Instructions:

- This assignment is meant to help you grok certain concepts we will use in the course. Please don't copy solutions from any sources.
- Avoid verbosity.
- Questions marked with * are relatively difficult. Don't be discouraged if you cannot solve them right away!
- The assignment needs to be written in latex using the attached tex file. The solution for each question should be written in the solution block in space already provided in the tex file. **Handwritten assignments will not be accepted.**

---

1. Suppose, a transformation matrix A, transforms the standard basis vectors of $R^3$ as follows :

$$\begin{bmatrix}1\\0\\0\end{bmatrix} => \begin{bmatrix}3\\8\\0\end{bmatrix} ; \begin{bmatrix}0\\1\\0\end{bmatrix} => \begin{bmatrix}-4\\9\\7\end{bmatrix} ; \begin{bmatrix}0\\0\\1\end{bmatrix} => \begin{bmatrix}-1\\2\\6\end{bmatrix}$$

(a) If the volume of a hypothetical parallelepiped in the un-transformed space is $100units^3$ what will be volume of this parallelepiped in the transformed space?

**Solution:** Consider the matrix $X = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$ and the transformed matrix $B = \begin{bmatrix}3 & -4 & -1\\8 & 9 & 2\\0 & 7 & 6\end{bmatrix}$

We know from the transformation equation $B = AX$, where A is the transformation matrix that(if the inverse of $X$ exists)

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

$$\implies A = \begin{bmatrix}3 & -4 & -1\\8 & 9 & 2\\0 & 7 & 6\end{bmatrix}$$

The volume($V_f$) of the parallelepiped in the transformed space is given by the formula

$$V_f = |A|V_o$$

where $V_o$ is the volume of the parellelepiped in the untransformed space.

$$|A| = 3*(54-14) - (-4)*(48-0) + (-1)*(56-0)$$

$$\implies |A| = 256$$

Hence

$$V_f = 256 * 100 units^3$$

$$\implies V_f = 25600 units^3$$

(b) What will be the volume if the transformation of the basis vectors is as follows :

$$\begin{bmatrix}1\\0\\0\end{bmatrix} => \begin{bmatrix}1\\1\\3\end{bmatrix} ; \begin{bmatrix}0\\1\\0\end{bmatrix} => \begin{bmatrix}-1\\2\\0\end{bmatrix} ; \begin{bmatrix}0\\0\\1\end{bmatrix} => \begin{bmatrix}0\\2\\2\end{bmatrix}$$

**Solution:** Consider the matrix $X = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$ and the transformed matrix $B = \begin{bmatrix}1 & -1 & 0\\1 & 2 & 2\\3 & 0 & 2\end{bmatrix}$

We know from the transformation equation $B = AX$, where A is the transformation matrix that(if the inverse of $X$ exists)

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

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

The volume($V_f$) of the parallelepiped in the transformed space is given by the formula

$$V_f = |A|V_o$$

where $V_o$ is the volume of the parellelepiped in the untransformed space.

$$|A| = 1 * (4 - 0) - (-1) * (2 - 6) + (0) * (0 - (-6))$$

$$\implies |A| = 0$$

Hence

$$V_f = 0 * 100 units^3$$

$$\implies V_f = 0 units^3$$

(c) Comment on the uniqueness of the second transformation.

**Solution:** The transformation matrix is such that it transforms all $3-d$ vectors into a $2-d$ plane, hence giving a volume of 0 for each transformed volume.

2. If $R^3$ is represented by following basis vectors : $\begin{bmatrix}5\\2\\0\end{bmatrix}$ , $\begin{bmatrix}8\\7\\-11\end{bmatrix}$ , $\begin{bmatrix}-4\\-9\\3\end{bmatrix}$

   (a) Find the representation of the vector $\begin{pmatrix}-3 & 1 & -2\end{pmatrix}^T$ (as represented in standard basis) in the above basis.

**Solution:** This problem is same as finding the vector $x$ in the equation $b = Ax$ where $b = \begin{bmatrix}-3\\1\\-2\end{bmatrix}$ and $A = \begin{bmatrix}5 & 8 & -4\\2 & 7 & -9\\0 & -11 & 3\end{bmatrix}$ If $A$ is invertible, then we can solve the above equation by taking $A^{-1}$.

$$|A| = 5*(21-99) - 8*(6-0) + (-4)*(-22-0)$$

$$\implies |A| = -350$$

Hence $A$ is invertible

$$\implies x = A^{-1}b$$

On solving for $A^{-1} = \frac{adj(A)}{|A|}$, we obtain $A^{-1}$ as :

$$A^{-1} = \frac{-1}{350}\begin{bmatrix}-78 & 20 & -44\\-6 & 15 & 37\\-22 & 55 & 19\end{bmatrix}$$

$$\implies x = \frac{-1}{350}\begin{bmatrix}-78 & 20 & -44\\-6 & 15 & 37\\-22 & 55 & 19\end{bmatrix}\begin{bmatrix}-3\\1\\-2\end{bmatrix}$$

$$\implies x = \begin{bmatrix}\frac{-171}{175}\\\frac{41}{350}\\\frac{-83}{350}\end{bmatrix}$$

Thus we can write the vector as :

$$\begin{bmatrix}-3\\1\\-2\end{bmatrix} = \frac{-171}{175}\begin{bmatrix}5\\2\\0\end{bmatrix} + \frac{41}{350}\begin{bmatrix}8\\7\\-11\end{bmatrix} + (\frac{-83}{350})\begin{bmatrix}-4\\-9\\3\end{bmatrix}$$

(b) We know that, orthonormal basis simplifies this transformation to a great extent. What would be the representation of vector $\begin{pmatrix}-3 & 1 & -2\end{pmatrix}^T$ in the orthogonal basis represented by : $\begin{bmatrix}1\\-1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$.

**Solution:** In the case of orthonormal vectors, we can use the property of inner product to simplify our calculations. Let the set of orthonormal basis vectors be $(v_1, v_2 \dots v_n)$. Any vector in the space spanned by these vectors can be represented as

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \dots \alpha_n v_n$$

Thus the constants $\alpha_i$ can be found by taking the inner product of $x$ with each orthonormal vector.

$$\alpha_i = \frac{\langle x, v_i \rangle}{\|v_i\|^2}$$

In this case

$$v_1 = \begin{bmatrix}1\\-1\\0\end{bmatrix}, v_2 = \begin{bmatrix}1\\1\\0\end{bmatrix}, v_3 = \begin{bmatrix}0\\0\\1\end{bmatrix} and x = \begin{bmatrix}-3\\1\\-2\end{bmatrix}$$

$$\alpha_1 = \frac{(-3)*(1) + (1)*(-1) + (0)*(-2)}{1^2 + (-1)^2} = -2$$

$$\alpha_2 = \frac{(-3)*(1) + (1)*(1) + (0)*(-2)}{1^2 + (1)^2} = -1$$

$$\alpha_3 = \frac{(-3)*(0) + (1)*(0) + (1)*(-2)}{1^2} = -2$$

Hence

$$\begin{bmatrix}-3\\1\\-2\end{bmatrix} = (-2)\begin{bmatrix}1\\-1\\0\end{bmatrix} + (-1)\begin{bmatrix}1\\1\\0\end{bmatrix} + (-2)\begin{bmatrix}0\\0\\1\end{bmatrix}$$

(c) Comment on the advantages of having orthonormal basis.

**Solution:** As mentioned earlier, the inner product definition simplifies our calculation by a great amount. Let the set of orthonormal basis vectors be $(v_1, v_2 \dots v_n)$. Any vector in the space spanned by these vectors can be represented as

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \dots \alpha_n v_n$$

Thus the constants $\alpha_i$ can be found by taking the inner product of $x$ with each orthonormal vector.

$$\alpha_i = \frac{\langle x, v_i \rangle}{\|v_i\|^2}$$

where the inner product is defined as

$$\langle x, y \rangle = x^T y$$

3. A square matrix is a Markov matrix if each entry is between zero and one and the sum along each row is one. Prove that a product of Markov matrices is Markov.

**Solution:** Let the square matrix be of dimension $(n, n)$. Consider the markov matrices $A$ and $B$. From the markov property,

$$\sum_{j=1}^{n} a_{ij} = 1 \ldots 1 \leq i \leq n, 0 \leq a_{ij} \leq 1 \tag{1}$$

$$\sum_{j=1}^{n} b_{ij} = 1 \ldots 1 \leq i \leq n, 0 \leq b_{ij} \leq 1 \tag{2}$$

Let $C = AB$, then

$$C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Let us test if $C$ satisfies the Markov property.

$$\sum_{j=1}^{n} C_{ij} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ik} b_{kj} \ldots 1 \leq i \leq n$$

Since $a_{ik}$ does not depend on $j$, we can change the summation as follows

$$\sum_{j=1}^{n} C_{ij} = \sum_{k=1}^{n} \sum_{j=1}^{n} a_{ik} b_{kj}$$

$$\implies \sum_{j=1}^{n} C_{ij} = \sum_{k=1}^{n} a_{ik} \sum_{j=1}^{n} b_{kj}$$

Using equation [1] and [2],

$$\sum_{j=1}^{n} C_{ij} = 1$$

Hence the product of two Markov matrices is Markov.

4. Give an example of a matrix A with the following three properties:

   (a) A has eigenvalues -1 and 2.

   (b) The eigenvalue -1 has eigenvector

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \tag{3}$$

   (c) The eigenvalue 2 has eigenvector

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \textit{and} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \tag{4}$$

**Solution:** Consider the matrix of eigenvectors $S = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 3 & 0 & 1 \end{bmatrix}$. We know that the number of linearly independent eigenvectors corresponding to a eigenvalue is its multiplicity. Hence the multiplicity of $\lambda = -1$ is 1 and the multiplicity of $\lambda = 2$ is 2. Consider the matrix with eigenvalues along the diagonal and rest of the elements as 0, $\Lambda = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Then an example of matrix $A$ would be $S^{-1}\Lambda S$.

$$S^{-1} = \frac{Adj(S)}{|S|}$$

$$\implies S^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -3 & 3 & -1 \end{bmatrix}$$

Hence

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

$$\implies A = \begin{bmatrix} 0.5 & 1.5 & -1.5 \\ -3 & 5 & -3 \\ -4.5 & 4.5 & -2.5 \end{bmatrix}$$

5. Perform the Gram-Schmidt process on each of these basis for $\mathbb{R}^3$. And convert the resulting orthogonal basis into orthonormal basis.

(a) $\langle \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} \rangle$

**Solution:** Let the orthonormal basis be $(u_1, u_2, u_3)$. Let the current basis vectors be $v_1 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$.

$$u_1 = \frac{v_1}{\|v_1\|}$$

$$\implies u_1 = \frac{1}{2\sqrt{3}} \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

Consider the intermediate vector $y_2$, where $u_2 = \frac{y_2}{\|y_2\|}$

$$y_2 = v_2 - \langle v_2, u_1 \rangle u_1$$

$$\implies y_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - 0 * \frac{1}{2\sqrt{3}} \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

$$y_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$\implies u_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Consider the intermediate vector $y_3$, where $u_3 = \frac{y_3}{\|y_3\|}$

$$y_3 = v_3 - \langle v_3, u_2 \rangle u_2 - \langle v_3, u_1 \rangle u_1$$

$$\implies y_3 = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} - (-1)\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \frac{8}{2\sqrt{3}}\frac{1}{2\sqrt{3}} \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

$$\implies y_3 = \begin{bmatrix} \frac{-5}{6} \\ \frac{5}{3} \\ \frac{-5}{6} \end{bmatrix}$$

$$\implies u_3 = \frac{\sqrt{6}}{5}\begin{bmatrix} \frac{-5}{6} \\ \frac{5}{3} \\ \frac{-5}{6} \end{bmatrix}$$

Hence the orthonormal basis obtained using Gram-Schmidt orthogonalization technique is $\langle \frac{1}{2\sqrt{3}}\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \frac{\sqrt{6}}{5}\begin{pmatrix} \frac{-5}{6} \\ \frac{5}{3} \\ \frac{-5}{6} \end{pmatrix}\rangle$

(b) $\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}\rangle$

**Solution:** Let the orthonormal basis be $(u_1, u_2, u_3)$. Let the current basis vectors be $v_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$.

$$u_1 = \frac{v_1}{\|v_1\|}$$

$$\implies u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Consider the intermediate vector $y_2$, where $u_2 = \frac{y_2}{\|y_2\|}$

$$y_2 = v_2 - \langle v_2, u_1\rangle u_1$$

$$\implies y_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - (\frac{-1}{\sqrt{2}}) * \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$y_2 = \begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix}$$

$$\implies u_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

Consider the intermediate vector $y_3$, where $u_3 = \frac{y_3}{\|y_3\|}$

$$y_3 = v_3 - \langle v_3, u_2\rangle u_2 - \langle v_3, u_1\rangle u_1$$

$$\implies y_3 = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} + \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 2.5 \\ 2.5 \end{bmatrix}$$

$$\implies y_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\implies u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Hence the orthonormal basis obtained using Gram-Schmidt orthogonalization technique is $\langle \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rangle$

6. Suppose, every year, 4% of the birds from Canada migrate to the US, and 1% of them travel to Mexico. Similarly, every year, 6% of the birds from US migrate to Canada, and 4% to Mexico. Finally, every year 10% of the birds from Mexico migrate to the US, and 0% go to Canada.

   (a) Represent the above probabilities in a transition matrix.

**Solution:** Let the current bird population in Canada, US and Mexico be $C_t, U_t$ and $M_t$ respectively. Let the bird population in the next year be denoted by $C_{t+1}, U_{t+1}$ and $M_{t+1}$ respectively.

$$C_{t+1} = C_t + 0.06U_t - 0.05C_t$$

$$\implies C_{t+1} = 0.95C_t + 0.06U_t$$

$$U_{t+1} = U_t + 0.1M_t - 0.1U_t + 0.04C_t$$

$$\implies U_{t+1} = 0.9U_t + 0.1M_t + 0.04C_t$$

$$M_{t+1} = M_t - 0.1M_t + 0.01C_t + 0.04U_t$$

$$\implies M_{t+1} = 0.9M_t + 0.01C_t + 0.04U_t$$

Writing the above equations in matrix form

$$\begin{bmatrix} C_{t+1} \\ U_{t+1} \\ M_{t+1} \end{bmatrix} = \begin{bmatrix} 0.95 & 0.06 & 0 \\ 0.04 & 0.9 & 0.1 \\ 0.01 & 0.04 & 0.9 \end{bmatrix} \begin{bmatrix} C_t \\ U_t \\ M_t \end{bmatrix}$$

Thus the transition matrix is

$$A = \begin{bmatrix} 0.95 & 0.06 & 0 \\ 0.04 & 0.9 & 0.1 \\ 0.01 & 0.04 & 0.9 \end{bmatrix}$$

(b) Is it possible that after some years, the number of birds in the 3 countries will become constant?

**Solution:** For the population to become constant after some years(say $t_o$)

$$Av_{t_o} = v_{t_o}$$

But $v_{t_o} = A^{t_o}v_o$ where $v_o$ is the initial distribution. Thus substituting back we get

$$A(A^{t_o}v_o) = A^{t_o}v_o$$

Since this is a Markov matrix, 1 is an eigenvalue of $A$ which means that for an appropriate $v_o$, it is possible that the population of birds becomes constant in the three countries.
The initial concentration can be written as a linear combination of the eigenvectors. Since this is a Markov matrix, all other eigen values have an absolute value less than or equal to 1. Let

$$v_o = a_1v_1 + a_2v_2 + a_3v_3$$

where $v_1, v_2$ and $v_3$ are the eigen vectors. Hence after $k$ years the population will become

$$v_t = a_1\lambda_1^k v_1 + a_2\lambda_2^k v_2 + a_3\lambda_3^k v_3$$

We are seeing that the eigen values less than 1 will vanish as $k$ becomes very large and the only term that will survive is the one with eigen value 1. Hence as $k$ tends to infinity, the population will become a constant.

7. (a) Show that any set of four unique vectors in $\mathbb{R}^2$ is linearly dependent.

**Solution:** Consider the vectors $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ in $\mathbb{R}^2$.
Consider the equation

$$\alpha_1 a + \alpha_2 b + \alpha_3 c + \alpha_4 d = 0$$

These 4 vectors are linearly independent if the above equation is satisfied only when $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$. If not, they are linearly dependent. The above equation can be split as

$$\alpha_1 a_1 + \alpha_2 b_2 + \alpha_3 c_3 + \alpha_4 d_4 = 0$$

$$\alpha_1 a_2 + \alpha_2 b_2 + \alpha_3 c_2 + \alpha_4 d_2 = 0$$

On solving this system on linear equations, we can get our variable values. However there are 2 equations and 4 variables, which leads to infinite solutions most of which will be non-zero. Hence any set of 4 unique vectors in $\mathbb{R}^2$ are linearly dependent.

(b) What is the maximum number of unique vectors that a linearly independent subset of $\mathbb{R}^2$ can have ?

**Solution:** Consider $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in $\mathbb{R}^2$. They are linearly independent as

$$c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

only when $c_1 = c_2 = 0$. The span of these two vectors is

$$c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

where $c_1, c_2 \in \mathbb{R}$. Hence the span of these 2 vectors is the whole $\mathbb{R}^2$ space. This means that any third unique vector in the $\mathbb{R}^2$ space can be represented as a linear combination of these 2 basis vectors, hence making all 3 linearly dependent. If the subset contains just 1 vector, then it is a linearly independent subset but here we are looking for the maximum number of unique vectors that are linearly independent in a subset of $\mathbb{R}^2$. Since every basis of a subspace consists of the same number of linearly independent vectors the maximum number of unique vectors that a linearly independent subset of $\mathbb{R}^2$ can have is 2.

8. (a) Determine if the vectors $\{v_1, v_2, v_3\}$ are linearly independent, where

$v_1 = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}$, $v_2 = \begin{bmatrix} 7 \\ 2 \\ -6 \end{bmatrix}$, $v_3 = \begin{bmatrix} 9 \\ 4 \\ -8 \end{bmatrix}$

Justify each answer

**Solution:** Take a matrix $A$, whose columns are made up of the vectors given. Let $A = \begin{bmatrix} 5 & 7 & 9 \\ 0 & 2 & 4 \\ 0 & -6 & -8 \end{bmatrix}$. If the columns are linearly independent, the only solution of $Av = 0$ is $v = 0$, where $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3. \end{bmatrix}$. We want to find $v$ such that,

$$\begin{bmatrix} 5v_1 + 7v_2 + 9v_3 = 0 \\ 2v_2 + 4v_3 = 0 \\ -6v_2 - 8v_3 = 0 \end{bmatrix}$$

From equation 2 and 3, we obtain

$$3(2v_2 + 4v_3) + (-6v_2 - 8v_3) = 0$$

$$\implies 4v_3 = 0$$

$$v_3 = 0$$

Hence substituting back into equation 2, we get $v_2 = 0$. Substituting back in equation 1, we get $v_1 = 0$. Hence $v = 0$, which means that the vectors are linearly independent.

(b) Prove that each set {f, g} is linearly independent in the vector space of all functions from $\mathbb{R}^+$to $\mathbb{R}$.

1. f(x) = x and g(x) = $\frac{1}{x}$
2. f(x) = cos(x) and g(x) = sin(x)
3. f(x) = $e^x$ and g(x) = ln(x)

**Solution:** We can use the Wronskian matrix to determine linear dependence or independence of functions. It is defined as

$$W = \begin{bmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{bmatrix}$$

If $|W| \neq 0$ for some $x$ in the domain, then the functions are linearly independent.
a)

$$W = \begin{bmatrix} x & \frac{1}{x} \\ 1 & \frac{-1}{x^2} \end{bmatrix}$$

Thus $|W| = \frac{-2}{x}$ which is non-zero for some value in the domain(example $x = 1$). Hence the functions are linearly independent.

b)

$$W = \begin{bmatrix} cos(x) & sin(x) \\ -sin(x) & cos(x) \end{bmatrix}$$

Thus $|W| = 1$ which is non-zero for all $x$ in the domain. Hence the functions are linearly independent.

c)

$$W = \begin{bmatrix} e^x & ln(x) \\ e^x & \frac{1}{x} \end{bmatrix}$$

Thus $|W| = e^x(\frac{1}{x} - ln(x))$ which is non-zero for some value in the domain(example $x = 2$). Hence the functions are linearly independent.

9. Let $t_\theta$ be

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{5}$$

(a) Show that $t_{\theta_1+\theta_2} = t_{\theta_1} * t_{\theta_2}$ (* here stands for matrix multiplication).

**Solution:**

$$t_{\theta_1} * t_{\theta_2} = \begin{bmatrix} cos\theta_1 & -sin\theta_1 \\ sin\theta_1 & cos\theta_1 \end{bmatrix} * \begin{bmatrix} cos\theta_2 & -sin\theta_2 \\ sin\theta_2 & cos\theta_2 \end{bmatrix}$$

$$\implies t_{\theta_1} * t_{\theta_2} = \begin{bmatrix} cos\theta_1 cos\theta_2 - sin\theta_1 sin\theta_2 & -cos\theta_1 sin\theta_2 - sin\theta_1 cos\theta_2 \\ sin\theta_1 cos\theta_2 + cos\theta_1 sin\theta_2 & -sin\theta_1 sin\theta_2 + cos\theta_1 cos\theta_2 \end{bmatrix}$$

$$\implies t_{\theta_1} * t_{\theta_2} = \begin{bmatrix} cos(\theta_1+\theta_2) & -sin(\theta_1+\theta_2) \\ sin(\theta_1+\theta_2) & cos(\theta_1+\theta_2) \end{bmatrix}$$

$$\implies t_{\theta_1} * t_{\theta_2} = t_{\theta_1+\theta_2}$$

(b) Show that $t_\theta^{-1} = t_{-\theta}$.

**Solution:**

$$t_\theta^{-1} = \frac{Adj(t_\theta)}{|t_\theta|}$$

$$\implies t_\theta^{-1} = \frac{1}{sin^2\theta + cos^2\theta} \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix}$$

$$\implies t_\theta^{-1} = \begin{bmatrix} cos(-\theta) & -sin(-\theta) \\ sin(-\theta) & cos(-\theta) \end{bmatrix}$$

$$\implies t_\theta^{-1} = t_{-\theta}$$

10. Given matrix has distinct eigenvalues

$$\begin{bmatrix} 1 & 2 & 1 \\ 6 & -1 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

(a) Diagonalize it.

**Solution:** The characteristic polynomial of this matrix is $|A - \lambda I|$.

$$p(\lambda) = |\begin{bmatrix} 1-\lambda & 2 & 1 \\ 6 & -1-\lambda & 0 \\ -1 & -2 & -1-\lambda \end{bmatrix}|$$

$$\implies p(\lambda) = (1-\lambda)((\lambda+1)^2) - 2(-6(1+\lambda)) + (-13-\lambda)$$
$$\implies p(\lambda) = -\lambda^3 - \lambda^2 + 12\lambda$$

Thus the characteristic equation of the matrix is

$$-\lambda^3 - \lambda^2 + 12\lambda = 0$$

where the roots of this equation are the eigenvalues of the matrix. Hence on solving the characteristic equation, the eigenvalues we obtain are $\lambda = 0, \lambda = 3, \lambda = -4$.
The diagonal matrix $\Lambda$ is composed of the eigenvalues

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

(b) Find a basis with respect to which this matrix has that diagonal representation

**Solution:** The basis with respect to which this matrix has diagonal representation is the set of eigenvectors corresponding to the eigenvalues obtained above. Consider $\lambda = 0$, the non-trivial solution of $Av = 0$ is an eigenvector.

$$\begin{bmatrix} v_1 + 2v_2 + v_3 \\ 6v_1 - v_2 \\ -v_1 - 2v_2 - v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

From equation 2 and 3, we obtain the eigenspace as $v_2 = 6v_1$ and $v_3 = -13v_1$. Thus an eigenvector corresponding to $\lambda = 0$ could be

$$v = \begin{bmatrix} 1 \\ 6 \\ -13 \end{bmatrix}$$

Consider $\lambda = 3$, the non-trivial solution of $(A - 3I)v = 0$ is an eigenvector.

$$\begin{bmatrix} -2v_1 + 2v_2 + v_3 \\ 3v_1 - 2v_2 \\ -v_1 - 2v_2 - 4v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

From equation 2 and 3, we obtain the eigenspace as $v_2 = \frac{3}{2}v_1$ and $v_3 = -v_1$. Thus an eigenvector corresponding to $\lambda = 3$ could be

$$v = \begin{bmatrix} 1 \\ \frac{3}{2} \\ -1 \end{bmatrix}$$

Consider $\lambda = -4$, the non-trivial solution of $(A + 4I)v = 0$ is an eigenvector.

$$\begin{bmatrix} 5v_1 + 2v_2 + v_3 \\ 6v_1 + 3v_2 \\ -v_1 - 2v_2 + 3v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

From equation 2 and 3, we obtain the eigenspace as $v_2 = -2v_1$ and $v_3 = -v_1$. Thus an eigenvector corresponding to $\lambda = 0$ could be

$$v = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$$

Thus the basis is $\langle \begin{pmatrix} 1 \\ 6 \\ -13 \end{pmatrix}, \begin{pmatrix} 1 \\ \frac{3}{2} \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} \rangle$

(c) Find the matrices P and $P^{-1}$ to effect the change of basis.

**Solution:** The matrix P is a matrix whose columns are the eigenvectors of the matrix. Hence

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 6 & \frac{3}{2} & -2 \\ -13 & -1 & -1 \end{bmatrix}$$

$$P^{-1} = \frac{Adj(P)}{|P|}$$

$$\implies = \frac{1}{84} \begin{bmatrix} -7 & 0 & -7 \\ 64 & 24 & 16 \\ 27 & -24 & -9 \end{bmatrix}$$

Let us verify the answer obtained. If our calculations are right, then $A = P\Lambda P^{-1}$.

$$P\Lambda P^{-1} = \frac{1}{84} \begin{bmatrix} 1 & 1 & 1 \\ 6 & \frac{3}{2} - 2 & \\ -13 & -1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} -7 & 0 & -7 \\ 64 & 24 & 16 \\ 27 & -24 & -9 \end{bmatrix}$$

$$\implies P\Lambda P^{-1} = \frac{1}{84} \begin{bmatrix} 0 & 3 & -4 \\ 0 & \frac{9}{2} & 8 \\ 0 & -3 & 4 \end{bmatrix} \begin{bmatrix} -7 & 0 & -7 \\ 64 & 24 & 16 \\ 27 & -24 & -9 \end{bmatrix}$$

$$\implies P\Lambda P^{-1} = \frac{1}{84} \begin{bmatrix} 84 & 168 & 84 \\ 504 & -84 & 0 \\ -84 & -168 & -84 \end{bmatrix}$$

$$\implies P\Lambda P^{-1} = \begin{bmatrix} 1 & 2 & 1 \\ 6 & -1 & 0 \\ -1 & -2 & -1 \end{bmatrix} = A$$

11. * **Induced Matrix Norms**

In case you didn't already know, a norm $\|.\|$ is any function with the following properties:

1. $\|x\| \geq 0$ for all vectors $x$.
2. $\|x\| = 0 \iff x = \mathbf{0}$.
3. $\|\alpha x\| = |\alpha| \|x\|$ for all vectors $x$, and real numbers $\alpha$.
4. $\|x + y\| \leq \|x\| + \|y\|$ for all vectors $x, y$.

Now, suppose we're given some vector norm $\|.\|$ (this could be L2 or L1 norm, for example). We would like to use this norm to measure the size of a matrix $A$. One way is to use the corresponding induced matrix norm, which is defined as $\|A\| = \sup_x \{\|Ax\| : \|x\| = 1\}$.

E.g.: $\|A\|_2 = \sup_x\{\|Ax\|_2 : \|x\|_2 = 1\}$, where $\|.\|_2$ is the standard L2 norm for vectors, defined by $\|x\|_2 = \sqrt{x^T x}$.
Note: sup stands for supremum.

Prove the following properties for an arbitrary induced matrix norm:

(a) $\|A\| \geq 0$.

**Solution:** Consider set $S = \{\|Ax\| : \|x\| = 1\}$. Let the dimension of $A$ be $(k, n)$ and the dimension of $x$ be $(n, 1)$. Hence the dimension of $Ax$ will be $(k, 1)$ which means that $Ax$ will be a vector. From the properties of norm of a vector, we know that $\|Ax\| \geq 0$. But $\|A\|$ is nothing but the least upper bound of $S$, thus $\|A\| \geq 0$.

(b) $\|\alpha A\| = |\alpha|\|A\|$ for any real number $\alpha$.

**Solution:** From the arguments above, we know that $Ax$ is a vector.

$$\|\alpha A\| = \sup_x\{\|\alpha Ax\| : \|x\| = 1\}$$

From the properties of inner product of a vector, $\|\alpha x\| = |\alpha|\|x\|$.

$$\|\alpha A\| = \sup_x\{|\alpha|\|Ax\| : \|x\| = 1\}$$

$$\|\alpha A\| = |\alpha| \sup_x\{\|Ax\| : \|x\| = 1\}$$

$$\|\alpha A\| = |\alpha|\|A\|$$

(c) $\|A + B\| \leq \|A\| + \|B\|$.

**Solution:**

$$\|A + B\| = \sup_x\{\big\|(A + B)x\big\| : \|x\| = 1\}$$

$$\implies \|A + B\| = \sup_x\{\|Ax + Bx\| : \|x\| = 1\}$$

But $Ax$ and $Bx$ both are vectors, thus $\|Ax + Bx\| \leq \|Ax\| + \|Bx\|$. Since the matrix norm is the supremum of the vector norm, the inequality sign can be transferred outside as the element which bounds this new set will increase.

$$\implies \|A + B\| \leq \sup_x\{\|Ax\| + \|Bx\| : \|x\| = 1\}$$

$$\implies \|A + B\| \leq \sup_x\{\|Ax\| : \|x\| = 1\} + \sup_x\{\|Bx\| : \|x\| = 1\}$$

$$\implies \|A + B\| \leq \|A\| + \|B\|$$

(d) $\|A\| = 0 \iff A = 0$.

**Solution:** If $\|A\| = 0$, then we are saying that 0 is an upper bound on the set of vector norms. But vector norms have to be greater than or equal to 0. This means that the set of vector norms $Ax$ such that $\|x\| = 1$ must be 0. This can happen only when $Ax = 0$ which implies that $A = 0$ as $\|x\| = 1$.
If $A = 0$, this implies that the only element in the set is 0 as $Ax = 0$ for all $x$ such that $\|x\| = 1$. The supremum of this set is 0 and hence $\|A\| = 0$.
Hence proved $\|A\| = 0 \iff A = 0$.

(e) $\|AB\| \leq \|A\|\|B\|$.

**Solution:** First, we need to show that $\|Ax\| \leq \|A\|\|x\|$. Suppose this does not hold true

$$\|Ax\| > \|A\|\|x\|$$
$$\implies \frac{\|Ax\|}{\|x\|} > \|A\|$$
$$\implies \left\| A\frac{x}{\|x\|} \right\| > \|A\|$$

where $\frac{x}{\|x\|}$ is a unit norm vector. But this definition contradicts the definition of the matrix norm. Hence proved $\|Ax\| \leq \|A\|\|x\|$.

$$\|AB\| = \sup_x \{\|ABx\| : \|x\| = 1\}$$

From the above obtained result

$$\implies \|AB\| \leq \sup_x \{\|A\|\|Bx\| : \|x\| = 1\}$$
$$\implies \|AB\| \leq \|A\| \sup_x \{\|Bx\| : \|x\| = 1\}$$
$$\implies \|AB\| \leq \|A\|\|B\|$$

(f) $\|A\|_2 = \sigma_{\max}(A)$, where $\sigma_{\max}$ is the largest singular value.

**Solution:** Consider the matrix $H = A^*A$. $X$ is a Hermitian matrix($H = H^*$). According to the finite dimensional spectral theorem, any Hermitian matrix can be diagonalized with a unitary matrix(conjugate transpose of a matrix is its inverse) which will result in a diagonal matrix with only real entries. Hence all eigenvalues of $H$ are real and its eigenvectors are linearly independent hence forming a basis. These eigenvectors are orthonormal as well. Let the eigenvectors be $v_1, v_2, \ldots v_n$ and the eigenvalues be $\lambda_1, \lambda_2 \ldots \lambda_n$.

Consider a vector $x$. Since the eigenvectors form a basis, $x$ can be represented as a linear combination of them.

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

$$\|x\|_2 = (x^T x)^{\frac{1}{2}}$$

$$\implies \|x\|_2 = \sqrt{\sum_{i=1}^{n} \alpha_i^2}$$

$$Hx = H(\sum_{i=1}^{n} \alpha_i v_i)$$

$$\implies Hx = \sum_{i=1}^{n} \alpha_i \lambda_i v_i$$

$$\|Ax\|_2 = \sqrt{\langle Ax, Ax\rangle} = \sqrt{\langle x, A^* Ax\rangle} = \sqrt{\langle x, Hx\rangle}$$

$$\implies \|Ax\|_2 = \sqrt{\langle \sum_{i=1}^{n} \alpha_i v_i, \sum_{i=1}^{n} \alpha_i \lambda_i v_i\rangle} = \sqrt{\sum_{i=1}^{n} \alpha_i \overline{\alpha_i \lambda_i}}$$

$$\|Ax\|_2 \leq \max_{1 \leq k \leq n} \sqrt{|\lambda_k|} \|x\|$$

But $\|A\|$ is nothing but $\sup_x \{\|Ax\|_2 : \|x\| = 1\}$. Hence $\|A\|_2 = \max_{1 \leq k \leq n} \sqrt{|\lambda_k|}$. This maximum value will occur for the largest eigen value, hence $\|A\|_2 = \sqrt{\lambda_o}$ where $\lambda_o$ is the largest eigen value.
Hence proved $\|A\|_2 = \sigma_{\max}(A)$.

12. Prove that the eigen vectors of a real symmetric($S_{n*n}$) matrix are linearly independent and form an orthogonal basis for $R^n$.

**Solution:** Consider two distinct eigenvalues $\lambda_1$ and $\lambda_2$ of the real symmetric matrix and their corresponding eigen vectors $v_1$ and $v_2$.

$$Av_1 = \lambda_1 v_1$$

Taking the transpose on both sides

$$v_1^T A^T = \lambda_1 v_1^T$$

Post-multiplying by $v_2$ on both sides, and since $A^T = A$

$$v_1^T A v_2 = \lambda_1 v_1^T v_2$$

$$v_1^T \lambda_2 v_2 = \lambda_1 v_1^T v_2$$
$$(\lambda_1 - \lambda_2) v_1^T v_2 = 0$$

This can either be when $\lambda_1 = \lambda_2$ or when $v_1^T v_2 = 0$. The former is not correct as it contradicts our initial statement, hence eigen vectors of distinct eigen values are orthogonal.
Even if one eigen value has a multiplicity of $r$, there will exist $r$ such linearly independent eigen vectors which can be made orthogonal using Gram-Schmidt orthogonalization. These eigen vectors will be orthogonal and linearly independent to other eigen vectors as shown above. Hence we have a set of $n$ linearly independent and orthogonal eigen vectors. We know that a set of $n$ linearly independent vectors forms a basis for $R^n$, hence the set of eigen vectors forms a basis.

13. **RAYLEIGH QUOTIENT**
Let A be an n×n real symmetric matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_n$ and corresponding orthonormal eigenvectors $v_1, \ldots, v_n$.

(a) Show that

$$\lambda_1 = \min_{x \neq 0} \frac{\langle x, Ax \rangle}{\|x\|^2} \quad and \quad \lambda_n = \max_{x \neq 0} \frac{\langle x, Ax \rangle}{\|x\|^2}. \tag{6}$$

Also, show directly that if $v \neq 0$ minimizes $\frac{\langle x, Ax \rangle}{\|x\|^2}$, then v is an eigenvector of A corresponding to the minimum eigenvalue of A.

**Solution:** Since the $n$ vectors are orthonormal, they form a basis for $R^n$ which means that any vector $x$ can be represented as a linear combination of these vectors. Thus

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

Since $v_i$ are eigen vectors, $Av_i = \lambda_i v_i$ where $\lambda_i$ is the corresponding eigen value.

$$\frac{\langle x, Ax \rangle}{\|x\|^2} = \frac{x^T A x}{\|x\|^2}$$

$$Ax = A(\sum_{i=1}^{n} \alpha_i v_i)$$

$$\implies Ax = \sum_{i=1}^{n} \alpha_i \lambda_i v_i$$

Remember that $Ax$ is still a vector.

$$x^T A x = x^T \sum_{i=1}^{n} \alpha_i \lambda_i v_i$$

$$\implies x^T Ax = \sum_{j=1}^{n} \alpha_j v_j^T \sum_{i=1}^{n} \alpha_i \lambda_i v_i$$

$$\implies x^T Ax = \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_j \alpha_i \lambda_i v_j^T v_i$$

But $v_i$ are orthogonal, hence this equation reduces to

$$x^T Ax = \sum_{i=1}^{n} \lambda_i \alpha_i^2$$

$$\implies \frac{x^T Ax}{\|x\|^2} = \frac{\sum_{i=1}^{n} \lambda_i \alpha_i^2}{\sum_{i=1}^{n} \alpha_i^2}$$

Let $\beta = \frac{x}{\|x\|}$, where each component $0 \leq \beta_i \leq 1$. Thus the equation reduces to

$$\implies \frac{x^T Ax}{\|x\|^2} = \sum_{i=1}^{n} \lambda_i \beta_i^2$$

This is minimum when most weight is given to $\lambda_1$. Thus

$$\min_{x \neq 0} \frac{\langle x, Ax \rangle}{\|x\|^2} = \lambda_1$$

The expression is maximum when most weight is given to $\lambda_n$. Thus

$$\max_{x \neq 0} \frac{\langle x, Ax \rangle}{\|x\|^2} = \lambda_n$$

Now $v$ minimizes the expression hence $v$ must be providing the most weight to $\lambda_1$. This is the case where $x = \alpha_1 v_1$. Without loss of generality we can take $\alpha_1 = 1$. Thus $x = v_1$ which is an eigenvector of $A$.

(b) show that

$$\lambda_2 = \min_{x \perp v_1, x \neq 0} \frac{\langle x, Ax \rangle}{\|x\|^2}. \tag{7}$$

**Solution:** $x \perp v_1$ tells us that the component of $x$ along $v_1$ is 0 i.e $\alpha_1 = 0 \implies \beta_1 = 0$. Thus the equation is reduced to

$$\frac{x^T Ax}{\|x\|^2} = \lambda_2 \beta_2^2 + \lambda_3 \beta_3^2 + \cdots + \lambda_n \beta_n^2$$

This will be minimum when largest weight is given to $\lambda_2$ as that is the smallest remaining real eigenvalue. Hence proved $\lambda_2 = \min_{x \perp v_1, x \neq 0} \frac{\langle x, Ax \rangle}{\|x\|^2}$

14. An m × n matrix has full row rank if its row rank is m, and it has full column rank if its column rank is n. Show that a matrix can have both full row rank and full column rank only if it is a square matrix.

**Solution:** Row rank is the number of independent rows in the matrix which is the rank of the matrix. Column rank is number of independent columns in the matrix which is the rank of the matrix. In this case, it is full column rank and full row rank which implies column rank $= n$ and row rank $= m$. But column rank is the rank of the matrix and row rank is also the rank of the matrix. Hence $m = n$, which is the condition for a square matrix.

15. Let A be a $m \times n$ matrix, and suppose $\vec{v}$ and $\vec{w}$ are orthogonal eigenvectors of $A^T A$. Show that $A\vec{v}$ and $A\vec{w}$ are orthogonal.

**Solution:** Let the eigenvalues corresponding to $\vec{v}$ and $\vec{w}$ be $\lambda_1$ and $\lambda_2$.Hence

$$A^T A\vec{v} = \lambda_1 \vec{v} \tag{8}$$

$$A^T A\vec{w} = \lambda_2 w \tag{9}$$

Since $\vec{v}$ and $\vec{w}$ are orthogonal, their inner product is 0.

$$\langle \vec{v}, \vec{w} \rangle = 0 \implies (\vec{v})^T \vec{w} = 0 \tag{10}$$

Consider the inner product of $A\vec{v}$ and $A\vec{w}$.

$$\langle A\vec{v}, A\vec{w} \rangle = (A\vec{v})^T (A\vec{w})$$

$$\langle A\vec{v}, A\vec{w} \rangle = (\vec{v})^T (A^T A\vec{w})$$

But from equation 8, we can substitute and write the equation as

$$\langle A\vec{v}, A\vec{w} \rangle = (\vec{v})^T (\lambda_2 \vec{w})$$

$$\langle A\vec{v}, A\vec{w} \rangle = \lambda_2 (\vec{v})^T \vec{w}$$

Thus from equation 10,

$$\langle A\vec{v}, A\vec{w} \rangle = 0$$

Hence proved that $A\vec{v}$ and $A\vec{w}$ are orthogonal.

16. Let $u_1, u_2, ...., u_n$ be a set of $n$ orthonormal vectors. Similarly let $v_1, v_2, ...., v_n$ be another set of $n$ orthonormal vectors.

(a) Show that $u_1 v_1^T$ is a rank-1 matrix.

**Solution:**

$$u_1 v_1^T = \begin{bmatrix} v_{11}u_1 & v_{12}u_1 & \dots & v_{1k}u_1 \end{bmatrix}$$

where $v_1 = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1k} \end{bmatrix}^T$.

Consider the $i^{th}$ column in this matrix $v_{1i}u_1$. It is just a scaled multiple of the first column by a scaling factor of $\frac{v_{1i}}{v_{11}}$. Hence they are linearly dependent. All of the columns are just a scaled version of the first column hence making all of them linearly dependent. So, there is just 1 linearly independent vector in the matrix hence it is a rank-1 matrix.

(b) Show that $u_1 v_1^T + u_2 v_2^T$ is a rank-2 matrix.

**Solution:**

$$A = u_1 v_1^T + u_2 v_2^T = \begin{bmatrix} v_{11}u_1 + v_{21}u_2 & v_{12}u_1 + v_{22}u_2 & \dots & v_{1k}u_1 + v_{2k}u_2 \end{bmatrix}$$

where $v_1 = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1k} \end{bmatrix}^T$ and $v_2 = \begin{bmatrix} v_{21} & v_{22} & \dots & v_{2k} \end{bmatrix}^T$.

Assume that the rank of this matrix is 1. Hence there is only 1 linearly independent vector in this matrix. Without loss of generality, we can assume the first vector to be linearly independent. This means that each other vector is a scaled version of the first vector. We get the set of equations

$$\begin{bmatrix} v_{12} = \alpha_1 v_{11} & v_{13} = \alpha_2 v_{11} & \dots & v_{1k} = \alpha_{k-1} v_{11} \end{bmatrix}$$

$$\begin{bmatrix} v_{22} = \alpha_1 v_{21} & v_{23} = \alpha_2 v_{21} & \dots & v_{2k} = \alpha_{k-1} v_{21} \end{bmatrix}$$

Note that the scaling constant are the same as the whole column is just a scaled version.

Let us take the dot product of $v_1$ and $v_2$.

$$\langle v_1, v_2 \rangle = \sum_{i=1}^{k} v_{1i} v_{2i}$$

But from our assumption we can write the summation as

$$\langle v_1, v_2 \rangle = v_{11} v_{21} (1 + \alpha_1^2 + \alpha_2^2 + \dots + \alpha_{k-1}^2)$$

But this can never be true as $v_1$ and $v_2$ are orthonormal($\langle v_1, v_2 \rangle = 0$)(We are not considering the case when $v_{11} = 0$ or $v_{21} = 0$ as that will lead to $v_1$ or $v_2$

becoming 0), hence this matrix cannot be a rank 1 matrix. Hence rank($A$)$\geq 2$.

Let $U = \begin{bmatrix} u_1 & u_2 \end{bmatrix}$ and $V = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$. Thus $A = UV^T$. We know that rank($UV^T$) $\leq min$(rank($U$),rank($V^T$)). Hence rank($A$) $\leq 2$. Thus, from our previous result obtained, rank($A$)$= 2$.

(c) Show that $\sum_{i=1}^{n} u_i v_i^T$ is a rank-n matrix.

**Solution:** Let $U = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix}$. Let $V = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$

$$A = \sum_{i=1}^{n} u_i v_i^T = UV^T$$

Because $v$ is a set of orthonormal vectors :

$$Av_i = u_i$$

Thus $u_i$ lies in the column space of $A$. We have $n$ such $u_i$. But since all of these $u_i$ are orthonormal, the rank of the column space will be greater than $n$ i.e rank($A$)$\geq n$.
We know that rank($UV^T$) $\leq min$(rank($U$),rank($V^T$)). Hence rank($A$) $\leq n$.
Hence from both the results obtained from above, rank($A$) $= n$.