1. (1 point) Honor code?

Eigenstory: Special Properties

2. (1 point) (a) Give a 3×3 matrix such that any two of it's eigenvectors corresponding to distinct eigenvalues are independent. Also, write the eigenvectors and their corresponding eigenvalue.

Solution:

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

respective eigen values $\lambda 1=1$, eigen values $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

For $\lambda 2=2$, eigen values $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

For $\lambda 2=3$, eigen values $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

(b) Give a 3×3 matrix (not Identity matrix) such that any two of it's eigenvectors corresponding to non-distinct eigenvalues are independent. Again, write the eigenvectors and their corresponding eigenvalue.

Solution:

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

respective eigen values $\lambda 1=1$, eigen values $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

$\lambda 2=2$, eigen values $\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$

3. (2 points) (a) Let A be a $K \times K$ square matrix. Prove that a scalar λ is an eigenvalue of A if and only if it is an eigenvalue of $A^\top$.

Solution: We have given that A is a $K \times K$ square matrix which can be written as below

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1k} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2k} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3k} \\ \dots & \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & a_{k3} & \dots & a_{kk} \end{bmatrix}$$

We can write A^T as below

$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{k1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{k2} \\ a_{13} & a_{23} & a_{33} & \dots & a_{k3} \\ \dots & \dots & \dots & \dots & \dots \\ a_{1k} & a_{2k} & a_{3k} & \dots & a_{kk} \end{bmatrix}$$

We have to prove that a scalar λ is an eigenvalue of A if and only if it is an eigenvalue of $A^\top$. This can be proved in two parts as below

1. **If part** A scalar λ is an eigenvalue of A if it is an eigenvalue of $A^\top$
2. **Only if part** If a scalar λ is an eigenvalue of A then it is an eigenvalue of $A^\top$

Now we know that, for any matrix, roots of the characteristics polynomial give the eigenvalues(i.e. solving $|A - \lambda I| = 0$). Hence, if the matrices A and A^T have the same characteristic polynomial, then if any eigenvalue is root of characteristic polynomial of A then it will also be root of characteristic polynomial of A^T and on the other hand if any eigenvalue is root of characteristic polynomial of A^T then it will also be root of characteristic polynomial of A
Hence, if the matrices A and A^T have the same characteristic polynomial, then they will be having the same eigenvalues.
So we have to show that the characteristic polynomial of A is the same as the characteristic polynomial of A^T for the eigen value λ.
Characteristic polynomial of A for the eigen value λ i.e. P($A(\lambda)$) can be written as below

$$P(A(\lambda)) = det(A - \lambda I) \tag{1}$$

Characteristic polynomial of A^T for the eigen value λ i.e. P($A^T(\lambda)$) can be written as below

$$P(A^T(\lambda)) = det(A^T - \lambda I)$$

By transpose property, we know that the transpose of identity matrix is I itself i.e. $I = I^T$. We get,

$$P(A^T(\lambda)) = det(A^T - \lambda I^T)$$

$$P(A^T(\lambda)) = det((A - \lambda I)^T)$$

By the property of determinant , we know that for any square matrix

$$det(A) = det(A^T)$$

Using this we get

$$P(A^T(\lambda)) = det((A - \lambda I)^T)$$
$$P(A^T(\lambda)) = det((A - \lambda I)) \tag{2}$$

Hence by equation(1) and equation(2), we get

$$P(A^T(\lambda)) = P(A(\lambda))$$

Hence , we got the characteristic polynomial of both A and A^T same. The eigen values are the roots of the characteristics polynomial.So the eigen values of both A and A^T are same.

Thus, if for any scalar λ which is an eigen value of A will also be the eigen value of A^T and if scalar λ is an eigen value of A^T then it will also be the eigen value of A.

Hence we proved that a scalar λ is an eigenvalue of A if and only if it is an eigenvalue of $A^\top$

(b) The product of the eigenvalues of a matrix is equal to its determinant. Prove that the diagonal elements of a triangular matrix are equal to its eigenvalues.

Solution: Let A be an lower triangular matrix or an Upper triangular matrix.

$$A = \begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{12} & a_{22} & 0 & \dots & 0 \\ a_{13} & a_{23} & a_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{nn} \end{bmatrix} \text{ or } \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

We know that the roots of the characteristic polynomial of A will give the eigenvalues of A. Thus we can obtain eigenvalues of A by solving the following characteristic equation

$$|A - \lambda I| = 0 \tag{1}$$

$A - \lambda I$ will also be a triangular matrix i.e. either lower or upper triangular matrix as we are subtracting λ from diagonal elements of A only and all off diagonal entries remains same.

$$A = \begin{bmatrix} a_{11}-\lambda & 0 & \dots & 0 \\ a_{12} & a_{22}-\lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn}-\lambda \end{bmatrix} \text{ or } \begin{bmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ 0 & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn}-\lambda \end{bmatrix}$$

By the property of determinant, we know that the determinant of triangular matrix either upper or lower is the product of the diagonal entries.

$$|A - \lambda I| = (a_{11} - \lambda)(a_{22} - \lambda)....(a_{nn} - \lambda) \tag{2}$$

Here n is the number of diagonal entries in the matrix and a_{ij} is the $(ij)^{th}$ element of matrix A.

Solving equation 1^{st} and 2^{nd} we get,

$$|A - \lambda I| = 0$$

$$(a_{11} - \lambda)(a_{22} - \lambda)....(a_{nn} - \lambda) = 0$$

Roots of above equations are the eigen values of A as below

$$\lambda = a_{11}, a_{22}, ..., a_{nn}$$

We can observe that the eigen values of A i.e. the given triangular matrix are simply the diagonal entries of given triangular matrix.
Hence we proved that the diagonal elements of a triangular matrix are equal to its eigenvalues.

4. (2 points) Let A be a $K \times K$ matrix. Let λ_k be one of the eigenvalues of A. Then prove that the geometric multiplicity of λ_k is less than or equal to its algebraic multiplicity.

Solution: We have given that A is a $K \times K$ matrix and λ_k be one of its eigenvalues of A.
We have to prove that the geometric multiplicity of λ_k is less than or equal to its algebraic multiplicity.
Let Geometric multiplicity of λ_k be n, i.e. λ_k has corresponding n linearly independent eigenvectors. Let these n eigenvectors be $v_1, v_2, ..., v_n$.
Since we know that $v_1, v_2, ..., v_n$ are the eigenvectors corresponding to λ_k eigenvalue of A, then we can write as below

$$Av_1 = \lambda_k v_1$$

$$Av_2 = \lambda_k v_2$$

$$\dots$$

$$Av_n = \lambda_k v_n$$

.

Let V be a $K \times K$ matrix described as below

1. First n column vectors are the eigenvectors corresponding to λ_k, i.e. $v_1, v_2, \ldots, v_n$.
2. The columns from $n+1$ to k are some randomly chosen K dimensional vectors $\mathrm{u}_{n+1}, \mathrm{u}_{n+2}, \ldots, \mathrm{u}_k$ such that all these vectors are linearly independent to each other and to the n eigenvectors of λ_k
3. As all the chosen columns are linearly independent, hence it makes V invertible and hence the rank of the matrix V will be K

Let us define M as a $K \times K$ matrix as below

$$M = V^{-1}AV$$

Here, M is similar to A, and by the property of similar matrices M has the same eigenvalues as A. Hence λ_k is an eigenvalue for M as well.

We can show matrix M as below

$$\begin{aligned} M &= V^{-1}AV \\ &= V^{-1}A \begin{bmatrix} v_1 & v_2 & v_3 & \ldots & v_n & \mathrm{u}_{n+1} & \ldots & \mathrm{u}_k \end{bmatrix} \\ &= V^{-1} \begin{bmatrix} Av_1 & Av_2 & Av_3 & \ldots & Av_n & A\mathrm{u}_{n+1} & \ldots & A\mathrm{u}_k \end{bmatrix} \end{aligned}$$

Since we know that $Av_i = \lambda_k v_i$

$$M = \begin{bmatrix} V^{-1}\lambda_k v_1 & V^{-1}\lambda_k v_2 & \ldots & V^{-1}\lambda_k v_n & V^{-1}A\mathrm{u}_{n+1} & \ldots & V^{-1}A\mathrm{u}_k \end{bmatrix} \tag{1}$$

Now we know that $V^{-1}V = I$, where I is the identity matrix. Hence we get,

$$V^{-1}V = I$$

$$V^{-1} \begin{bmatrix} v_1 & v_2 & v_3 & \ldots & v_n & \mathrm{u}_{n+1} & \ldots & \mathrm{u}_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

$$\begin{bmatrix} V^{-1}v_1 & V^{-1}v_2 & \ldots & V^{-1}v_n & V^{-1}\mathrm{u}_{n+1} & \ldots & V^{-1}\mathrm{u}_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

Now for the k-dimensional space, let $b_1, b_2, \dots b_k$ be the k standard basis vectors as below

$$b_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, b_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, ..., b_k = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Then we can observe that $V^{-1}v_1 = b_1$, $V^{-1}v_2 = b_2$, $\dots$ $V^{-1}v_n = v_n$.
Let us use this in the matrix M from equation(1) as below

$$M = \begin{bmatrix} V^{-1}\lambda_k v_1 & V^{-1}\lambda_k v_2 & \dots & V^{-1}\lambda_k v_n & V^{-1}A\mathrm{u}_{n+1} & \dots & V^{-1}A\mathrm{u}_k \end{bmatrix}$$

$$M = \begin{bmatrix} \lambda_k * b_1 & \lambda_k * b_2 & \dots & \lambda_k * b_n & V^{-1}A\mathrm{u}_{n+1} & \dots & V^{-1}A\mathrm{u}_k \end{bmatrix}$$

$$M = \begin{bmatrix} \lambda_k & 0 & 0 & \dots & 0 & * & \dots & * \\ 0 & \lambda_k & 0 & \dots & 0 & \vdots & Q_1 & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_k & * & \dots & * \\ 0 & 0 & 0 & \dots & 0 & * & \dots & * \\ 0 & 0 & 0 & \dots & 0 & \vdots & Q_2 & \vdots \\ 0 & 0 & 0 & \dots & 0 & * & \dots & * \end{bmatrix}$$

Here,it can be observed that the columns $n+1$ to K are the columns which are formed due to the product of $V^{-1}A\mathrm{u}_{n+i} \forall i = 1$ to $K-n$.
To find the eigenvalues of M, we can solve the characteristic equation $\det(M - \lambda I) = 0$ as below

$$det(M - \lambda I) = 0$$

Using Cofactor expansion, we get the following determinant

$$(\lambda_k - \lambda)^n * det(Q_2 - \lambda * I_{K-n}) = 0$$

There are 3 observations as below

1. $(\lambda_k - \lambda)^n = 0$ implies that λ_k has an Algebraic Multiplicity(AM) of n
2. $det(Q_2 - \lambda * I_{k-n}) = 0$ implies that λ can take ≥ 1 different eigenvalues in addition to λ_k, out of which some might or not be the same as λ_k
3. Hence the Algebraic multiplicity(AM) of λ_k is ALTEAST n

Now, we know that the eigenvalues of M are the same as the eigenvalues of A, hence the above 3 observations also apply to A also.
Hence, we can enlist the facts as below

1. Geometric multiplicity of λ_k is n
2. Algebraic multiplicity of $\lambda_k \geq n$

Hence we proved that for any given eigenvalue λ_k, the Geometric Multiplicity is less than or equal to the Algebraic Multiplicity i.e. $GM \leq AM$

5. (1 point) Prove if A and B are positive definite then so is A + B.

Solution: We have to prove that if A and B are positive definite then so is A + B.
Matrix A is positive definite if the quadratic form of A is always positive $x^T Ax > 0$ for any $x \neq 0$
If A is symmetric and positive definite then its eigenvalues are positive
We know that, if matrix A is positive definite, then for all vectors $x \neq 0$ we have

$$x^T Ax > 0$$

Also if matrix B is positive definite, then for all vectors $x \neq 0$ we have

$$x^T Bx > 0$$

Now in order to show $A+B$ positive definite, then for all vectors $x \neq 0$, we will have to show that $x^T(A+B)x > 0$
Now let us derive $x^T(A+B)x$ for all vectors $x \neq 0$ as below

$$x^T(A+B)x = x^T Ax + x^T Bx$$

Since we have $x^T Ax > 0$ and $x^T Bx > 0$ thus both the terms in the above equation is greater than 0 so there addition will also be greater than 0.

$$x^T(A+B)x = x^T Ax + x^T Bx > 0 + 0 > 0$$
$$x^T(A+B)x > 0$$

Hence for all vectors $x \neq 0$, we got $x^T(A+B)x > 0$, so $A+B$ is positive definite
Hence if A and B are positive definite then so is A + B.

Eigenstory: Special Matrices

6. (2 points) Consider the matrix $R = I - 2\mathbf{u}\mathbf{u}^\top$ where $\mathbf{u}$ is a unit vector $\in \mathbb{R}^n$.

(a) Show that R is symmetric and orthogonal. (How many independent vectors will R have?)

Solution:

Solution: 1. Symmetric: $R^\top = (I - 2\mathbf{u}\mathbf{u}^\top)^\top$
$\Rightarrow R^\top = I^\top - (2\mathbf{u}\mathbf{u}^\top)^\top$ — $((A+B)^\top = A^\top + B^\top)$
$\Rightarrow R^\top = I - 2(\mathbf{u}^\top)^\top\mathbf{u}^\top$ — $(I^\top = I$ and $(AB)^\top = B^\top A^\top)$
$\Rightarrow R^\top = I - 2(\mathbf{u}\mathbf{u}^\top)$
$\Rightarrow R^\top = R$
$\Rightarrow$ R is symmetric.

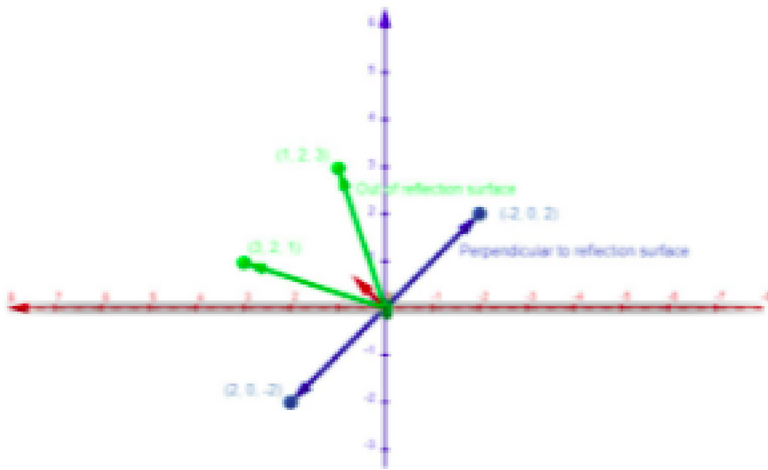
2. Orthogonal: $R^\top R = RR = R^2$—-(Because $R^\top = R$)
$\Rightarrow R^\top R = (I - 2\mathbf{u}\mathbf{u}^\top)^2$
$\Rightarrow R^\top R = I^2 - 4\mathbf{u}\mathbf{u}^\top + 4\mathbf{u}\mathbf{u}^\top\mathbf{u}\mathbf{u}^\top$
$\Rightarrow R^\top R = I - 4 + 4$——(Because $I^2 = I$ and $\mathbf{u}\mathbf{u}^\top = 1$)
$\Rightarrow R^\top R = I$
$\Rightarrow R$ is orthogonal.

R will have n independent eigenvectors because R is symmetric and has dimensions $n \times n$ since $u \in \mathbb{R}^n$

(b) Let $\mathbf{u} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix}$. Draw the line passing through this vector in geogebra (or any tool of your choice). Now take any vector in $\mathbf{R}^3$ and multiply it with the matrix R (i.e., the matrix R as defined above with $\mathbf{u} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix}$). What do you observe or what do you think the matrix R does or what would you call matrix R? (Hint: the name starts with R)

Solution:

Solution:
Below is the image from geogebra. The red vector is u.
The green colored vectors are outside the surface formed by u. The green colored vectors are reflection of each other over the surface formed by u.
The blue colored vector are perpendicular to surface formed by u. The blue colored vectors are reflection of each other over the surface formed by u.



$R^n = I$. This is because here R is a **Reflection matrix**. It means R maps any vector into another vector which is the reflection of input vector over the plane $u_1x + u_2y + u_3z = 0$.

(c) Compute the eigenvalues and eigenvectors of the matrix R as defined above with $\mathbf{u} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix}$

Solution:

Solution:

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

$(R - \lambda I) = \begin{bmatrix} -\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{bmatrix}$

$det(R - \lambda I) = 0 \longrightarrow \lambda^2(1-\lambda) - 1(1-\lambda) = 0$

Therefore, the eigen values of R are $\lambda = 1, 1, -1$

Eigen vector corresponding to $\lambda = 1$ is $\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right)$

Eigen vector corresponding to $\lambda = -1$ is $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

(d) I believe that irrespective of what $\mathbf{u}$ is any such matrix R will have the same eigenvalues as you obtained above (with one of the eigenvalues repeating). Can you reason why this is the case? (Hint: think about how we reasoned about the eigenvectors of the projection matrix P even without computing them.)

Solution:

Solution:

Any vector x can be in three regions with respect to reflection surface.

Here, R is reflection matrix with reflection surface being a plane with equation $u_1x + u_2y + u_3z = 0$

1. Vector x is contained in the reflection surface)

In such cases, $Rx = 1x$. This is because, reflection of x over the reflection surface is the vector itself. So in such cases, eigen value is 1.

2. Vector x is in the space perpendicular to reflection surface

In such cases, $Rx = -1x$. This is because, reflection of x over the reflection surface is the negative of the vector itself(this is shown as blue vector in b part of question). So, in such cases, eigen value is -1

3. Vector x is not present in the reflection surface nor perpendicular to the reflection surface.

In such cases, x cannot be expressed as $Rx = \lambda x : \lambda > 0$

Therefore, possible eigen values are $1, -1$.

7. (2 points) Let Q be a $n \times n$ real orthogonal matrix (i.e., all its elements are real and its columns are orthonormal). State with reason whether the following statements are True or False (provide a proof if the statement is True and a counter-example if it is False).

(a) If λ is an eigenvalue of Q then $\lambda^2 = 1$. (0.5 marks)

Solution:

Solution:

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

Now let's find the eigenvalue of Q.

$$det(Q - \lambda I) = 0$$

$$\begin{vmatrix} \frac{1}{\sqrt{2}} - \lambda & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \lambda \end{vmatrix} = 0$$

$$(\frac{1}{\sqrt{2}} - \lambda)^2 + \frac{1}{2} = 0$$

$$\frac{1}{2} - \sqrt{2}\lambda + \lambda^2 + \frac{1}{2} = 0$$

$$\lambda^2 - \sqrt{2}\lambda + 1 = 0$$

Now let's find roots $\sqrt{D} = \sqrt{b^2 - 4ac} = \sqrt{2 - 4(1)(1)} = \sqrt{-2} = i\sqrt{2}$
$\lambda = \frac{-b+\sqrt{D}}{2a}$ and $\lambda = \frac{-b-\sqrt{D}}{2a}$
$\lambda = \frac{\sqrt{2}+i\sqrt{2}}{2}$ and $\lambda = \frac{\sqrt{2}-i\sqrt{2}}{2}$

$\lambda = \frac{1+i}{\sqrt{2}}$ and $\lambda = \frac{1-i}{\sqrt{2}}$

Now take λ^2

$$\lambda^2 = (\frac{1+i}{\sqrt{2}})^2$$

$$\lambda^2 = \frac{1 + 2i + i^2}{2}$$

$$\lambda^2 = \frac{1 + 2i + -1}{2}$$

$$\lambda^2 = \frac{2i}{2}$$

$$\lambda^2 = i \neq 1$$

Hence it's False.

(b) The eigen vectors of Q are orthogonal. Just state yes or no.(0.25 marks)

Solution:

Solution: False,

Simple example would be indentity matrix. Identity matrix is Orthogonal. So,

$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
All the vectors in 2D space are eigen vectors of the above matrix.
Lets take two eigen vectors $X_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $X_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$
The above two eigen vectors are not orthogonal to each other.

(c) Q is always diagonalizable, and if it is diagonisable only under some particular condition, give prove for that.(1.25 marks)

Solution:

Solution:
A real orthogonal matrix Q containing all real elements over $\mathbb{R}$ is diagonalizable if and only if Q is symmetric ($Q^\top = Q$).
Now we need to prove both the parts.

Part 1: If a real orthogonal matrix Q is diagonalizable, then Q is symmetric.

Let us assume that Q is a real orthogonal matrix which is diagonalizable, then there exists an orthogonal matrix U and a diagonal matrix D both over $\mathbb{R}$ such that $Q = UDU^{-1} = UDU^\top (\because U^{-1} = U^\top)$
Now, $Q^\top = (UDU^\top)^\top$
$\Rightarrow Q^\top = UD^\top U^\top$
$\Rightarrow Q^\top = UDU^\top$

$\Rightarrow Q^\top = Q$.
Hence, Q is symmetric.

Part 2: If a real orthogonal matrix Q is symmetric, then Q is diagonalizable.

We will proof the above statement by induction.
It can be observed that this statement is true for 1×1 matrix Q. If $Q = [q]$ then $Q = [1][q][1] = UQU^T$
We assume that this is true for $(n-1) \times (n-1)$ matrix.
Now, we have to show that all $n \times n$ orthogonal matrix which are symmetric are diagonalizable.
Now we consider a $n \times n$ symmetric matrix Q which has real eigenvalues λ_i with real eigenvector v_1. The orthonormal basis in R^n is $P = [v_1 v_2 \ldots v_n]$.
The matrix $P^{-1}QP$ is symmetric since P is orthogonal.
Let, e_1 be an orthonormal basis vector, then the first column is

$$P^{-1}QPe_1 = P^{-1}Qv_1 = P^{-1}\lambda_1 v_1 = \lambda_1 P^{-1} v_1 = \lambda_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Thus, we can write $P^{-1}QP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & B \end{bmatrix}$ where B is $(n-1) \times (n-1)$ symmetric matrix.
Now, according to our assumption B is orthogonally diagonalizable. So, there exists an orthogonal matrix A and diagonal matrix D' such that $B = AD'A^{-1}$
Now, for $n \times n$ matrix $R = \begin{bmatrix} 1 & 0 \\ 0 & A \end{bmatrix}$
$R^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & A^{-1} \end{bmatrix}$ (Since $\begin{bmatrix} 1 & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & A^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$)
Let, $U = PR$ (U is orthogonal matrix)
$U^{-1}QU = (R^{-1}P^{-1})Q(PR)$
$\Rightarrow R^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & B \end{bmatrix} R$ ($\because P^{-1}QP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & B \end{bmatrix}$)
$\Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & A^{-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & A \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & A^{-1}BA \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & D' \end{bmatrix}$
$\Rightarrow Q = U \begin{bmatrix} \lambda_1 & 0 \\ 0 & D' \end{bmatrix} U^{-1}$
Since , $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & D' \end{bmatrix}$ is a diagonal matrix, thus Q is diagonalizable.
Hence $Q = UDU^{-1} = UDU^T$.
A real orthogonal matrix Q containing all real elements over $\mathbb{R}$ is diagonalizable if and only if Q is symmetric ($Q^\top = Q$). It is proved.

8. (1½ points) Any rank one matrix can be written as $\mathbf{u}\mathbf{v}^\top$.

   (a) Prove that the eigenvalues of any rank one matrix are $\mathbf{v}^\top \mathbf{u}$ and 0.

Solution:

Solution: Given, A=$uv^\top$

Let x be the eigen vector corresponding to non-zero eigen value $v^\top u$

then we have

$uv^\top x = v^\top ux$

$$\implies x = \frac{u(v^\top x)}{v^\top u}$$

We know that x,u and v are column vectors

then we have $v^\top x$ and $v^\top u$ are scalars then $\frac{v^\top x}{v^\top u}$ will also be scalar say c

$\implies x = cu$

we know that if cu is an eigen vector then u is also an eigen vector

$\implies$ u is the eigen vector of A corresponding to non-zero eigen value $v^\top u$

We have (n-1) zeros as eigen values for matrix A let x be the corresponding eigen vector then

$uv^\top x = 0$

we know u, v and x cannot be zero vectors

Hence $v^\top x$ must be zero

$\implies v^\top x = 0$

i.e., x (required eigen vector)is orthogonal to the vector v

Now x can be any vector that lies in the hyperplane, where the hyperplane is perpendicular to the vector v

In an n dimensional space we have (n-1) dimensional hyperplane and n-1 independent vectors that span this hyperplane

Therefore the required eigen vectors are the basis of the hyperplane which is perpendicular to the vector v

Therefore the eigen vectors are

1.)u and its corresponding eigen value is $v^\top u$

2.)basis vectors of the hyperplane which is perpendicular to vector v and their corresponding eigen values are 0

(b) How many times does the value 0 repeat?

Solution:

Solution: We found in the part(a) of solution that we have $\lambda = 0$ as one of the eigenvalue of rank one matrix A, so the corresponding eigenvectors can be obtained by solving $Ax = 0$
So the eigenvectors corresponding to 0 eigenvalue will be any non zero vector x which is in null space of A.
Now since rank of A is 1, hence by rank nullity theorem the dimension of null space of A will be $n-1$, hence there will be $n-1$ independent vectors in the basis of null space. Hence there will be $n-1$ independent eigenvectors corresponding to 0 eigenvalue
We know that for a eigenvalue λ, Geometric Multiplicity is number of independent eigenvectors for λ, hence in our case for $\lambda = 0$ eigenvalue the Geometric Multiplicity(GM) is $n-1$
We know that for a eigenvalue λ, Algebraic Multiplicity is the number of repetitions of λ among the eigenvalues, hence in our case for $\lambda = 0$, the Algebraic Multiplicity(AM) is number of repetitions of 0 among the eigenvalues
We also know that, for a given eigenvalue λ, geometric multiplicity(GM) is less than equal to the algebraic multiplicity(AM).

$$GM \leq AM$$

Hence for $\lambda = 0$, we get

$$n - 1 \leq AM$$

Hence the algebraic multiplicity i.e. number of repetitions of $\lambda = 0$ among the eigenvalues will be at least $n-1$
Hence A will be having 0 eigenvalue at least $n-1$ times.
Hence 0 repeat at least $n-1$ times
Hence when $v^T u = 0$, 0 repeats n times, but when $v^T u \neq 0$ then 0 repeats $n-1$ times.

(c) What are the eigenvectors corresponding to these eigenvalues?

Solution:

Solution: The eigenvector corresponding to $\lambda = v^\top u$ is u. —-(Already shown in part (a))
The eigenvectors corresponding to $\lambda = 0$ are:
$uv^\top x = 0x$
$\Rightarrow uv^\top x = 0$
$\Rightarrow$ Eigenvectors corresponding to $\lambda = 0$ is the null space of $uv^\top$,i.e., the null space of the rank one matrix.

9. (2 points) Consider a $n \times n$ Markov matrix.

(a) Prove that the dominant eigenvalue of a Markov matrix is 1

Solution:

Solution:

Let $M = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ m_{21} & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nn} \end{bmatrix}$ be a Markov matrix having sum of each row $= 1$ and each element of $M \geq 0$.

Proof (part 1): 1 is an eigenvalue of a Markov matrix

Let us consider a vector $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ such that,

$$
\begin{aligned}
M\mathbf{x} &= \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ m_{21} & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} m_{11} + m_{12} + \dots + m_{1n} \\ m_{21} + m_{22} + \dots + m_{2n} \\ \vdots \\ m_{n1} + m_{n2} + \dots + m_{nn} \end{bmatrix}
\end{aligned}
$$

Since sum of each row is 1 we can write $(m_{i1} + m_{i2} + \cdots + m_{in} = 1)\ \forall i \in \{1, 2, \ldots, n\}$. Using this property in above equation we get,

$$M\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$
$$M\mathbf{x} = \mathbf{x}$$

From above we can say $\mathbf{x}$ is an eigenvector of a Markov matrix M corresponding to eigenvalue 1. Therefore 1 is an eigenvalue of a Markov matrix.

Proof (part 2): Absolute value of any eigenvalue is ≤ 1

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ be an eigenvector of M corresponding to the eigenvalue λ then,

$$M\mathbf{x} = \lambda\mathbf{x}$$
$$\begin{bmatrix} m_{11} & m_{12} & \ldots & m_{1n} \\ m_{21} & m_{22} & \ldots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \ldots & m_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}$$
$$\begin{bmatrix} m_{11}x_1 + m_{12}x_2 + \cdots + m_{1n}x_n \\ m_{21}x_1 + m_{22}x_2 + \cdots + m_{2n}x_n \\ \vdots \\ m_{n1}x_1 + m_{n2}x_2 + \cdots + m_{nn}x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}$$

Let us consider an element of $\mathbf{x}$, x_k such that $|x_k| = max\{|x_1, ||x_2|, \ldots, |x_n|\}$. Then on comparing k^{th} row on both sides we get,

$$\begin{aligned} \lambda x_k &= m_{k1}x_1 + m_{k2}x_2 + \cdots + m_{kn}x_n \\ |\lambda x_k| &= |m_{k1}x_1 + m_{k2}x_2 + \cdots + m_{kn}x_n| \\ |\lambda||x_k| &\leq |m_{k1}x_1| + |m_{k2}x_2| + \cdots + |m_{kn}x_n| \\ &\leq m_{k1}|x_1| + m_{k2}|x_2| + \cdots + m_{kn}|x_n| \qquad [m_{kj} \geq 0] \\ &\leq (m_{k1} + m_{k2} + \cdots + m_{kn})|x_k| \\ &\leq |x_k| \qquad [m_{k1} + m_{k2} + \cdots + m_{kn} = 1] \end{aligned}$$

We can say $|x_k| > 0$ otherwise $\mathbf{x}$ will be a zero vector and contradicts with our assumption that $\mathbf{x}$ is an eigenvector. Therefore using this arguement we can conclude from above equation that,

$$|\lambda| \leq 1$$

Now from Part 1 and Part 2 we can say that 1 is an eigenvalue of a Markov matrix and absolute values of all the eigenvalues are less than or equal to 1 which implies 1 is the dominant eigenvalue of a Markov matrix.

(b) Consider any 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $a + b = c + d = k$. Show that one of the eigenvalues of such a matrix is k. (I hope you notice that a Markov matrix is a special case of such a matrix where $a + b = c + d = 1$.)

Solution:

Solution:

Consider a vector $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$,

Then $\begin{bmatrix} a & b \\ c & d \end{bmatrix} x = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a+b \\ c+d \end{bmatrix} = (a+b) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \because a+b = c+d$

Therefore vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigen vector and (a+b) is an eigen value.

(c) Does the result extend to $n \times n$ matrices where the sum of the elements of a row is the same for all the n rows? (Explain with reason)

Solution:

Solution: Yes
Proof:

Say, matrix A is a $n \times n$ matrix whose sum of elements of rows are equal and it equal to k.
$\therefore \sum_{j=1}^{n} a_{i,j} = k, \forall i \in 1, 2, .., n$

Now, consider a vector $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ .. \\ 1 \end{bmatrix}$, all ones vector of size $n \times 1$.

Now, $A\mathbf{x} = \begin{bmatrix} \sum_{j=1}^{n} a_{1,j} \\ \sum_{j=1}^{n} a_{2,j} \\ \cdots \\ \sum_{j=1}^{n} a_{n,j} \end{bmatrix}$, which means $(A\mathbf{x})_{i,1} = \sum_{j=1}^{n} a_{i,j}$

$\implies A\mathbf{x} = (\sum_{j=1}^{n} a_{1,j})\mathbf{x}, \qquad \because$ sum of elements of rows are equal.

This implies $(\sum_{j=1}^{n} a_{1,j})$ is an eigen value of matrix A, defined as above.

(d) What is the corresponding eigenvector?

Solution:

Solution:
The corresponding eigen vector is all ones vector of size $n \times n$.

$\implies$ eigen vector = $\begin{bmatrix} 1 \\ 1 \\ ... \\ 1 \end{bmatrix}$

Proof:
Consider, matrix $A_{n \times n}$ whose sum of all elements of the rows are equal, and

$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ ... \\ 1 \end{bmatrix}$

Now, $A\mathbf{x} = \begin{bmatrix} \sum_{j=1}^{n} a_{1,j} \\ \sum_{j=1}^{n} a_{2,j} \\ ... \\ \sum_{j=1}^{n} a_{n,j} \end{bmatrix}$, which means $(A\mathbf{x})_{i,1} = \sum_{j=1}^{n} a_{i,j}$

$\implies A\mathbf{x} = (\sum_{j=1}^{n} a_{1,j})\mathbf{x}$, $\quad \because$ sum of elements of rows are equal.

This implies $\mathbf{x}$ is an eigen vector for the corresponding eigen value.

Eigenstory: Special Relations

10. (4 points) For each of the statements below state True or False with reason.

(a) If i(complex number) is an eigen value of A , then it follows that i is an eigen value of A^{-1}.

Solution:

Solution:

We know that:

1) If λ is an eigen value of a real matrix B, then complex conjugate of λ is also an eigen value.

Proof:

Say, λ is an eigen value and x is its corresponding eigen vector for B, then $Bx = \lambda x$

Taking conjugate on both side

$\therefore \bar{B}\bar{x} = \bar{\lambda}\bar{x} \implies B\bar{x} = \bar{\lambda}\bar{x}$,because B is real

This implies if λ is an eigen value, then $\bar{\lambda}$ is also an eigen value.

Now consider, real matrix A, then if i is an eigen value

$\implies Ax = ix$, where x is the correspoing eigen vector for i

Multiplying both side with A^{-1} we get,

$A^{-1}Ax = iA^{-1}x \implies A^{-1}x = -ix$

This implies that $-i$ is an eigen value of A^{-1}. And using 1) we know that the conjugate of $-i$ is also an eigen value of A^{-1}

$\implies \bar{-i} = i$

Hence, i is also an eigen value of A^{-1}.

(b) If the characteristic equation of a matrix A is $\lambda^5 + 7\lambda^3 - 6\lambda^2 + 128 = 0$ then sum of eigen values is -7.

Solution: False, trace(A)=-$\frac{coef.\lambda^4}{coef.\lambda^5}$ which is 0.

(c) If A is 3×3 matrix with eigenvector as $\begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$ then $\begin{bmatrix} 16 \\ -12 \\ 8 \end{bmatrix}$ is also an eigen vector of A.

Solution:

Solution: True

Say λ is the corresponding eigen value for the eigen vector $v = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$

Now, consider $A \begin{bmatrix} 16 \\ -12 \\ 8 \end{bmatrix} = 4A \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix} = 4\lambda \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix} = \lambda \begin{bmatrix} 16 \\ -12 \\ 8 \end{bmatrix}$

Hence $\begin{bmatrix} 16 \\ -12 \\ 8 \end{bmatrix}$ is an eigen vector of A.

(d) If A is symmetric matrix then the algebraic and geometric multiplicity is same for every eigen value.

Solution:

Solution: TRUE
If A is a real symmetric matrix, then by Schur's Theorem we can say that Matrix A is always diagonalizable. And a matrix is diagonalizable if and only if for each eigenvalue of the matrix , the algebraic multiplicity of is equal to the geometric multiplicity.
Hence we can conclude that, If A is symmetric matrix then the algebraic and geometric multiplicity is same for every eigen value.

(e) If $\mathbf{x}$ is an eigenvector of A and B then it is also an eigenvector of both AB and BA, even if the eigenvalues of A and B corresponding to $\mathbf{x}$ are different.

Solution:

Solution: True

Say, λ_a is the corresponding eigen value for A and λ_b is the corresponding eigen value for B

$\implies A\mathbf{x} = \lambda_a \mathbf{x}$ and $B\mathbf{x} = \lambda_b \mathbf{x}$

Now consider, $AB\mathbf{x} = A\lambda_b \mathbf{x} = \lambda_a \lambda_b \mathbf{x} \implies \mathbf{x}$ is an eigen vector for AB

Consider, $BA\mathbf{x} = B\lambda_a \mathbf{x} = \lambda_b \lambda_a \mathbf{x} \implies \mathbf{x}$ is an eigen vector for BA

(f) If $\mathbf{x}$ is and eigenvector of A and B then it is also an eigenvector of $A + B$

Solution:

Solution: True

Say, λ_a is the corresponding eigen value for A and λ_b is the corresponding eigen value for B

Now consider $(A+B)\mathbf{x} = A\mathbf{x} + B\mathbf{x} = \lambda_a\mathbf{x} + \lambda_b\mathbf{x} = (\lambda_a + \lambda_b)\mathbf{x}$

This implies $\mathbf{x}$ is also an eigen vector for $A + B$

(g) The non-zero eigenvalues of $AA^\top$ and $A^\top A$ are equal.

Solution:

Solution: True

1) Consider λ_1 is an eigen value and $\mathbf{x}$ is the corresponding eigen vector of AA^T, then $AA^T\mathbf{x} = \lambda_1\mathbf{x}$

Multiplying both side by A^T then we get
$\implies A^TAA^T\mathbf{x} = \lambda_1 A^T\mathbf{x}$

$\implies A^TA(A^T\mathbf{x}) = \lambda_1(A^T\mathbf{x})$.
This means that λ_1 is also an eigen value of A^TA matrix.

2) Consider λ_2 is an eigen value and $\mathbf{y}$ is the corresponding eigen vector of A^TA, then $A^TA\mathbf{y} = \lambda_2\mathbf{y}$

Multiplying both side by A then we get
$\implies AA^TA\mathbf{y} = \lambda_2 A\mathbf{y}$

$\implies AA^T(A\mathbf{y}) = \lambda_2(A\mathbf{y})$.
This means that λ_2 is also an eigen value of AA^T matrix.

Using 1) and 2) we know that all the eigen values of A^TA and AA^T are equal.

(h) The eigenvectors of $AA^\top$ and $A^\top A$ are always same.

Solution:

Solution: False

Consider $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $AA^T = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $A^TA = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

Eigen values and vectors of AA^T are:

eigen vector for eigen value $2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

eigen vector for eigen value $1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

eigen vector for eigen value $0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

Eigen values and vectors of A^TA are:

eigen vector for eigen value $2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

eigen vector for eigen value $1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

Eigenstory: Change of basis

11. (2 points) Consider the following two basis. Basis 1: $\mathbf{u_1} = \frac{1}{5}\begin{bmatrix} 3 \\ 4 \end{bmatrix}, \mathbf{u_2} = \frac{1}{5}\begin{bmatrix} 4 \\ -3 \end{bmatrix}$, and Basis 2: $\mathbf{u_1} = \frac{1}{5}\begin{bmatrix} 3 \\ -4 \end{bmatrix}, \mathbf{u_2} = \frac{1}{5}\begin{bmatrix} -4 \\ -3 \end{bmatrix}$,. Consider a vector $\mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix}$ in Basis 1 (i.e., $\mathbf{x} = a\mathbf{u_1} + b\mathbf{u_2}$). How would you represent it in Basis 2?

Solution:

Solution: We will first convert the vector x, which is in Basis 1 to the vector in standard basis by multiplying vector x by the Basis 1 matrix B_1, where B_1 is given as -

$$B_1 = \frac{1}{5}\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} \tag{11.1}$$

Let x_0 be the vector in standard basis corresponding to the vector x in Basis 1, now x_1 can be obtained by multiplying B_1 and x, i.e., B_1x and is given as -

$$x_0 = B_1x = \frac{1}{5}\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 3a+4b \\ 4a-3b \end{bmatrix} \tag{11.2}$$

Now we have vector x_0 which is in the standard basis, now if we have to represent this vector in Basis 2, then we have to multiply the vector x_0 with the inverse of the Basis 2 matrix, let B_2 be the Basis 2 matrix and it is given as -

$$B_2 = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix} \tag{11.3}$$

Now the inverse of B_2 is represented as B_2^{-1} and given as -

$$B_2^{-1} = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix} \tag{11.4}$$

Now let x_2 be the vector in the Basis 2 corresponding to the vector x in Basis 1. Now x_2 is given as -

$$B_2^{-1}x_0 = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}\frac{1}{5}\begin{bmatrix} 3a+4b \\ 4a-3b \end{bmatrix} \tag{11.5}$$

On solving equation 11.5, we get -

$$x_2 = B_2^{-1}x_0 = \frac{1}{25}\begin{bmatrix} -7a+24b \\ -24a-7b \end{bmatrix} \tag{11.6}$$

12. (1 point) Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors in the standard basis. Let $T(\mathbf{u})$ and $T(\mathbf{v})$ be the representation of these vectors in a different basis. Prove that $\mathbf{u} \cdot \mathbf{v} = T(\mathbf{u}) \cdot T(\mathbf{v})$ if and only if the basis represented by T is an orthonormal basis (i.e., dot products are preserved only when the new basis is orthonormal).

Solution:

Solution: If part: If T is an orthonormal basis, then $\mathbf{u} \cdot \mathbf{v} = T(\mathbf{u}) \cdot T(\mathbf{v})$
$T(u)$ = Representation of u in the basis T = Tu
Now, $T(u).T(v) = (Tu)^\top Tv$——(Definition of dot)
product) $\Rightarrow T(u).T(v) = u^\top T^\top Tv$———$((AB)^\top = B^\top A^\top$
$\Rightarrow T(u).T(v) = u^\top Iv$ ——-(T is orthonormal)
$\Rightarrow T(u).T(v) = u^\top v$
$\Rightarrow T(u).T(v) = u.v$
Hence, dot products are preserved.
Only if part: If $\mathbf{u} \cdot \mathbf{v} = T(\mathbf{u}) \cdot T(\mathbf{v})$, then T is orthonormal.
$u.v = Tu.Tv$
$\Rightarrow u^\top v = (Tu)^\top (Tv)$
$\Rightarrow u^\top v = u^\top T^\top Tv$
This is only possible when $T^\top T = I$, i.e, T is orthonormal. (Because u and v can be any vectors in the standard basis and not just zero vectors).

Eigenstory: PCA and SVD

13. (1 point) We are familiar with the following equation: $A = U \sum V^T$, where A is a real valued $m \times n$ matrix and other symbols have their usual meanings. State how this equation is related to Principal Component Analysis. State the correct dimensions of the three matrices at the RHS of the given equation. (No vague answers please.)

Solution:

Solution: PCA:
PCA aims to find linearly uncorrelated orthogonal axes, which are also known as principal components (PCs) in the m dimensional space to project the data points onto those PCs. The first PC captures the largest variance in the data.
The PCs can be determined via eigen decomposition of the covariance matrix C. After all, the geometrical meaning of eigen decomposition is to find a new coordinate system of the eigenvectors for C through rotations. $C = W\lambda W^{-1}$, Eigendecomposition of the covariance matrix C.

SVD :
SVD is another decomposition method for both real and complex matrices. It decomposes a matrix into the product of two unitary matrices (U, V^*) and a rectangular diagonal matrix of singular values Σ
$X = U\Sigma V^*$

Relationship between PCA and SVD:

PCA and SVD are closely related approaches and can be both applied to decompose any rectangular matrices. We can look into their relationship by performing SVD on the covariance matrix C:

$$\begin{aligned} C &= \frac{X^T X}{n} \\ &= \frac{V\Sigma U^T U \Sigma V^T}{n-1} \\ &= V\frac{\Sigma^2}{n}V^T \\ &= V\frac{\Sigma^2}{n}V^{-1} \end{aligned}$$

(Since V is unitary)
From the above derivation, we notice that the result is in the same form with eigen decomposition of C, we can easily see the relationship between singular values (Σ) and eigenvalues (λ): $\lambda = \frac{\Sigma^2}{n}$
Relationship between eigenvalue and singular values
We can actually perform PCA using SVD, or vice versa.

In fact, most implementations of PCA actually use performs SVD under the hood rather than doing eigen decomposition on the covariance matrix because SVD can be much more efficient and is able to handle sparse matrices.

The correct dimensions of the three matrices at the RHS of the given equation are:

Let A be $m \times n$ matrix: We know that,
A=$U\Sigma D$, where A is a SVD
Dimension of U= $m \times m$
Dimension of Σ= $m \times n$
Dimension of V^T= $n \times n$

$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V_{n\times n}$

14. (1½ points) Consider the matrix $\begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix}$

(a) Find Σ and V, *i.e.*, the eigenvalues and eigenvectors of $A^\top A$

Solution:

$$A = \begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix}$$

Calculating Σ -

Σ(the same dimensions as A) has singular values and is diagonal matrix with its diagonal entries square root of eigenvalues of $A^T A$.

$$A^T A = \begin{bmatrix} 5 & -4 \\ 5 & 4 \end{bmatrix} \begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix} = \begin{bmatrix} 41 & 9 \\ 9 & 41 \end{bmatrix}$$

Eigen values of $A^T A$ is the roots of the characteristic equations. For finding the eigenvalues of $A^T A$ the characteristic equation of $A^T A$ is given as :

$$|A^T A - \lambda I| = 0$$

$$\begin{vmatrix} 41 - \lambda & 9 \\ 9 & 41 - \lambda \end{vmatrix} = 0$$

$$\lambda_1 = 50 \text{ and } \lambda_2 = 32$$

Eigen values of $A^T A$ are $\lambda_1 = 50$ and $\lambda_2 = 32$.

We know that,

$$A^T A = V \Sigma^T \Sigma V$$

Here, $\Sigma^T \Sigma$ is the eigen values of $A^T A$. Thus, Σ is the square roots of the eigen values of $A^T A$.

$$\Sigma = \begin{bmatrix} \sqrt{50} & 0 \\ 0 & \sqrt{32} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix} \tag{Ans}$$

Calculating V -
V is the eigen vectors of $A^T A$. We should have -

$$(A^T A - \lambda I)\mathrm{x} = 0$$

For $\lambda_1 = 50$

$$\begin{bmatrix} 41-50 & 9 \\ 9 & 41-50 \end{bmatrix} \mathrm{x} = 0$$

$$\begin{bmatrix} -9 & 9 \\ 9 & -9 \end{bmatrix} \mathrm{x} = 0$$

The eigen vector corresponding to $\lambda_1 = 50$ is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Normalised eigen vector is $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$

For $\lambda_2 = 32$

$$\begin{bmatrix} 41-32 & 9 \\ 9 & 41-32 \end{bmatrix} \mathrm{x} = 0$$

$$\begin{bmatrix} 9 & 9 \\ 9 & 9 \end{bmatrix} \mathrm{x} = 0$$

The eigen vector corresponding to $\lambda_2 = 32$ is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Normalised eigen vector is $\begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$

Thus,

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{Ans}$$

(b) Find Σ and U, *i.e.*, the eigenvalues and eigenvectors of $AA^\top$

Solution:

Solution: Given,

$$A = \begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix}$$

Calculating Σ -
Σ(the same dimensions as A) has singular values and is diagonal matrix with

its diagonal entries square root of eigenvalues of AA^T.

$$AA^T = \begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix} \begin{bmatrix} 5 & -4 \\ 5 & 4 \end{bmatrix} = \begin{bmatrix} 50 & 0 \\ 0 & 32 \end{bmatrix}$$

This matrix is diagonal matrix with its diagonal entries 50 and 32. We know that, eigen values of diagonal matrix is the same as diagonal entries of that matrix. so,

$$\lambda_1 = 50 \text{ and } \lambda_2 = 32$$

Thus, it shows that Eigen values of AA^T are $\lambda_1 = 50$ and $\lambda_2 = 32$.
We know that,

$$AA^T = U\Sigma\Sigma^T U$$

Here, $\Sigma\Sigma^T$ is the eigen values of AA^T. Thus, Σ is the square roots of the eigen values of AA^T.

$$\Sigma = \begin{bmatrix} \sqrt{50} & 0 \\ 0 & \sqrt{32} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix} \qquad \text{(Ans)}$$

Calculating U -
V is the eigen vectors of AA^T. We should have -

$$(AA^T - \lambda I)\mathrm{x} = 0$$

For $\lambda_1 = 50$

$$\begin{bmatrix} 50-50 & 0 \\ 0 & 32-50 \end{bmatrix} \mathrm{x} = 0$$

$$\begin{bmatrix} 0 & 0 \\ 0 & -18 \end{bmatrix} \mathrm{x} = 0$$

The Normalised eigen vector corresponding to $\lambda_1 = 50$ is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

For $\lambda_2 = 32$

$$\begin{bmatrix} 50-32 & 0 \\ 0 & 32-32 \end{bmatrix} \mathrm{x} = 0$$

$$\begin{bmatrix} 18 & 0 \\ 0 & 0 \end{bmatrix} \mathrm{x} = 0$$

The Normalised eigen vector corresponding to $\lambda_2 = 32$ is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Thus, The columns of the matrix U are the normalized eigen vectors of AA^T:

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{(Ans)}$$

(c) Now compute $U\Sigma V^\top$. Did you get back A? If yes, good! If not, what went wrong?

Solution:

Solution: Given:

$$A = \begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix} \tag{1}$$

From Part a and Part b we have ,

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

$$\Sigma = \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix} \tag{3}$$

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{4}$$

Lets compute $U\Sigma V^\top$ and check whether we get back A or not.

$$U\Sigma V^\top = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T$$

$$U\Sigma V^\top = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$U\Sigma V^\top = \begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix}$$

$$U\Sigma V^\top = A$$

Here by computation of $U\Sigma V^\top$ we get back A. But it's not always the case that we get back A. The reason would be we calculated U and V matrices irrespective of each other.

As we know that,there can be multiple orthogonal basis for any subspace. Thus,there can be multiple orthogonal basis for A^TA and AA^T which leads to have multiple U and V matrix. Thus, choosing U and V irrespective of each other may give A back.

In our case,while calculating eigen vector of A^TA for $\lambda_2 = 32$ if we consider eigen vector to be $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ then we can't get back the A by computing $U\Sigma V^T$.

Thus we should first compute either U or V and then compute the other one by taking reference of computed one. So we will get U and V compatible to each other such that

$$U\Sigma V^\top = A$$

One of the approach of finding U and V is to find V first and then compute each vectors of U .
Since we know that,

$$AV = U\Sigma$$

V and U matrix have many eigen vectors.Thus, we can defined i^{th} vector of U i.e. $\mathbf{u_i}$ in terms of i^{th} eigen value σ_i and i^{th} vector of V i.e. $\mathbf{v_i}$

$$A\mathbf{v_i} = \mathbf{u_i}\sigma_i$$

$$\mathbf{u_i} = \frac{1}{\sigma_i}A\mathbf{v_i}$$

If we find U and V in such a way that they are compatible to each other then we will always be able to get back A.

Please refer to following lectures of Prof. Gilbert Strang to understand what went wrong and then correct your answer (if it was wrong):

- `https://www.youtube.com/watch?v=TX_vooSnhm8&t=1177s` (starts at 1177 seconds)
- `https://www.youtube.com/watch?v=HgC1l_6ySkc&feature=youtu.be&t=1731)` (starts at 1731 seconds)

15. (2 points) Prove that the matrices U and V that you get from the SVD of a matrix A contain the basis vectors for the four fundamental subspaces of A. (this is where the whole course comes together: fundamental subspaces, basis vectors, orthonormal vectors, eigenvectors, and our special symmetric matrices $AA^\top$, $A^\top A$!)

Solution:

Solution: Let's say we have a $m \times n$ matrix **A** whose SVD is $A = U_1 \Sigma_1 V_1^T$ where $U_1 = U_{1 \to r}$, an $m \times r$ matrix, $V_1 = V_{1 \to r}$, an $n \times r$ matrix.

For column space of A:
Let $b \in C(A)$
$b = Ax = U_1 \Sigma_r V_1^T x = U_1 x*$

Hence, $b \in C(U)_1$
Let's assume that $b \in C(U)_1$. Given $V_1^T V_1 = I$ and Σ_r is invertible, $U_1 = AV_1 \Sigma_r^{-1}$
Then,
$b = U_1 y = AV_1 \Sigma_r^{-1} y = Ax$

Thus $b \in C(A)$, Therefore $C(A) = C(U)_1$, U_1 is the orthonormal basis for C(A)

For row space of A:
Let $b \in C(A^T)$
$b = A^T y = V_1 \Sigma_r U_1^T x = V_1 y*$

Hence, $b \in C(V)_1$
Let's assume that $b \in C(V)_1$. Given $U_1^T U_1 = I$ and Σ_r is invertible, $V_1 = A^T U_1 \Sigma_r^{-1}$
Then,
$b = V_1 z = A^T U_1 \Sigma_r^{-1} z = A^T y$

Thus $b \in C(A^T)$, Therefore $C(A^T) = C(V)_1$, V_1 is the orthonormal basis for $C(A^T)$

For null space of A:
Let $A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \Sigma \begin{bmatrix} V_1 & V_2 \end{bmatrix}^T$

1. $U_1 = U_{1 \to r}$, $m \times r$ matrix
2. $U_2 = U_{r+1 \to m}$, $m \times m - r$ matrix
3. $V_1 = V_{1 \to r}$, $n \times r$ matrix
4. $V_2 = V_{r+1 \to n}$, $n \times n - r$ matrix

and Σ is defined as $\Sigma = \begin{bmatrix} \Sigma_{r \times r} & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{bmatrix}$

$\Sigma = \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} A \begin{bmatrix} V_1 & V_2 \end{bmatrix}$
$\Sigma = \begin{bmatrix} U_1^T AV_1 & U_1^T AV_2 \\ U_2^T AV_1 & U_2^T AV_2 \end{bmatrix}$
$\begin{bmatrix} U_1^T AV_2 \\ U_2^T AV_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$AV_2 \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$AV_2 = 0$

Since there are n-r columnsin V_2 which is equal to dim(N(A)) and $V_2 \in N(A)$, orthoinormal columns of V_2 for basis N(A)

For left null space of A:
From above proof we see that
$\Sigma = \begin{bmatrix} U_1^T AV_1 & U_1^T AV_2 \\ U_2^T AV_1 & U_2^T AV_2 \end{bmatrix}$
$\begin{bmatrix} U_2^T AV_1 \\ U_2^T AV_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$U_2^T \begin{bmatrix} V_1^T & V_2^T \end{bmatrix} = 0$
$U_2^T AV = 0$
$U_2^T A = 0$
Since there are m-r columns in U_2 which is equal to $dim(N(A^T))$ and $U_2 \in N(A^T)$ orthonormal columns of U_2 for basis $N(A^T)$

1. U_1 is the orthonormal basis for $C(A)$
2. U_2 is the orthonormal basis for $C(A^T)$
3. V_1 is the orthonormal basis for $N(A)$
4. V_2 is the orthonormal basis for $N(A^T)$