$$
$$= \lambda_1\mathbf{v}_2 \bullet \mathbf{v}_1 = \lambda_1\mathbf{v}_1 \bullet \mathbf{v}_2$$

and so we conclude that

$$\lambda_1(\mathbf{v}_1 \bullet \mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \bullet \mathbf{v}_2).$$

10. We deduce at once from the result of Question 9 that

$$(\lambda_1 - \lambda_2)\mathbf{v}_1 \bullet \mathbf{v}_2 = 0$$

and so given that $\lambda_1 \neq \lambda_2$ we may cancel that factor to conclude that $\mathbf{v}_1 \bullet \mathbf{v}_2 = 0$, which is to say that eigenvectors of distinct eigenvalues of a symmetric matrix are mutually orthogonal.

Problem Set 6 Matrices and Analytical geometry

1. In two dimensions, the standard basis vectors are $\mathbf{u} = (1, 0)^T$ and $\mathbf{v} = (0, 1)^T$. We get the columns of A and of A^{-1} respectively by rotating each of them in turn through $-30°$ and $30°$ to give the transformation matrices:

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

2. This is a linear transformation of the plane and so the columns of the transformation matrix A are the images of the respective standard basis vectors, $\mathbf{u} = \binom{1}{0}$ and $\mathbf{v} = \binom{0}{1}$ under the mapping. We see that $\mathbf{u}$ maps onto the unit vector making an angle of $2 \times \frac{\pi}{8} = \frac{\pi}{4}$ with the axes, which is $\frac{1}{\sqrt{2}}\binom{1}{1}$, while reflecting $\mathbf{v}$ in the line $\theta = \frac{\pi}{8}$ maps it into the direction $\theta = \frac{\pi}{8} - \left(\frac{\pi}{2} - \frac{\pi}{8}\right) = -\frac{\pi}{4}$, whence $\mathbf{v}$ is mapped to $\frac{1}{\sqrt{2}}\binom{1}{-1}$. Hence the required matrix is:

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

3. The columns of A are the respective images of the standard basis vectors $\binom{1}{0}$ and $\binom{0}{1}$. The former is mapped by this *shearing* transformation to $\binom{1}{-5}$ while the latter is *invariant* (fixed) by this linear transformation.

$$A = \begin{bmatrix} 1 & 0 \\ -5 & 1 \end{bmatrix}.$$

4. Following the procedure of the first two questions, we see that in general the matrix for a rotation of θ anti-clockwise about the origin is given by:

$$\text{Rot}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Geometrically it is clear that $\mathrm{Rot}(\theta)\mathrm{Rot}(\phi) = \mathrm{Rot}(\theta + \phi)$. This can be verified algebraically as well:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} \cos\theta\sin\phi - \sin\theta\sin\phi & -\cos\theta\sin\phi - \sin\theta - \sin\theta\cos\phi \\ \sin\theta\cos\phi + \cos\theta\sin\phi & -\sin\theta\sin\phi + \cos\theta\sin\phi + \cos\theta\cos\phi \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix}.$$

5. And for reflection in the line making an angle θ:

$$\mathrm{Ref}(\theta) = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix};$$

for example, to find the second column entries we need the coodinates of the tip of the position vector $\mathbf{v} = \binom{\mathbf{0}}{\mathbf{1}}$ after reflection in the line. For $0 \le \theta \le \frac{\pi}{2}$, the angle of the image vector is $\frac{\pi}{2} - 2(\frac{\pi}{2} - \theta) = 2\theta - \frac{\pi}{2}$. Hence the respsective x- and y-coordinates are:

$$\cos(2\theta - \frac{\pi}{2}) = \cos 2\theta \cos(\frac{\pi}{2}) + \sin 2\theta \sin(\frac{\pi}{2}) = \sin 2\theta;$$

$$\sin(2\theta - \frac{\pi}{2}) = \sin 2\theta \cos(\frac{\pi}{2}) - \cos 2\theta \sin\frac{\pi}{2} = -\cos 2\theta.$$

For $\frac{\pi}{2} \le \theta \le \pi$ the angle of the image vector is $\frac{\pi}{2} + 2(\theta - \frac{\pi}{2}) = 2\theta - \frac{\pi}{2}$, which yields the same result.

6. The transformation matrix is the product:

$$\begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos 2\phi\cos\theta + \sin 2\phi\sin\theta & -\cos 2\phi\sin\theta + \sin 2\phi\cos\theta \\ \sin 2\phi\cos\theta - \cos 2\phi\sin\theta & -\sin 2\phi\sin\theta - \cos 2\phi\cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos(2\phi - \theta) & \sin(2\phi - \theta) \\ \sin(2\phi - \theta) & -\cos(2\phi - \theta) \end{bmatrix},$$

which is the transformation of $\mathrm{Ref}(\phi - \frac{\theta}{2})$.

7. In this case we put $\theta = -\frac{\pi}{6}$ and $\phi = \frac{\pi}{8}$ so the product of the corresponding matrices represents the transformation $\mathrm{Ref}(\phi - \frac{\theta}{2}) = \mathrm{Ref}(\frac{\pi}{8} + \frac{\pi}{12}) = \mathrm{Ref}(\frac{5\pi}{24})$. The matrix of the transformation is

$$\frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} \sqrt{6} - \sqrt{2} & \sqrt{6} + \sqrt{2} \\ \sqrt{6} + \sqrt{2} & \sqrt{2} - \sqrt{6} \end{bmatrix},$$

and therefore we may infer that the respective values of $\cos(2\phi - \theta) = \cos\frac{5\pi}{12}$ and $\sin\frac{5\pi}{12}$:

$$\cos\frac{5\pi}{12} = \frac{\sqrt{6} - \sqrt{2}}{4}, \ \sin\frac{5\pi}{12} = \frac{\sqrt{6} + \sqrt{2}}{4}.$$

Comment These values can be verified directly. For example:

$$\cos\frac{5\pi}{12} = \cos(\frac{\pi}{4} + \frac{\pi}{6}) = \cos\frac{\pi}{4}\cos\frac{\pi}{6} - \sin\frac{\pi}{4}\sin\frac{\pi}{6}$$

$$= \frac{1}{\sqrt{2}}(\frac{\sqrt{3}}{2} - \frac{1}{2}) = \frac{\sqrt{6}-\sqrt{2}}{4}.$$

8. Reversing the order of the product from Question 6 gives

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{bmatrix} = \begin{bmatrix} \cos\theta\cos 2\phi - \sin\theta\sin 2\phi & \cos\theta\sin 2\phi + \sin\theta\cos 2\phi \\ \sin\theta\cos 2\phi + \cos\theta\sin 2\phi & \sin\theta\sin 2\phi - \cos\theta\cos 2\phi \end{bmatrix}$$

$$= \begin{bmatrix} \cos(2\phi+\theta) & \sin(2\phi+\theta) \\ \sin(2\phi+\theta) & -\cos(2\phi+\theta) \end{bmatrix},$$

which is the matrix of $\mathrm{Ref}(\phi + \frac{\theta}{2})$.

9. By Question 5, the matrix for $\mathrm{Ref}(\phi)\mathrm{Ref}(\theta)$ is

$$\begin{bmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{bmatrix} \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} = \begin{bmatrix} \cos 2\phi\cos 2\theta + \sin 2\phi\cos 2\theta & \cos 2\phi\sin 2\theta - \sin 2\phi\cos 2\theta \\ \sin 2\phi\cos 2\theta - \cos 2\phi\sin 2\theta & \sin 2\phi\sin 2\theta + \cos 2\phi\cos 2\theta \end{bmatrix}$$

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

Hence $\mathrm{Ref}(\phi)\mathrm{Ref}(\theta) = \mathrm{Rot}(2(\phi-\theta))$.

10.

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

The inverse A^{-1} will represent the linear transformation such that $\mathbf{i} \mapsto \mathbf{k} \mapsto \mathbf{j}$, which gives:

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

To find the eigenvalues λ we solve $|A - \lambda I| = 0 \Leftrightarrow$

$$\begin{vmatrix} -\lambda & 0 & 1 \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{vmatrix} = 0;$$

Expanding by the first row gives the equation:

$$\Leftrightarrow -\lambda(\lambda^2 - 0) - 0 + (1 - 0) = 0 \Leftrightarrow \lambda^3 = 1 \Leftrightarrow \lambda = 1.$$

To find the corresponding eigenspace we solve $(A - I)\mathbf{x} = \mathbf{0}$, which yields the equations $-x + z = x - y = y - z = 0 \Leftrightarrow x = y = z$. Hence eigenvectors are the non-zero multiples of $\mathbf{x} = (1, 1, 1)^T$.

Comment The action of this linear transformation is that of rotation through $\frac{2\pi}{3}$ about the axis $x = y = z$. Any linear mapping that is a rotation will fix its axis of rotation pointwise and so must have 1 as an eigenvalue and the axis of rotation as the corresponding eigenspace, with no other eigenvectors.

Problem Set 7 Linear Independence and bases

1. We determine the rank of the matrix of vectors.

$$\begin{bmatrix} -1 & -2 & 2 \\ 2 & 1 & 4 \\ 0 & 1 & -1 \\ -3 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} -1 & -2 & 2 \\ 0 & 1 & 6 \\ 0 & 1 & -1 \\ 0 & 6 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & 6 \\ 0 & 0 & -7 \\ 0 & 0 & 3 \end{bmatrix},$$

which is rank 3 so the given set of 3 vectors is independent.

2.

$$M = \begin{bmatrix} 1 & -1 & 3 & 2 \\ -1 & 3 & -2 & 2 \\ 2 & 1 & 2 & -1 \\ -1 & 0 & 2 & 7 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 3 & 2 \\ 0 & 2 & 1 & 4 \\ 0 & 3 & -4 & -5 \\ 0 & -1 & 5 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 3 & 2 \\ 0 & 1 & -5 & -9 \\ 0 & 2 & 1 & 4 \\ 0 & 3 & -4 & -5 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & -1 & 3 & 2 \\ 0 & 1 & -5 & 9 \\ 0 & 0 & 11 & 22 \\ 0 & 0 & 11 & 22 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 3 & 2 \\ 0 & 1 & -5 & 9 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence the subspace has dimension 3 and has a basis $\{(1, -1, 3, 2), (0, 1, -5, 9), (0, 0, 1, 2)\}$.

3. Put $x_2 = c_1$, $x_3 = c_2$, $x_4 = c_3$ so that the solution of the equation of the hyperlane is

$$\mathbf{x} = (\frac{3}{2}c_1 - 2c_2 + \frac{1}{2}c_3, c_1, c_2, c_3)$$

$$= c_1(\frac{3}{2}, 1, 0, 0) + c_2(-2, 0, 1, 0) + c_3(\frac{1}{2}, 0, 0, 1).$$

Therefore a basis for the hyperplane is, for example

$$\{(3, 2, 0, 0), (-2, 0, 1, 0), (1, 0, 0, 2)\}.$$

4. Suppose that

$$\mathbf{u} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k = b_1\mathbf{u}_1 + b_2\mathbf{u}_2 + \cdots + b_k\mathbf{u}_k$$

$$\Rightarrow (a_1 - b_1)\mathbf{u}_1 + (a_2 - b_2)\mathbf{u}_2 + \cdots + (a_k - b_k)\mathbf{u}_k = \mathbf{0},$$

and since $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k\}$ is an independent set, $a_1 - b_1 = a_2 - b_2 = \cdots = a_k - b_k = 0$, which is to say that $a_1 = b_1, a_2 = b_2 \cdots, a_k = b_k$.

5. Let $A = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k\}$ be a finite subset of a vector space V and suppose that $\mathbf{u}_j = a_1\mathbf{u}_1 + \cdots + a_i\mathbf{u}_i$ for some $1 \leq i < j \leq k$. Then

$$a_1\mathbf{u}_1 + \cdots + a_i\mathbf{u}_i + 0\mathbf{u}_{i+1} + \cdots + 0\mathbf{u}_{j-1} - \mathbf{u}_j + 0\mathbf{u}_{j+1} + \cdots + 0\mathbf{u}_k = \mathbf{0}$$

showing that A is not linearly independent. It follows by the contrapositive that if A is linearly independent then no member of A is a linear combination of its predecessors.

Conversely, suppose that A is not linearly independent so that we have

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k = \mathbf{0}$$

with not all the a_i equal to zero. Let j be the greatest subscript i such that $a_i \neq 0$. Note that $j \geq 2$ as $\mathbf{u}_1 \neq \mathbf{0}$. Then

$$\mathbf{u}_j = -\frac{a_1}{a_j}\mathbf{u}_1 - \frac{a_2}{a_j}\mathbf{u}_2 - \cdots - \frac{a_{j-1}}{a_j}\mathbf{u}_{j-1}$$

so that $\mathbf{u}_j$ is a linear combination of its predecessors in the list. Hence it follows that if no member of A is a linear combination of the predecessors in the list, then A is independent.

6. Let $R = \langle \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n \rangle$ where $A = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n\}$ is the set of row vectors of M. Swapping two rows does not alter the generating set A of R and so does not alter the subspace generated by that set. Suppose next that $\mathbf{r}_i$ is replaced by $a\mathbf{r}_i$ for some $a \neq 0$. Then for any linear combination $\mathbf{u} = a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + a_n\mathbf{r}_n$ of the members of A we have

$$\mathbf{u} = a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + \frac{a_i}{a}(a\mathbf{r}_i) + \cdots + a_n\mathbf{r}_n$$

so that $R \subseteq S = \langle \mathbf{r}_1, \mathbf{r}_2, \cdots, a\mathbf{r}_i, \cdots, \mathbf{r}_n \rangle$. By the same argument, replacing the generator $a\mathbf{r}_i$ by $\mathbf{r}_i$ shows that $S \subseteq R$. Therefore $R = S$ so that the row space is unchanged by this row operation.

Finally, let us replace the generator $\mathbf{r}_j$ of A by $\mathbf{r} = \mathbf{r_j} + a\mathbf{r_i}$, again calling the subspace generated S . Then since $\mathbf{r} \in R$ it follows that $S \subseteq R$. However, since $\mathbf{r}_j = \mathbf{r} - a\mathbf{r_i}$ it follows in the same way that each member of A lies in the subspace S so that $R \subseteq S$ and once more we have the required equality $R = S$ of subspaces.

7. By Question 6, the row space R of M and of its echelon form E are the same. The dimension of R equals the size of any maximal independent set of rows of either matrix as any such set is a basis for R. Clearly E is spanned by its non-zero rows, which number $m-k$, so that the common row rank of M and E is $m-k$.

8. Each non-zero row of the echelon form E of M has one pivotal 1. Consider the corresponding set B of columns of E. Each such column C has a unique non-zero entry in its pivotal 1 in row i say. Since every other member of B has a zero in row i, it follows that C is not a linear combination of the other columns. Since this applies to each member of B, it follows from Question 7 that B is an independent set.

9. From Question 8 it follows that the column rank of E is at least $m-k$, the common value of the row rank of E and M. In particular the row rank of M is less than or equal to the column rank of M.

10. By Question 9 we have that

$$\text{colrank}M = \text{rowrank}(M^T) \leq \text{colrank}(M^T) = \text{rowrank}(M);$$

therefore it follows from this and Question 9 that the row rank and column ranks of M are equal and that their common value is $m-k$, the number of non-zero rows in the echelon from of M.

Comment This common value can therefore be referred to simply as the *rank* of the matrix M.

Problem Set 8

1.
$$L(\mathbf{0}) = L(\mathbf{0}+\mathbf{0}) = L(\mathbf{0}) + L(\mathbf{0})$$
whereupon, subtracting $L(\mathbf{0})$ from both sides gives $L(\mathbf{0}) = \mathbf{0}$.

Comment Bear in mind that this is saying that $L(\mathbf{0}_U) = \mathbf{0}_v$, the zeros of the respective domain and range spaces.

2. The case where $k=2$ is given and the $k=1$ case follows from this by taking $a_2 = 0$. Hence assume that $k \geq 3$. We then bracket as below and apply the $k=2$ case, and then the inductive hypothesis for the $k-1$ case as follows:

$$L(a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k) = L(a_1\mathbf{u}_1 + (a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k))$$
$$= a_1L(\mathbf{u}_1) + L(a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k)$$
$$= a_1L(\mathbf{u}_1) + a_2L(\mathbf{u}_2) + \cdots + a_kL(\mathbf{u}_k).$$

Comment In the same way we can show that if a set is closed under the taking of linear combinations of two vectors, then the same is true of arbitrary linear combinations.

3. By Question 1, $L(\mathbf{0}) = \mathbf{0}$ so that the kernel of L is not empty. Suppose now that $L(\mathbf{u}) = L(\mathbf{v}) = \mathbf{0}$ and that a, b are scalars. Then

$$L(a\mathbf{u} + b\mathbf{v}) = aL(\mathbf{u}) + bL(\mathbf{v}) = a\mathbf{0}+b\mathbf{0} = \mathbf{0}+\mathbf{0} = \mathbf{0}.$$

Hence $\ker(L)$ is a subspace of the domain space U.

4. By Question 1, the range set $L(U)$ contains $\mathbf{0}_V$ and so is not empty. Let $\mathbf{x}, \mathbf{y} \in L(U)$ so that $\mathbf{x} = L(\mathbf{u})$ and $\mathbf{y} = L(\mathbf{v})$ say and let a, b be scalars. Then

$$L(a\mathbf{u} + b\mathbf{v}) = aL(\mathbf{u}) + bL(\mathbf{v}) = a\mathbf{x} + b\mathbf{y}$$

and so $L(U)$ is closed under the taking of linear combinations and is therefore a subspace of the codomain space V.

5. Suppose that L is one-to-one and that $\mathbf{u}, \in \ker(L)$. Then by Question 1 we then have $L(\mathbf{u}) = L(\mathbf{0}) = \mathbf{0}$ and since L is injective (one-to-one) it follows that $\mathbf{u} = \mathbf{0}$. Therefore if L is injective then $\ker(L) = \{\mathbf{0}\}$. Conversely suppose that $L(\mathbf{u}) = L(\mathbf{v})$. Then by linearity of L we hav that $L(\mathbf{u}-\mathbf{v}) = L(\mathbf{u}) - L(\mathbf{v}) = \mathbf{0}$ so that $\mathbf{u}-\mathbf{v} \in \ker(L)$. If we suppose now that $\ker(L)$ is the trivial subspace $\{\mathbf{0}\}$ then we have $\mathbf{u}-\mathbf{v} = \mathbf{0}$ so that $\mathbf{u} = \mathbf{v}$ and therefore L is one-to-one. In

conclusion a linear mapping L is one-to-one if and only if the kernel of L is trivial.

6. A typical member of the range is $M\mathbf{u}$ where $\mathbf{u}^T = (u_1, u_2, \cdots, u_n)$ say. We then have that the ith entry of the vector $M\mathbf{u}$ equals $a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{im}u_n$ and so, writing $\mathbf{c}_j$ for the jth column of M we infer that:

$$M\mathbf{u} = u_1\mathbf{c}_1 + u_2\mathbf{c}_2 + \cdots + u_n\mathbf{c}_n.$$

Hence the range space $L(U)$ is contained in the span of the columns of M. Conversely taking $u_i = 1$ and $u_j = 0$ for all $j \neq i$ we see that each $\mathbf{c}_j$ lies in $L(U)$. Therefore the columns of M form a aspanning set for the range space $L(U)$.

7. Let $\mathbf{c}_j = L(\mathbf{e}_j)$, where $\mathbf{e}_j$ is the jth standard basis vector of $\mathbb{R}^n$. Let M be the $m \times n$ matrix whose jth column is $\mathbf{c}_j$. Then we have using Questions 5 and 6:

$$L(\mathbf{u}) = L(u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \cdots + u_n\mathbf{e}_n) = u_1L(\mathbf{e_1}) + u_2L(\mathbf{e}_2) + \cdots + u_nL(\mathbf{e}_n)$$

$$= u_1\mathbf{c}_1 + u_2\mathbf{c}_2 + \cdots + u_n\mathbf{c}_n = M(\mathbf{u}).$$

Therefore the action of L is that of the matrix M the columns of which are the images of the each of the standard basis vectors taken in the natural order.

8. Since U and W are subspaces of V we certainly have $\mathbf{0} \in U \cap W$. Let $\mathbf{u}, \mathbf{v} \in U \cap W$ and take scalars a, b. Then, again since each is a subspace, $a\mathbf{u} + b\mathbf{v} \in U \cap W$, whence it follows that $U \cap W$ is a subspace of V.

9. We have $A \subseteq \langle B \rangle$. Take an arbitrary member of $\langle A \rangle$, which may be written as $\mathbf{u} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n$ where each $\mathbf{u}_i \in A$. Then each $\mathbf{u}_i \in \langle B \rangle$ and since $\langle B \rangle$ is a subspace it is closed under the taking of arbitrary linear combinations so that, in particular, $\mathbf{u} \in \langle B \rangle$. Since $\mathbf{u}$ was an arbitrary member of $\langle A \rangle$ it follows that $\langle A \rangle \subseteq \langle B \rangle$, as required.

10. Since B_1 is independent and B_2 is a spanning set for V, it follows that $|B_1| \leq |B_2|$. Interchanging B_1 and B_2 is the previous reasoning gives the opposite inequality and so $|B_1| = |B_2|$. Therefore all bases of V have the same number of elements, the *dimension* of V.

Problem Set 9

1. Certainly $\mathbf{0} = \mathbf{0} + \mathbf{0} \in U + W$ so $U + W \neq \emptyset$. Let $\mathbf{x}_1, \mathbf{x}_2 \in U + W$ so that $\mathbf{x}_1 = \mathbf{u}_1 + \mathbf{v}_1$ say and $\mathbf{x}_2 = \mathbf{u}_2 + \mathbf{v_2}$. Take scalars a,b. Then

$$a\mathbf{x}_1 + b\mathbf{x}_2 = a(\mathbf{u}_1 + \mathbf{v}_1) + b(\mathbf{u}_2 + \mathbf{v}_2) = (a\mathbf{u}_1 + b\mathbf{u}_2) + (a\mathbf{v}_1 + b\mathbf{v}_2) \in U + W,$$

and so $U + W$ is a subspace of V.

2. The dimension of the domain space is given as n and, by Question 7 of Set 8 it follows that the dimension of the range space $L(U)$ is the rank $m \leq n$

of M. Now the kernel of L is the solution space of the system $M\mathbf{x} = \mathbf{0}$, which is that of $E\mathbf{x} = \mathbf{0}$, where E is the echelon form of M. Each unknown x_i for which the ith column of E is *not* a pivot column may be assigned freely, with the other unknowns expressed in terms of these free variables. This leads to a basis of order $n - m$ for the kernel of L. Therefore

$$\dim(\text{kernel}(L)) + \dim(\text{range}(L)) = \dim(\text{domain}(L)).$$

3. Since $A = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k\}$ is an independent subset of V and $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m\}$ spans V we know from the Exchange Lemma that $k \leq m$. If A does not span V then some member of S, without loss we may assume it is $\mathbf{v}_1$, does lie in $\langle A \rangle$. Then $A \cup \{\mathbf{v}_1\} = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{v}_1\}$ is independent as no member of the set is a linear combination of its predecessors. We may repeat this argument as often as required until we have extended A to a basis B of V. Therefore any independent subset of a finite dimensional vector space may be extended to a basis for V.

4. Let $\mathbf{u}$ and $\mathbf{v}$ respectively be solutions to the systems $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x} = \mathbf{b}$. Then

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0}+\mathbf{b} =\mathbf{b}$$

so it follows that $U + \mathbf{v}$ consists of solutions of the inhomogeneous system $A\mathbf{x} = \mathbf{b}$, where U is the subspace of solutions of the homgeneous system.

Conversely let $\mathbf{w}$ be any solution of $A\mathbf{x} = \mathbf{b}$. Then

$$A(\mathbf{w} - \mathbf{v}) = A\mathbf{w} - A\mathbf{v} = \mathbf{b}-\mathbf{b} = \mathbf{0}$$

so that $\mathbf{v} - \mathbf{w} \in U$ and since $\mathbf{w} = (\mathbf{w} - \mathbf{v}) + \mathbf{v} \in U + \mathbf{v}$. Therefore the solution set of $A\mathbf{x} = \mathbf{b}$ is $U + \mathbf{v}$ where U is the subspace of solutions of $A\mathbf{x} = \mathbf{0}$ and $\mathbf{v}$ is *any* solution of the system $A\mathbf{x} = \mathbf{b}$.

5. Let U be a subspace of V. By Question 3, any basis B of U may be extended to a basis B' of V. Hence $\dim(U) \leq \dim(V)$. Moreover if we have equality then since B may be extended to a basis for V and all bases for V have the same number of elements, it follows that B is a basis for V and so $U = V$. Therefore a subspace U of a finite dimensional vector space V has dimension no larger than that of V with equality of dimension if and only if $U = V$.

6. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k\}$ be a spanning set for V and let B be a maximal independent subset of S. If B were not a basis for V there would exist some $\mathbf{u}_i \in S$ such that $\mathbf{u}_i \notin \langle B \rangle$ as otherwise $S \subseteq \langle B \rangle$ and then $V = \langle S \rangle \subseteq \langle B \rangle$ in which case B would be a basis for V. But then $B \cup \{\mathbf{u}_i\}$ is an independent set that strictly contains B as, attaching $\mathbf{u}_i$ to the end of any list of the elements of B gives a set in which no members is a linear combination of its predecessors. However this now contradicts that B is a maximal independent subset of S. Therefore any maximal independent subset of a spanning set for V is a basis for V.

7. By Question 5, $U \cap W$ is a finite dimensional with basis B_1say with k elements. By Question 4, $B_1 = \{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ can be extended to a finite basis $B_2 = B_1 \cup C$ where $C = \{\mathbf{c}_1, \cdots, \mathbf{c}_m\}$ say of U and extended to a finite basis

$B_3 = B_1 \cup D$ where $D = \{\mathbf{d}_1, \cdots, \mathbf{d}_n\}$ say of W with both these unions being *disjoint*, meaning that $B_2 \cap C = \emptyset = B_3 \cap D$. Note also that $C \cap D = \emptyset$ as no member of $C \cup D$ lies in $U \cap W$. We claim that $B = B_1 \cup C \cup D$ is a basis for $U + W$. First we show that B is independent. To this end consider the set B listed in the order

$$B = \{\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{c}_1, \cdots, \mathbf{c}_m, \mathbf{d}_1, \cdots, \mathbf{d}_n\}.$$

Since B_2 spans U and B_3 spans W it follows that $B = B_2 \cup B_3$ spans the vector space $U + W$. We next show that B is independent. Since B_2 is an independent set, no $\mathbf{u}_i$ or $\mathbf{c}_j$ is a linear combination of its predecessors in the above list for B. Suppose however that for some $\mathbf{d}_i$ we have

$$\mathbf{d}_i = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k + a_{k+1}\mathbf{c}_1 + \cdots + a_{k+m}\mathbf{c}_m + a_{k+m+1}\mathbf{d}_1 + \cdots + a_{k+m+i-1}\mathbf{d}_{i-1} \quad (1)$$

$$\Rightarrow \mathbf{d}_i - a_{k+m+1}\mathbf{d}_1 - \cdots - a_{k+m-i}\mathbf{d}_{i-1} = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k + a_{k+1}\mathbf{c}_1 + \cdots + a_{k+m}\mathbf{c}_m \quad (2)$$

However the RHS of (2) lies in U while the LHS lies in W so that both sides represent a common member $\mathbf{x} \in U \cap W$. Hence the RHS of (2) may therefore be written as a linear combination of the vectors of B_1 it follows that $\mathbf{d}_i$ is a linear combination of $\mathbf{u}_1, \cdots, \mathbf{u}_k, \mathbf{d}_1, \cdots, \mathbf{d}_{i-1}$, contradicting that B_3 is a basis for W. It follows that B is indeed an independent set and therefore a basis for $U + W$. Moreover it now follows that the order of B is $k + m + n = (k + m) + (k + n) - k$, which is equivalent to the required statement:

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

8. Let $\mathbf{x} = \mathbf{u} + \mathbf{v}$ where $\mathbf{u} \in U$ and $\mathbf{v} \in W$ and let $\mathbf{y} \in U \cap W$. Then $\mathbf{y}, -\mathbf{y} \in U \cap W$ so that $\mathbf{u} + \mathbf{y} \in U$ and $\mathbf{u} - \mathbf{y} \in W$ and so $\mathbf{x} = (\mathbf{u} + \mathbf{y}) + (\mathbf{v} - \mathbf{y})$. It follows that if the representation of *any* member $\mathbf{x} \in U + W$ is unique then $U \cap W = \{\mathbf{0}\}$.

Conversely suppose that $U \cap W = \{\mathbf{0}\}$ and let $\mathbf{x} = \mathbf{u}_1 + \mathbf{v_1} = \mathbf{u_2} + \mathbf{v_2}$ where $\mathbf{u}_1, \mathbf{u}_2 \in U$ and $\mathbf{v}_1, \mathbf{v}_2 \in W$. Then $\mathbf{u}_1 - \mathbf{u}_2 = \mathbf{v}_1 - \mathbf{v}_2 \in U \cap W = \{\mathbf{0}\}$ so that $\mathbf{u_1} = \mathbf{u_2}$ and $\mathbf{v}_1 = \mathbf{v}_2$.

9. Certainly $\mathbf{0} \in U^\perp$ so that $U^\perp \neq \emptyset$. Take any $\mathbf{v}, \mathbf{w} \in U^\perp$ and scalars a, b. Then for any $\mathbf{u} \in U$ we have

$$\mathbf{u} \bullet (a\mathbf{v} + b\mathbf{w}) = \mathbf{u} \bullet (a\mathbf{v}) + \mathbf{u} \bullet (b\mathbf{w})$$

$$= a(\mathbf{u} \bullet \mathbf{v}) + b(\mathbf{u} \bullet \mathbf{w}) = a0 + b0 = 0.$$

Therefore $U^\perp$ is a subspace of V. Moreover if $\mathbf{v} \in U \cap U^\perp$ then $\mathbf{v} \bullet \mathbf{v} = 0$, whence $\mathbf{v} = \mathbf{0}$ and so $U \cap U^\perp = \{\mathbf{0}\}$.

10. Let $A = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k\}$ be a basis of U. Then $U^\perp$ is exactly the solution set of $M\mathbf{x} = \mathbf{0}$ where the rows of M are exactly the members of A. The solution space is the kernel of the linear mapping defined by left multiplication by M, the rank of which, since A is independent, is k. It follows by Question 6 of Set 8 that the dimension of $U^\perp$ is given by $n - k$. It now follows from

Question 6 that $\dim(U \oplus U^{\perp}) = k + (n-k) - 0 = n$. Therefore, by Question 4 we conclude that

$$U \oplus U^{\perp} = V.$$

Problem Set 10

1. Let us write

$$\mathbf{v} = a_1\mathbf{v_1} + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k$$

$$\Rightarrow \mathbf{v}\bullet\mathbf{v}_1 = a_1\mathbf{v}\bullet\mathbf{v}_1 + a_2\mathbf{v}\bullet\mathbf{v}_2 + \cdots + a_k\mathbf{v}\bullet\mathbf{v}_k$$

$$= a_1||\mathbf{v}_1||^2 + a_2(0) + \cdots + a_k(0) = a_1$$

And so we see that $a_1 = \mathbf{v}\bullet\mathbf{v}_1$ and by the same argument we obtain generally that $a_i = \mathbf{v}\bullet\mathbf{v}_i$ $(1 \leq i \leq k)$, as required.

2. Suppose that

$$\mathbf{v} = a_1\mathbf{v_1} + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k = \mathbf{0}.$$

Then $a_i = \mathbf{v}\bullet\mathbf{v}_i = \mathbf{v}\bullet\mathbf{0} = 0$, for all $1 \leq i \leq k$, and therefore $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ is independent.

3.

$$||\mathbf{v}_1|| = \frac{1}{2^2}(3+1) = 1,\ ||\mathbf{v}_2|| = \frac{1}{2^2}((-1)^2 + 3) = 1,$$

so both vectors are unit vectors. Moreover $\mathbf{v}_1\bullet\mathbf{v}_2 = \frac{1}{2^2}(-\sqrt{3}+\sqrt{3}) = 0$ so that the pair $B = \{\mathbf{v}_1, \mathbf{v}_2\}$form an orthonormal set of order 2. By Question 2, B is independent and is therefore B is a basis of the 2-dimensional vector space $\mathbb{R}^2$.

Given $\mathbf{v}$ = (-2,3) we have the B-coordinates of $\mathbf{v}$ are given by

$$(\mathbf{v}\bullet\mathbf{v}_1, \mathbf{v}\bullet\mathbf{v}_2) = (-\sqrt{3} + \frac{3}{2}, 1 + \frac{3\sqrt{3}}{2}).$$

4. We first verify that $W = \{\mathbf{w}_1, \cdots, \mathbf{w}_k\}$ is an orthogonal set of vectors. The equations are:

$$\mathbf{w}_1 = \mathbf{v}_1,\ \mathbf{w}_i = \mathbf{v}_i - \frac{\mathbf{v}_i\bullet\mathbf{w}_1}{\mathbf{w}_1\bullet\mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{v}_i\bullet\mathbf{w}_2}{\mathbf{w}_2\bullet\mathbf{w}_2}\mathbf{w}_2 - \cdots - \frac{\mathbf{v}_i\bullet\mathbf{w}_{i-1}}{\mathbf{w}_{i-1}\bullet\mathbf{w}_{i-1}}\mathbf{w}_{i-1},\ 2 \leq i \leq k. \tag{3}$$

Suppose inductively that for all $j < i$ we have $\mathbf{w}_i\bullet\mathbf{w}_j = 0$, which holds by default when $i = 1$. Now take $i \geq 2$ and suppose that the claim holds for all lesser values of i. Then from (3) and the inductive hypothesis we have

$$\mathbf{w}_i\bullet\mathbf{w}_j = \mathbf{v}_i\bullet\mathbf{w}_j - \frac{\mathbf{v}_i\bullet\mathbf{w}_j}{\mathbf{w}_j\bullet w_j}\mathbf{w}_j\bullet\mathbf{w}_j = 0,$$

and so the induction continues and therefore the set W consists of mutually orthogonal vectors. To complete the proof we need to prove that no member of

W is the zero vector, which is true for $\mathbf{w}_1 = \mathbf{v}_1$ as the orginal set of vectors is independent. Again we may now check this inductively. By construction, the set $\{\mathbf{w}_1, \cdots, \mathbf{w}_{i-1}\}$ lies in the span of $\{\mathbf{v}_1, \cdots, \mathbf{v}_{i-1}\}$. If now $\mathbf{w}_i = \mathbf{0}$ it would follow from (3) that $\mathbf{v}_i$ was in the span of $\{\mathbf{w}_1, \cdots, \mathbf{w}_{i-1}\}$ and hence in the span of $\{\mathbf{v}_1, \cdots, \mathbf{v}_{i-1}\}$ contrary to the independence of the original basis set $\{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$. Therefore the set W is an orthogonal set of non-zero vectors and the correpsonding set of unit vectors forms an orthogonal basis of the vector space spanned by $\{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$.

5. The given homogeneous system of equations gives rise to the row reduction as follows:

$$\begin{array}{rcccccccl} x & + & y & + & z & + & w & = & 0 \\ -x & + & y & & & + & w & = & 0 \end{array}$$

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

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

putting $z = 2c$ and $w = d$ for arbitrary constants c and d gives $x = -c, y = -c-d$. Hence the solution vector $\mathbf{x}$ is given by $\mathbf{x} = c(-1, -1, 2, 0)+d(0, -1, 0, 1)$. Hence as a basis for the solution space we may choose $\{\mathbf{v}_1, \mathbf{v}_2\} = \{(1, 1, -2, 0), (0, 1, 0, -1)\}$.

Applying the Gram-Schmidt equations we have $\mathbf{w}_1 = \mathbf{v}_1 = (1, 1, -2, 0)$. Using the Gram-Schmidt equation we then have:

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \bullet \mathbf{w}_1}{\mathbf{w}_1 \bullet \mathbf{w}_1}\mathbf{w}_1 = (0, 1, 0, -1) - \frac{0+1+0+0}{1+1+4+0}(1, 1, -2, 0)$$

$$= (0, 1, 0, -1) - \frac{1}{6}(1, 1, -2, 0) = \frac{1}{6}(-1, 5, 2, -6).$$

Hence $||\mathbf{w}_1|| = \sqrt{1+1+4+0} = \sqrt{6}$ and $||\mathbf{w}_2|| = \frac{1}{6}\sqrt{1+25+4+36} = \frac{1}{6}\sqrt{66}. =$ The corresponding orthonmormal basis for the solution space of the system therefore is

$$\{\frac{\mathbf{w}_1}{||\mathbf{w}_1||}, \frac{\mathbf{w}_2}{||\mathbf{w}_2||}\} = \frac{1}{\sqrt{6}}(1, 1, -2, 0), \frac{1}{\sqrt{66}}(-1, 5, 2, -6).$$

6.

$$\begin{bmatrix} -4 & 1 & 1 & 1 & 0 & 0 \\ 5 & -3 & -2 & 0 & 1 & 0 \\ -1 & 2 & 1 & -1 & -1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \\ -1 & 5 & 3 & 4 & 3 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & -1 & 1 & 1 & 0 \\ 5 & -3 & -2 & 0 & 1 & 0 \\ -1 & 2 & 1 & -1 & -1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \\ -1 & 5 & 3 & 4 & 3 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & -1 & 1 & 1 & 0 \\ 0 & 7 & 3 & -5 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \\ 0 & 3 & 2 & 5 & 4 & 5 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & -2 & -1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \\ 0 & 7 & 3 & -5 & -4 & 0 \\ 0 & 3 & 2 & 5 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{2} & -5 & -4 & -\frac{7}{2} \\ 0 & 0 & -\frac{1}{2} & 5 & 4 & \frac{7}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 10 & 8 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 2 & 0 & -10 & -8 & -6 \\ 0 & 0 & 1 & 10 & 8 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & -5 & -4 & -3 \\ 0 & 0 & 1 & 10 & 8 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence the change-of-basis matrix from **B**- to **C**-coordinates is

$$P = \begin{bmatrix} 1 & 1 & 1 \\ -5 & -4 & -3 \\ 10 & 8 & 7 \end{bmatrix}.$$

Comment This calculation that both sets are independent and generate the same space as the row reduced matrix has exactly two rows of zeros.

Continuing our question, the vector $\mathbf{v} = 2(1, 0, -1, 0, 4)^T - (0, 1, -1, 0, 3)$ has **B**-coords $2(1, 0, 0) - (0, 1, 0) = (2, -1, 0)$ and so the **C**-coordinates of **v**is $P\mathbf{v}$

$$= \begin{bmatrix} 1 & 1 & 1 \\ -5 & -4 & -3 \\ 10 & 8 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -6 \\ 12 \end{bmatrix}.$$

7. First we find the eigenvalues of the transformation:

$$\begin{vmatrix} 1-\lambda & 5 \\ 2 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) - 10 = 0$$

$$\Rightarrow \lambda^2 - 5\lambda - 6 = (\lambda + 1)(\lambda - 6) = 0,$$

so that $\lambda \in \{-1, 6\}$. The equation for an eigenvector with $\lambda_1 = -1$ is $2x+5y = 0$ so that $\mathbf{v}_1$=$(5, -2)^T$. For $\lambda_2 = 6$ the equation is $-5x + 5y = 0$ so that $\mathbf{v}_2 = (1, 1)^T$ is an eigenvector. Hence our matrix P of eigenvectors and its inverse P^{-1} are given by

$$P = \begin{bmatrix} 5 & 1 \\ -2 & 1 \end{bmatrix} \quad P^{-1} = \frac{1}{7} \begin{bmatrix} 1 & -1 \\ 2 & 5 \end{bmatrix} \quad D = \begin{bmatrix} -1 & 0 \\ 0 & 6 \end{bmatrix}$$

and so $A = PDP^{-1}$ as can be checked directly. Direct calculation now gives that for any vector $\mathbf{v} = (a, b)^T$:

$$A^n\mathbf{v} = PD^nP^{-1}\mathbf{v} = \frac{1}{7} \begin{bmatrix} (-1)^n 5(a-b) + 6^n(2a+5b) \\ 2(-1)^{n+1}(a-b) + 6^n(2a+5b) \end{bmatrix}.$$

Hence, as long as $2a+5b \neq 0$, which is to say **v** is not a multiple of the eigenvector of the smaller eigenvalue, for large n, the direction of $A^n\mathbf{v}$ approaches that of $\mathbf{v}_2 = (1, 1)^T$.

8. Let $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ denote an orthonormal basis for $\mathbb{R}^n$. Write $\mathbf{v} = a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n$ and $\mathbf{w} = b_1\mathbf{u}_1 + \cdots + b_n\mathbf{u}_n$ so that, by orthogonality and the fact that $\mathbf{u}_i \bullet \mathbf{u}_i = 1$ we have

$$\mathbf{v} \bullet \mathbf{w} = a_1b_1 + \cdots + a_nb_n = (\mathbf{v} \bullet \mathbf{u}_1)(\mathbf{w} \bullet \mathbf{u}_1) + \cdots + (\mathbf{v} \bullet \mathbf{u}_n)(\mathbf{w} \bullet \mathbf{u}_n).$$

Putting $\mathbf{w} = \mathbf{v}$ is this result gives

$$\mathbf{v} \bullet \mathbf{v} = ||\mathbf{v}||^2 = (\mathbf{v} \bullet \mathbf{u}_1)^2 + \cdots + (\mathbf{v} \bullet \mathbf{u}_n)^2.$$

9. We extend the orthonormal set $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ to an orthonormal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ for $\mathbb{R}^n$ by the Gram-Schmidt algorithm as this will not alter the first listed set of k orthonormal vectors. All the terms on the right in Parseval's equality are non-negative and so, deleting all but the first k terms gives Bessel's inequality:

$$(\mathbf{v} \bullet \mathbf{u}_1)^2 + \cdots + (\mathbf{v} \bullet \mathbf{u}_k)^2 \leq ||\mathbf{v}||^2.$$

10. The least squares approximation to a set $(x_1, y_1), \cdots, (x_n, y_n)$ is the line $y = mx + b$ where for $\mathbf{1} = (1, 1, \cdots 1)$, $\mathbf{x} = (x_1, \cdots, x_n)$, $\mathbf{y} = (y_1, \cdots, y_n)$

$$m = \frac{(\mathbf{1} \bullet \mathbf{1})(\mathbf{x} \bullet \mathbf{y}) - (\mathbf{1} \bullet \mathbf{x})(\mathbf{1} \bullet \mathbf{y})}{(\mathbf{1} \bullet \mathbf{1})(\mathbf{x} \bullet \mathbf{x}) - (\mathbf{1} \bullet \mathbf{x})^2} \quad b = \frac{(\mathbf{x} \bullet \mathbf{x})(\mathbf{1} \bullet \mathbf{y}) - (\mathbf{x} \bullet \mathbf{y})(\mathbf{1} \bullet \mathbf{x})}{(\mathbf{1} \bullet \mathbf{1})(\mathbf{x} \bullet \mathbf{x}) - (\mathbf{1} \bullet \mathbf{x})^2}.$$

In this example $\mathbf{x} = (-1, 1, 3, 5)$ and $\mathbf{y} = (1, -1, -4, -4)$ and so

$$m = \frac{(4)(-34) - (8)(-8)}{(4)(36) - 8^2} = -\frac{9}{10}$$

$$b = \frac{(36)(-8) - (-34)(8)}{(4)(36) - 8^2} = -\frac{1}{5}.$$

The line of best fit is therefore $y = -\frac{9}{10}x - \frac{1}{5}$.