Honor code: I pledge on my honor that: I have completed all steps in the below homework on my own, I have not used any unauthorized materials while completing this homework, and I have not given anyone else access to my homework.

Varun Gumma

Name and Signature

1. (1 point) Have you read and understood the honor code?

> **Solution:** YES

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:** Consider the matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. Eigenvalues are $1, 2$ and -1 and corresponding eigenvectors are $\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^\top$, $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^\top$ and $\begin{bmatrix} 1 & -2 & 1 \end{bmatrix}^\top$ respectively. Here all three eigenvalues are distinct and their corresponding eigenvectors are independent.

(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:** Consider $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. The eigenvalues are $1, 2$ and 2 and corresponding eigenvectors are $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^\top$, $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^\top$ and $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\top$ respectively. Here the eigenvectors corresponding to eigenvalue 2 are $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^\top$ and $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\top$ which are independent.

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: By definition, the scalar λ is an eigenvalue of M if and only if it satisfies the characteristic equation of M, i.e. $det(M - \lambda I) = 0$.

λ is eigenvalue of A
$\iff det(A - \lambda I) = 0$
$\iff det((A - \lambda I)^T) = 0$ (as $det(A) = det(A^T)$)
$\iff det(A^T - \lambda I^T) = 0$
$\iff det(A^T - \lambda I) = 0$
$\iff \lambda$ is eigenvalue of A^T.

Hence, A and A^T have same set of eigenvalues and share the same characteristic equation.

(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 any triangular matrix. $\therefore A - \lambda I$ is just A with a λ term subtracted from all the diagonal elements. $\therefore A - \lambda I$ is also a triangular matrix and since determinant of a triangular matrix is the product of the diagonal terms, we have $det(A - \lambda I) = \prod_{i=1}^{n}(a_{ii} - \lambda)$. $\therefore$ for the eigenvalues of A, we have $det(A - \lambda I) = \prod_{i=1}^{n}(a_{ii} - \lambda) = 0$. Hence, if $\lambda = a_{ii}$ $(1 \leq i \leq n)$, the determinant becomes zero. $\therefore a_{ii}$ $(1 \leq i \leq n)$, or the diagonal elements of A are the eigenvalues of A. It also follows the condition that for a matrix, the product of eigenvalues is equal to its determinant as here $\prod_{i=1}^{n} \lambda_i = \prod_{i=1}^{n} a_{ii} = det(A)$.

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: Let the geometric multiplicity of λ_k be r. $\therefore$ we have r independent eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\ldots$ $\mathbf{v}_r$ associated with λ_k. Let $V = \begin{bmatrix}\mathbf{v}_1 & \ldots & \mathbf{v}_r\end{bmatrix}$ (size $k \times r$). Let's augument $s = K - r$ independent vectors to this set so that they form a basis for K dimensions. $\therefore$ now $V = \begin{bmatrix}\mathbf{v}_1 & \ldots & \mathbf{v}_r & \mathbf{w}_1 & \ldots & \mathbf{w}_s\end{bmatrix}$. Since, they form a basis, any K dimensional vector can be expressed in terms of them. $\therefore$ we can have $A\mathbf{w}_i = \sum_{j=1}^{r} c_{ji}\mathbf{v}_j + \sum_{j=1}^{s} d_{ji}\mathbf{w}_j$ or $A\begin{bmatrix}\mathbf{w}_1 & \ldots & \mathbf{w}_s\end{bmatrix} = \begin{bmatrix}\mathbf{v}_1 & \ldots & \mathbf{v}_r & \mathbf{w}_1 & \ldots & \mathbf{w}_s\end{bmatrix}\begin{bmatrix}C \\ D\end{bmatrix} = V\begin{bmatrix}C \\ D\end{bmatrix}$, where C and D are matrices of size $r \times s$ and $s \times s$ respectively.

$$
\begin{aligned}
&\therefore (A - \lambda_k I_K)V \\
&= \begin{bmatrix}(A - \lambda_k I_K)\mathbf{v}_1 & \ldots & (A - \lambda_k I_K)\mathbf{v}_r & (A - \lambda_k I_K)\mathbf{w}_1 & \ldots & (A - \lambda_k I_K)\mathbf{w}_s\end{bmatrix} \\
&= \begin{bmatrix}(\lambda - \lambda_k)\mathbf{v}_1 & \ldots & (\lambda - \lambda_k)\mathbf{v}_r & A\mathbf{w}_1 - \lambda\mathbf{w}_1 & \ldots & A\mathbf{w}_s - \lambda\mathbf{w}_s\end{bmatrix} \\
&= \begin{bmatrix}-\lambda_k\mathbf{v}_1 & \ldots & -\lambda_k\mathbf{v}_r & A\mathbf{w}_1 & \ldots & A\mathbf{w}_s\end{bmatrix} + \begin{bmatrix}\lambda\mathbf{v}_1 & \ldots & \lambda\mathbf{v}_r & -\lambda\mathbf{w}_1 & \ldots & -\lambda\mathbf{w}_s\end{bmatrix}
\end{aligned}
$$

$$= \begin{bmatrix} -\lambda_k \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_r \end{bmatrix} & V \begin{bmatrix} C \\ D \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \lambda \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_r \end{bmatrix} & -\lambda \begin{bmatrix} \mathbf{w}_1 & \dots & \mathbf{w}_s \end{bmatrix} \end{bmatrix}$$

$$= V(\begin{bmatrix} -\lambda_k I_r & C \\ \mathbf{0} & D \end{bmatrix} + \begin{bmatrix} \lambda I_r & \mathbf{0} \\ \mathbf{0} & -\lambda I_s \end{bmatrix})$$

$$= V \begin{bmatrix} (\lambda - \lambda_k) I_r & C \\ \mathbf{0} & D - \lambda I_s \end{bmatrix}$$

$$= V \begin{bmatrix} (\lambda - \lambda_k) I_r & C \\ \mathbf{0} & D - \lambda I_s \end{bmatrix}$$

$\therefore det((A - \lambda_k I_K)V) = det(A - \lambda_k I_K) \cdot det(V) = det(V \begin{bmatrix} (\lambda - \lambda_k) I_r & C \\ \mathbf{0} & D - \lambda I_s \end{bmatrix}) = det(V) \cdot det(\begin{bmatrix} (\lambda - \lambda_k) I_r & C \\ \mathbf{0} & D - \lambda I_s \end{bmatrix})$ or $det(A - \lambda_k I_r) = det(\begin{bmatrix} (\lambda - \lambda_k) I_r & C \\ \mathbf{0} & D - \lambda I_s \end{bmatrix})$. Here, $det(V)$ can be cancelled as V has all independent columns and is full rank and hence has non-zero determinant.

$\therefore$ the characteristic equation with λ_k is $det(A - \lambda_k I_r) = det((\lambda - \lambda_k) I_r (D - \lambda I_s)) = (\lambda - \lambda_k)^r \cdot det(D - \lambda I_s)$. As we can see, the first term has λ_k as a root r times and the second term may or may not have λ_k as a root. Hence, the whole product has λ_k as a root atleast r times. Hence, the algebraic multiplicity of λ_k is atleast r while the geometric multiplicity is exactly r (as per the assumption). $\therefore$ for any eigenvalue, algebraic multiplicity is greater than or equal to geometric multiplicity.

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

Solution: By definition, matrix M is positive definite if for all non-zero vectors $\mathbf{v}$ of appropriate dimension, $\mathbf{v}^T M \mathbf{v} > 0$.

Let $A + B = C$ and $\mathbf{v}$ be any non-zero vector of appropriate dimension, then $\mathbf{v}^T C \mathbf{v} = \mathbf{v}^T (A + B) \mathbf{v} = \mathbf{v}^T A \mathbf{v} + \mathbf{v}^T B \mathbf{v}$. Since A, B are positive definite $\mathbf{v}^T A \mathbf{v} > 0$ and $\mathbf{v}^T B \mathbf{v} > 0$. $\therefore \mathbf{v}^T C \mathbf{v} > 0$ as it is a sum of two strictly positive values. $\therefore C$ is also positive definite.

Eigenstory: Special Matrices

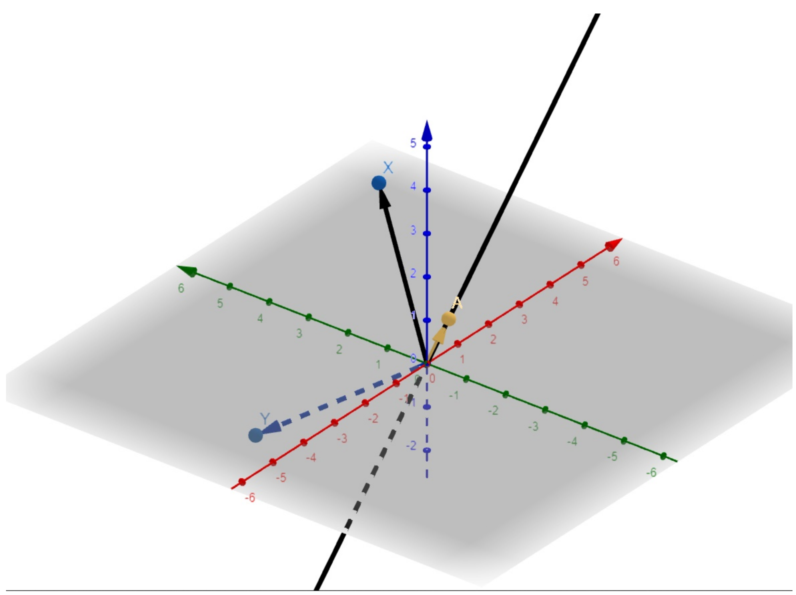
6. (2 points) Consider the matrix $R = I - 2\mathbf{u}\mathbf{u}^\top$ where $\mathbf{u}$ is a unit vector $\in \mathbb{R}^n$.
   (a) Show that R is symmetric and orthogonal. (How many independent vectors will R have?)

Solution: $R = I - 2\mathbf{u}\mathbf{u}^T \therefore R^T = (I - 2\mathbf{u}\mathbf{u}^T)^T = I^T - 2(\mathbf{u}\mathbf{u}^T)^T = I - 2\mathbf{u}\mathbf{u}^T = R$. As $R = R^T$, R is symmetric. $RR^T = R^TR = (I - 2\mathbf{u}\mathbf{u}^T)(I - 2\mathbf{u}\mathbf{u}^T) = I^2 - 2\mathbf{u}\mathbf{u}^T - 2\mathbf{u}\mathbf{u}^T + 4\mathbf{u}\mathbf{u}^T\mathbf{u}\mathbf{u}^T$. Since, $\mathbf{u}$ is a unit vector $\mathbf{u}^T\mathbf{u} = ||\mathbf{u}||_2^2 = 1^2 = 1$. $\therefore$ the above eqaution boils down to $I - 4\mathbf{u}\mathbf{u}^T + 4\mathbf{u}\mathbf{u}^T = I$. $\therefore$ as $RR^T = R^TR = I$, R is orthogonal. Since $RR^T = I$, $R^{-1} = R^T$. As R is invertible, R must be a full rank matrix with all its columns being independent. $\therefore$ R has n independent columns (dimension of R is same as dimension of $\mathbf{u}\mathbf{u}^\mathbf{T}$ which is $n \times n$).

(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: Here $R = I - 2\mathbf{u}\mathbf{u}^T = \begin{bmatrix}0 & 0 & -1\\0 & 1 & 0\\-1 & 0 & 0\end{bmatrix}$. Let's choose $\mathbf{x} = \begin{bmatrix}1\\2\\3\end{bmatrix}$.

$\therefore R\mathbf{x} = \mathbf{y} = \begin{bmatrix}-3\\2\\-1\end{bmatrix}$. R is rotation matrix and it moves a vector (with tail fixed at origin) with the norm remaining the same ($||\mathbf{x}|| = ||\mathbf{y}||$).



(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: Here $R = I - 2\mathbf{u}\mathbf{u}^T = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$. For eigenvalues, $det(R - \lambda I) =$ $\begin{vmatrix} -\lambda & 0 & -1 \\ 0 & 1-\lambda & 0 \\ -1 & 0 & -\lambda \end{vmatrix} = (\lambda+1)(1-\lambda)(\lambda-1) = 0$ or $\lambda = -1, 1, 1$.

1. *Eigenvector for 1*: Substituting $\lambda = 1$ in $R - \lambda I$, we have $\begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{bmatrix}$. Subtracting R_1 from R_3, $\begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. $\therefore$ we have two free variables x_2 and x_3 and the null space is $\begin{bmatrix} -x_3 & x_2 & x_3 \end{bmatrix}^\top = x_3 \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^\top + x_2 \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^\top$. $\therefore$ the eigenvectors are $\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^\top$ and $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^\top$

2. *Eigenvector for -1*: Substituting $\lambda = -1$ in $R - \lambda I$, we have $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}$. Adding R_1 to R_3 and dividing R_2 by its pivot, $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. $\therefore$ we have one free variable x_3 and the null space is $\begin{bmatrix} x_3 & 0 & x_3 \end{bmatrix}^\top = x_3 \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^\top$. $\therefore$ the eigenvector is $\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^\top$.

(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: Since R (for any $\mathbf{u}$) is orthonormal, $||R\mathbf{x}|| = (R\mathbf{x})^T R\mathbf{x} = \mathbf{x}^T R^T R\mathbf{x} = \mathbf{x}^T I\mathbf{x} = \mathbf{x}^T\mathbf{x} = ||\mathbf{x}||$. $\therefore$ if any vector is multiplied by R, the product has the same norm as the original vector and hence the matrix R just performs rotation. A vector $\mathbf{x}$ will be an eigenvector for a rotation matrix iff upon multiplication it points opposite to $\mathbf{x}$ or in the same direction of $\mathbf{x}$, i.e. gets rotated by π or 2π.

1. *Vector is rotated by π*: The new vector will be $-\mathbf{x}$ (points in the opposite direction as $\mathbf{x}$ with same norm as $\mathbf{x}$). $\therefore R\mathbf{x} = -1 \cdot \mathbf{x}$. Hence, one eigenvalue

is -1.

2. *Vector is rotated by* 2π: The new vector will be $\mathbf{x}$ (points in the same direction as $\mathbf{x}$ with same norm as $\mathbf{x}$). $\therefore R\mathbf{x} = 1\cdot\mathbf{x}$. Hence, other eigenvalue is 1.

Any other vector which gets rotated by any other angle upon multiplication will lose its orientation and cannot be an eigenvector. $\therefore$ the only possible eigenvalues are -1 and 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: False. Let $Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Q is orthogonal as $Q^TQ = I$. For eigenvalues of Q we have $\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0$ or $\lambda = \pm i$. Here, $\lambda^2 \neq 1$ for either eigenvalue.

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

Solution: Yes

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

Solution: True, Q is always diagonalizable (over complex numbers). Since, the eigenvectors and eigenvalues of an orthonormal matrix can be complex, Schur's theorem for orthonormal matrices is $Q = VTV^*$ where $V^* = \overline{V}^T$ and $VV^* = I$ and T is an upper triangular complex matrix with its diagonal elements as the eigenvalues of Q. As Q is orthonormal with real entries, $QQ^* = Q^*Q = I$ (here $Q^* = Q^T$). $\therefore VTV^*VT^*V^* = VT^*V^*VTV^* = I$ or $TT^* = T^*T = I$.

Let $T = \begin{bmatrix} t_{11} & t_{12} & t_{13} & \dots & t_{1n} \\ 0 & t_{22} & t_{23} & \dots & t_{2n} \\ 0 & 0 & t_{33} & \dots & t_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & t_{nn} \end{bmatrix}$, $T^* = \begin{bmatrix} \overline{t_{11}} & 0 & 0 & \dots & 0 \\ \overline{t_{12}} & \overline{t_{22}} & 0 & \dots & 0 \\ \overline{t_{13}} & \overline{t_{23}} & \overline{t_{33}} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{t_{1n}} & \overline{t_{2n}} & \overline{t_{3n}} & \dots & \overline{t_{nn}} \end{bmatrix}$

k^{th} diagonal element of $T^*T = \sum_{i=k}^{n} |t_{ki}|^2$ but it is also equal to k^{th} diagonal element of I which is 1. $\therefore \sum_{i=k}^{n} |t_{ki}|^2 = 1$. But from the property of Schur's decomposition, t_{kk} is an eigenvalue and hence $|t_{kk}|^2 = 1$. $\therefore \sum_{i=k}^{n} |t_{ki}|^2 = |t_{kk}|^2 + \sum_{i=k+1}^{n} |t_{ki}|^2 = 1 + \sum_{i=k+1}^{n} |t_{ki}|^2 = 1$ or $\sum_{i=k+1}^{n} |t_{ki}|^2 = 0$. Since sum of squares of some terms is 0, each term in that summation must be 0. $\therefore |t_{ki}|^2 = 0$ or $t_{ki} = 0$ (for $i > k$). As this holds for each throughout the matrix (i.e. for all diagonal terms), all upper diagonal terms of T are zero, or T is diagonal (from now referred as Λ) with diagonal elements being the eigenvalues of Q. $\therefore$ as $Q = V\Lambda V^*$ or $\Lambda = V^*QV$, orthonormal matrices are diagonalizable.

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: Let $\mathbf{u}$, $\mathbf{v}$ be $n \times 1$ vectors. As $\mathbf{u}\mathbf{v}^T$ (let it be denoted by A) is rank one (not full rank), it is singular and hence has a non-zero dimension for null space. $\therefore \exists \mathbf{b}$ such that $A\mathbf{b} = \mathbf{0} = 0 \cdot \mathbf{b}$. Hence, 0 is an eigenvalue of $\mathbf{u}\mathbf{v}^T$.

   Let the eigenvalue be λ with a corresponding eigenvector $\mathbf{x}$. $\therefore A\mathbf{x} = \lambda\mathbf{x}$ or $\mathbf{u}\mathbf{v}^T\mathbf{x} = \lambda\mathbf{x}$. As $\mathbf{v}^T\mathbf{x}$ is scalar (say k), $\mathbf{x} = \frac{k}{\lambda}\mathbf{u}$. $\mathbf{x}$ is an eigenvector of A and also a multiple of $\mathbf{u}$ and hence, even $\mathbf{u}$ is an eigenvector and has the same eigenvalue λ (if $\mathbf{x}$ is eigenvector of M, $M(t\mathbf{x}) = t(M\mathbf{x}) = t(k\mathbf{x}) = k(t\mathbf{x})$ or $t\mathbf{x}$ is also an eigenvector with eigenvalue k). $\therefore (\mathbf{u}\mathbf{v}^T)\mathbf{u} = \mathbf{u}(\mathbf{v}^T\mathbf{u}) = \lambda\mathbf{u}$. $\therefore$ another eigenvalue is $\mathbf{v}^T\mathbf{u}$.

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

   Solution: Number of 0 eigenvalues is same the dimension of null space of $\mathbf{u}\mathbf{v}^T$. If $\mathbf{u}$, $\mathbf{v}$ are $n \times 1$ vectors, $\mathbf{u}\mathbf{v}^T$ is $n \times n$ matrix and has rank 1 $\therefore$ dimension of null space is $n-1$ by rank-nullity theorem. $\therefore$ eigenvalue 0 repeats $n-1$ times.

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

   Solution: Eigenvectors of 0: Basis vectors of null space of $\mathbf{u}\mathbf{v}^T$.
   Eigenvector of $\mathbf{v}^T\mathbf{u}$: $(\mathbf{u}\mathbf{v}^T)\mathbf{x} = (\mathbf{v}^T\mathbf{u})\mathbf{x}$. $\therefore \mathbf{x} = \mathbf{u}$ satisfies, i.e. $(\mathbf{u}\mathbf{v}^T)\mathbf{u} = (\mathbf{v}^T\mathbf{u})\mathbf{u}$ (since $\mathbf{v}^T\mathbf{u}$ is a scalar). Hence, the eigenvector is $\mathbf{u}$.

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

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

Solution: Every Markov matrix has non-negative elements and sum of elements of each row is 1.

1. *1 is an eigenvalue*: Let A be a markov matrix and $\mathbf{o}$ denote the vector $\begin{bmatrix}1 & 1 & 1 & \dots & 1\end{bmatrix}^T$. $\therefore A\mathbf{o} = \begin{bmatrix}\sum_{j=1}^{n} a_{1j} & \sum_{j=1}^{n} a_{2j} & \dots & \sum_{j=1}^{n} a_{nj}\end{bmatrix}^T = \begin{bmatrix}1 & 1 & 1 & \dots & 1\end{bmatrix}^T$. Since $A\mathbf{o} = \mathbf{o}$, $\mathbf{o}$ is an eigenvector with eigenvalue 1. $\therefore$ 1 is one of the eigenvalue.

2. *Eigenvalues are atmost 1*: Let λ be some other eigenvalue of Markov matrix A with corresponding eigenvector $\mathbf{x}$. $\therefore A\mathbf{x} = \lambda\mathbf{x}$. Let $|v_k| = max(|v_1|, |v_2|, \dots |v_n|)$. $\therefore$ we have $a_{k1}v_1 + a_{k2}v_2 + \dots a_{kn}v_n = \lambda v_k$. Taking a norm on both sides:

   $$\begin{aligned}
   &|\lambda v_k| = |\lambda||v_k| \\
   &= |a_{k1}v_1 + a_{k2}v_2 + \dots + a_{kn}v_n| \\
   &\leq |a_{k1}v_1| + |a_{k2}v_2| + \dots + |a_{kn}v_n| \text{ (Cauchy-Schwarz Inequality)} \\
   &= a_{k1}|v_1| + a_{k2}|v_2| + \dots + a_{kn}|v_n| \text{ (since } a_{ij} \text{ are non-negative, } |a_{ij}| = a_{ij}) \\
   &\leq a_{k1}|v_k| + a_{k2}|v_k| + \dots + a_{kn}|v_k| \text{ (since } \forall i\ |v_i| \leq |v_k|) \\
   &= (a_{k1} + a_{k2} + \dots + a_{kn})|v_k| \\
   &= |v_k| \text{ (as } \textstyle\sum_{j=1}^{n} a_{kj} = 1)
   \end{aligned}$$

   $\therefore$ from the above steps we have, $|\lambda||v_k| \leq |v_k|$ and as we have a non-zero $\mathbf{v}$, $v_k > 0$ and hence $|\lambda| \leq 1$ or $-1 \leq \lambda \leq 1$. $\therefore$ the eigenvalues are atmost 1.

$\therefore$ from the above two cases, we observe that 1 is the dominant eigenvalue.

(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: With $A = \begin{bmatrix}a & b \\ c & d\end{bmatrix}$, the eigenvalues of A are roots of $det(A - \lambda I) = 0$. $det(A - \lambda I) = \begin{vmatrix}a - \lambda & b \\ c & d - \lambda\end{vmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + ad - bc = 0$. Discriminant of this quadratic is $(a+d)^2 - 4(ad - bc) = a^2 + d^2 + 2ad - 4ad + 4bc = a^2 + d^2 - 2ad + 4bc = (a - d)^2 + 4bc$. Since, $a + b = c + d$, $a - d = c - b$. $\therefore (a - d)^2 + 4bc = (c - b)^2 + 4bc = c^2 + b^2 + 2bc = (c + b)^2$. Hence, the one of root of this quadratic (which an eigenvalue) is $\frac{a+d+\sqrt{(c+b)^2}}{2} = \frac{a+d+b+c}{2} = \frac{2k}{2} = k$. If $a + b = c + d = 1$ (as in a Markov matrix), one of the eigenvalues will be $k = 1$.

(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: Let A be an $n \times n$ matrix such that for each row, all the row elements add up to a constant value, say k, i.e $\forall i, \sum_{j=0}^{n} a_{ij} = k$. The eigenvalues for this matrix can be obtained as solutions to $det(A - \lambda I) = 0$. This determinant will have $a_{ii} - \lambda$ aong the principle diagonal. As determinant doesn't alter with column-ops, we perform $C_1 \leftarrow \sum_{i=1}^{n} C_i$. Now, the first element of each row contains the sum of the elements of that row, which will be $\forall i$, $a_{i1} + a_{i2} + \cdots + a_{ii} - \lambda + \cdots + a_{in} = (k - \lambda)$. Since after this column-op, all the elements of the first column will be $(k - \lambda)$, we can pull it out of the determinant as a common factor and the first column will be all $1s$. Let this newly formed determinant have a value $g(\lambda)$ (since it still has λ terms in it). $\therefore$ as per out steps, $det(A - \lambda I) = (k - \lambda) \cdot g(\lambda)$. If we have $\lambda = k$, $det(A - kI) = (k - k) \cdot g(k) = 0 \cdot g(k) = 0$. Since the determinant of $A - \lambda I = 0$ for $\lambda = k$, i.e. the sum of the values of any row, it is an eigenvalue.

(d) What is the corresponding eigenvector?

Solution: For eigenvalue $k = \sum_{j=0}^{n} a_{ij}$ (for any i), $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \ldots & x_n \end{bmatrix}^\top$ be the corresponding eigenvector. $\therefore (A - kI)\mathbf{x} = \mathbf{0}$. The multiplication can be written as linear combinations of columns of $A - kI$ as $x_1 \begin{bmatrix} a_{11} - k & a_{21} & a_{31} \ldots a_{n1} \end{bmatrix}^\top$ $+ x_2 \begin{bmatrix} a_{12} & a_{22} - k & a_{32} & \ldots & a_{n2} \end{bmatrix}^\top + \cdots + x_n \begin{bmatrix} a_{1n} & a_{n2} & a_{n3} & \ldots & a_{nn} - k \end{bmatrix}^\top =$ $\begin{bmatrix} \sum_{j=1}^{n} x_j a_{1j} - kx_1 & \sum_{j=1}^{n} x_j a_{2j} - kx_2 & \ldots & \sum_{j=1}^{n} x_j a_{nj} - kx_n \end{bmatrix}^\top = \mathbf{0}$. As $\forall i$, $\sum_{j=1} x_j a_{ij} - kx_i = 0$ or $\sum_{j=1}^{n} (\frac{x_j}{x_i}) \cdot a_{ij} = k = \sum_{j=1}^{n} a_{ij}$. This is possible then $\forall i, j$ $\frac{x_j}{x_i} = 1$. $\therefore$ all elements of $\mathbf{x}$ are equal and $\mathbf{x} = \begin{bmatrix} x & x & x & \ldots & x \end{bmatrix}^\top$. With $x = 1$, we get the basis for this span which is also the primary eigenvector as $\begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \end{bmatrix}^\top$

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: True. Given A is a real matrix, its inverse will also be a real matrix and hence its characteristic equation will have real coefficients. If i is an eigenvalue for A, $A\mathbf{x} = i\mathbf{x}$. Multiplying both sides with A^{-1}, $A^{-1}A\mathbf{x} = iA^{-1}\mathbf{x}$ or $A^{-1}\mathbf{x} = -i\mathbf{x}$. $\therefore -i$ is an eigenvalue of A^{-1}. But as the coefficients of characteristic equation are real, the complex roots appear as conjugate pairs (only then upon addition or multiplication, their imaginary terms cancel out and we get real coefficients). $\therefore$ as $-i$ is an eigenvalue, i is also an eigenvalue.

(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. If the characteristic equation is of the form $(\lambda - a)(\lambda - b)(\lambda - c)(\lambda - d)(\lambda - e)$ (for 5^{th} degree polynomial), where $a, b, c, d, e \in \mathbb{C}$, we have $-(a+b+c+d+e)$ (minus of sum of eigenvalues) as the coefficient of λ^4. Here, the coefficient of $\lambda^4 = 0$. $\therefore$ sum of eigenvalues is 0 and not -7.

(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:** True. If $\begin{bmatrix} 4 & -3 & 2 \end{bmatrix}^\top = \mathbf{v}$, then $\begin{bmatrix} 16 & -12 & 8 \end{bmatrix}^\top = \mathbf{w} = 4\mathbf{v}$. Given $\mathbf{v}$ is an eigenvector for A, $A\mathbf{v} = \lambda\mathbf{v}$. $\therefore A\mathbf{w} = A(4\mathbf{v}) = 4A\mathbf{v} = 4\lambda\mathbf{v} = \lambda(4\mathbf{v}) = \lambda\mathbf{w}$. $\therefore$ as $A\mathbf{w} = \lambda\mathbf{w}$, $\mathbf{w} = \begin{bmatrix} 16 & -12 & 8 \end{bmatrix}^\top$ is also an eigenvector.

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

> **Solution:** True. Since by spectral theorem, any symmetric matrix can be diagonalized as $A = Q\Lambda Q^{-1}$ where Q matrix has eigenvectors of A as its columns, we observe that Q is invertible, i.e. it has full rank and all its columns are independent. $\therefore$ is A is $n \times n$, we have n independent eigenvectors. $\therefore$ for each diagonal element (λ_i) of Λ we have the corresponding eigenvector in V ($\mathbf{v}_i$) and even if two diagonal elements are the same (repeating eigenvalues), their eigenvectors are independent. For each eigenvalue, the number of times it appears on the diagonal of Λ, that many corresponding independent eigenvectors it has in V or algebraic multiplicity is equal to geometric multiplicity for all eigenvalues.

(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:** True. Given $\mathbf{x}$ is an eigenvector for A and B, $A\mathbf{x} = \lambda_a\mathbf{x}$, $B\mathbf{x} = \lambda_b\mathbf{x}$.
> $B(A\mathbf{x}) = B\lambda_a\mathbf{x} = \lambda_a B\mathbf{x} = \lambda_a\lambda_b\mathbf{x}$. If if assume $\lambda_a\lambda_b = \lambda$, we have $(BA)\mathbf{x} = \lambda\mathbf{x}$.
> $\therefore$ $\mathbf{x}$ is an eigenvector for BA with eigenvalue $\lambda_a\lambda_b$.
> $A(B\mathbf{x}) = A\lambda_b\mathbf{x} = \lambda_b A\mathbf{x} = \lambda_b\lambda_a\mathbf{x}$. If if assume $\lambda_a\lambda_b = \lambda$, we have $(AB)\mathbf{x} = \lambda\mathbf{x}$.
> $\therefore$ $\mathbf{x}$ is an eigenvector for AB with eigenvalue $\lambda_a\lambda_b$.
> $\therefore$ $\mathbf{x}$ is an eigenvector to both AB and BA with the same eigenvalue.

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

Solution: True. Let $A\mathbf{x} = \lambda_a \mathbf{x}$ and $B\mathbf{x} = \lambda_b \mathbf{x}$. Let $A + B = C$. $\therefore C\mathbf{x} = (A+B)\mathbf{x} = A\mathbf{x} + B\mathbf{x} = \lambda_a \mathbf{x} + \lambda_b \mathbf{x} = (\lambda_a + \lambda_b)\mathbf{x} = \lambda_c \mathbf{x}$. Hence, $\mathbf{x}$ is also an eigenvector to $A + B$ with its eigenvalue being the sum of eigenvalues with A and B.

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

Solution: True. Let λ be any non-zero eigen value of AA^T and let $\mathbf{x}$ be the corresponding eigenvector, i.e. $AA^T\mathbf{x} = \lambda\mathbf{x}$. Left mutiplying by A^T on both sides, we get $A^TAA^T\mathbf{x} = A^T\lambda\mathbf{x} = \lambda A^T\mathbf{x}$. If we take $A^T\mathbf{x} = \mathbf{y}$, $A^TA\mathbf{y} = \lambda\mathbf{y}$ or λ is the eigenvalue of A^TA and its corresponding eigenvector is $\mathbf{y}$. Hence, AA^T and A^TA share same non-zero eigenvalues.

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

Solution: False. Let $A = \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}$.

$A^TA = \begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix}$ and its eigenvectors are $\begin{bmatrix} 2+\sqrt{5} \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2-\sqrt{5} \\ 1 \end{bmatrix}$.

$AA^T = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$ and its eigenvectors are $\begin{bmatrix} -1-\sqrt{5} \\ 2 \end{bmatrix}$ and $\begin{bmatrix} -1+\sqrt{5} \\ 2 \end{bmatrix}$.

As observed from this example, eigenvectors of AA^T and A^TA need not be equal.

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: Let $\mathbf{u_1} = \frac{1}{5}\begin{bmatrix} 3 \\ 4 \end{bmatrix}, \mathbf{u_2} = \frac{1}{5}\begin{bmatrix} 4 \\ -3 \end{bmatrix}$ (basis-1) and $\mathbf{v_1} = \frac{1}{5}\begin{bmatrix} 3 \\ -4 \end{bmatrix}, \mathbf{v_2} = \frac{1}{5}\begin{bmatrix} -4 \\ -3 \end{bmatrix}$ (basis-2). To transform $\mathbf{x} = a\mathbf{u_1} + b\mathbf{u_2}$ into basis-2, we transform the basis vectors of basis-1, i.e. $(\mathbf{u_1}, \mathbf{u_2})$ into that of basis-2. Hence, we express $\mathbf{u_1} = \alpha\mathbf{v_1} + \beta\mathbf{v_2}$ and $\mathbf{u_2} = \gamma\mathbf{v_1} + \delta\mathbf{v_2}$ or $\begin{bmatrix} \mathbf{u_1} & \mathbf{u_2} \end{bmatrix} = \begin{bmatrix} \mathbf{v_1} & \mathbf{v_2} \end{bmatrix}\begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}$.

$$\therefore \frac{1}{5}\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}\begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}$$

$$\therefore \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix} = \begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}^{-1}\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} = \frac{1}{25}\begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} = \begin{bmatrix} -\frac{7}{25} & \frac{24}{25} \\ -\frac{24}{25} & -\frac{7}{25} \end{bmatrix}.$$

$\therefore \mathbf{x} = a\mathbf{u_1} + b\mathbf{u_2}$
$= a(\alpha\mathbf{v_1} + \beta\mathbf{v_2}) + b(\gamma\mathbf{v_1} + \delta\mathbf{v_2})$
$= (a\alpha + b\gamma)\mathbf{v_1} + (a\beta + b\delta)\mathbf{v_2}$
$= (\frac{-7a+24b}{25})\mathbf{v_1} + (\frac{-24a-7b}{25})\mathbf{v_2}$
$= (\frac{-7a+24b}{25}) \cdot \frac{1}{5}\begin{bmatrix} 3 \\ -4 \end{bmatrix} + (\frac{-24a-7b}{25}) \cdot \frac{1}{5}\begin{bmatrix} -4 \\ -3 \end{bmatrix}$

$\therefore$ in terms if basis-2 vectors it will be $\frac{1}{25}\begin{bmatrix} -7a+24b \\ -24a-7b \end{bmatrix}$

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: Let the matrix for T be M. Now, M contains the basis of the transformed space. Let us start with vectors $\mathbf{u}$, $\mathbf{v}$ in the standard basis and let $T(\mathbf{u}) = \hat{\mathbf{u}}$ and $T(\mathbf{v}) = \hat{\mathbf{v}}$. Hence, $\mathbf{u} = M\hat{\mathbf{u}}$, $\mathbf{v} = M\hat{\mathbf{v}}$

1. *If case*: Assuming basis represented by T is orthonormal, i.e. the columns of M are orthonormal (as columns of M denote where the std basis vectors land after transformation), $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = (M\hat{\mathbf{u}})^T M\hat{\mathbf{v}} = \hat{\mathbf{u}}^T M^T M\hat{\mathbf{v}} = \hat{\mathbf{u}}^T I\hat{\mathbf{v}} = \hat{\mathbf{u}}^T\hat{\mathbf{v}} = (T(\mathbf{u}))^T T(\mathbf{v}) = T(\mathbf{u}) \cdot T(\mathbf{v})$
2. *Only If case*: Assuming $\mathbf{u} \cdot \mathbf{v} = T(\mathbf{u}) \cdot T(\mathbf{v})$, we have $\mathbf{u} \cdot \mathbf{v} = M\hat{\mathbf{u}} \cdot M\hat{\mathbf{u}} = (M\hat{\mathbf{u}})^T M\hat{\mathbf{v}} = \hat{\mathbf{u}}^T M^T M\hat{\mathbf{v}}$ and $T(\mathbf{u}) \cdot T(\mathbf{v}) = (T(\mathbf{u}))^T T(\mathbf{v}) = \hat{\mathbf{u}}^T\hat{\mathbf{v}}$. $\therefore \hat{\mathbf{u}}^T M^T M\hat{\mathbf{v}} = \hat{\mathbf{u}}^T\hat{\mathbf{v}}$ and this is only possible when $M^T M = I$ or M is orthonormal.

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:

Dimensions of U, Σ and V are $m \times m$, $m \times n$ and $n \times n$ respectively. We can perform PCA using SVD as follows:

1. The sample covariance matrix S assuming A is zero-centered is $\frac{1}{m}A^TA = \frac{1}{m}(U\Sigma V^T)^T U\Sigma V^T = \frac{1}{m}V\Sigma^T U^T U\Sigma V^T = V(\frac{\Sigma^T\Sigma}{m})V^T$.

2. Since S is a symmetric matrix, $S = V(\frac{\Sigma^T\Sigma}{m})V^T$ is its eigenvalue decomposition where V is the matrix of eigenvectors and $\frac{\Sigma^T\Sigma}{m}$ is the diagonal eigenvalues matrix.

3. The $\mathbf{v}$'s form the principle directions of A and the eigenvalues of S are $\lambda_i = \frac{\sigma_i^2}{m}$ where σ_i are the singular values of A.

4. The principle components are $AV = U\Sigma V^T V = U\Sigma$. Hence the top-$k$ principle components (which have the maximum information) are given by $U_k\Sigma_k$ where Σ_k has the highest k eigenvalues along the diagonal and V_k is the set of the eigenvectors corresponding to aforementioned eigenvalues.

5. Hence, the best k rank approximation of A is obtained by $\hat{A}_k = U_k\Sigma_k V_k^T = \sum_{i=1}^{k}\sigma_i \mathbf{u}_i\mathbf{v}_i^T$ (Eckart-Young theorem).

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: With $A = U\Sigma V^T$, $A^TA = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T\Sigma V^T$ ($U^TU = I$ as U is a orthonormal matrix). This is the eigenvalue decomposition of A^TA, and V is eigenvector matrix of A^TA and $\Sigma^T\Sigma$ is the diagonal matrix with eigenvalues of A^TA. Here, diagonal of $\Sigma^T\Sigma$ has squares of the diagonal terms of Σ.

$A^TA = \begin{bmatrix} 5 & -4 \\ 5 & 4 \end{bmatrix}\begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix} = \begin{bmatrix} 41 & 9 \\ 9 & 41 \end{bmatrix}$. For eigenvalues, $det(A^TA - \lambda I) = \begin{vmatrix} 41-\lambda & 9 \\ 9 & 41-\lambda \end{vmatrix} = (41-\lambda)^2 - 81 = \lambda^2 - 82\lambda + 1600 = (\lambda-50)(\lambda-32) = 0$. $\therefore$ the eigenvalues are 50, 32.

1. *Eigenvector of* 50: Substituting 50 in $A - \lambda I$ we have $\begin{bmatrix} -9 & 9 \\ 9 & -9 \end{bmatrix}$. Adding R_1 to R_2 and dividing R_1 by its pivot, $\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$. $\therefore x_2$ is a free variable and the null space is of the form $x_2\begin{bmatrix} 1 & 1 \end{bmatrix}^\top$. $\therefore$ the basis and the eigenvector is $\begin{bmatrix} 1 & 1 \end{bmatrix}^\top$.

2. *Eigenvector of* 32: Substituting 32 in $A - \lambda I$ we have $\begin{bmatrix} 9 & 9 \\ 9 & 9 \end{bmatrix}$. Subtracting

R_1 from R_2 and dividing R_1 by its pivot, $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. $\therefore x_2$ is a free variable and the null space is of the form $x_2 \begin{bmatrix} -1 & 1 \end{bmatrix}^\top$. $\therefore$ the basis and the eigenvector is $\begin{bmatrix} -1 & 1 \end{bmatrix}^\top$.

$\therefore \Sigma = \begin{bmatrix} \sqrt{50} & 0 \\ 0 & \sqrt{32} \end{bmatrix}$ and $V = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ (eigenvectors with unit norm since $VV^T = V^TV = I$)

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

Solution: With $A = U\Sigma V^T$, $AA^T = U\Sigma V^T(U\Sigma V^T)^T = U\Sigma V^T V \Sigma^T U^T = U\Sigma\Sigma^T U^T$ ($V^TV = I$ as V is a orthonormal matrix). This is the eigenvalue decomposition of AA^T, and U is eigenvector matrix of AA^T and $\Sigma\Sigma^T$ is the diagonal matrix with eigenvalues of AA^T. Here, diagonal of $\Sigma\Sigma^T$ has squares of the diagonal terms of Σ.

$AA^T = \begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix} \begin{bmatrix} 5 & -4 \\ 5 & 4 \end{bmatrix} = \begin{bmatrix} 50 & 0 \\ 0 & 32 \end{bmatrix}$. For eigenvalues, $det(AA^T - \lambda I) = \begin{vmatrix} 50-\lambda & 0 \\ 0 & 32-\lambda \end{vmatrix} = (50-\lambda)(32-\lambda) = 0$. $\therefore$ the eigenvalues are 50, 32.

1. *Eigenvector of* 50: Substituting 50 in $A - \lambda I$ we have $\begin{bmatrix} 0 & 0 \\ 0 & -18 \end{bmatrix}$. $\therefore x_1$ is a free variable and the null space is of the form $x_1 \begin{bmatrix} 1 & 0 \end{bmatrix}^\top$. $\therefore$ the basis and the eigenvector is $\begin{bmatrix} 1 & 0 \end{bmatrix}^\top$.

2. *Eigenvector of* 32: Substituting 32 in $A - \lambda I$ we have $\begin{bmatrix} 18 & 0 \\ 0 & 0 \end{bmatrix}$. $\therefore x_2$ is a free variable and the null space is of the form $x_2 \begin{bmatrix} 0 & 1 \end{bmatrix}^\top$. $\therefore$ the basis and the eigenvector is $\begin{bmatrix} 0 & 1 \end{bmatrix}^\top$.

$\therefore \Sigma = \begin{bmatrix} \sqrt{50} & 0 \\ 0 & \sqrt{32} \end{bmatrix}$ and $U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (eigenvectors with unit norm since $UU^T = U^TU = I$)

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

Solution: $U\Sigma V^T = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\sqrt{50} & 0\\ 0 & \sqrt{32}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}$
$= \begin{bmatrix}5\sqrt{2} & 0\\ 0 & 4\sqrt{2}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix} = \begin{bmatrix}5 & 5\\ -4 & 4\end{bmatrix} = A.$

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: Given A (size $m \times n$) has a SVD as $A = U\Sigma V^T$, let's assume A is rank-r i.e. r independent columns. Here, U is $m \times m$ and orthonormal ($U^TU = UU^T = I$), V is $n \times n$ and orthonormal ($V^TV = VV^T = I$), and Σ is $m \times n$ and diagonal ($\Sigma^T = \Sigma$). Σ is of the form $\begin{bmatrix}\sigma_{r\times r} & \mathbf{0}_{r\times n-r}\\ \mathbf{0}_{m-r\times r} & \mathbf{0}_{m-r\times n-r}\end{bmatrix}$

Let's analyse $A\mathbf{v}_j$ ($\mathbf{v}_j \in V$). $\therefore A\mathbf{v}_j = U\Sigma V^T\mathbf{v}_j = U\Sigma\mathbf{e}_j$ (where $\mathbf{e}_j$ is the j^{th} std basis). Since multiplying by $\mathbf{e}_j$ selects the j^{th} column of a matrix, we have $U\Sigma\mathbf{e}_j = U(\sigma_j\mathbf{e}_j) = \sigma_j U\mathbf{e}_j = \sigma_j\mathbf{u}_j$. $\therefore A\mathbf{v}_j = \sigma_j\mathbf{u}_j$. Now two cases arise:

1. *if* $j \leq r$: This implies $\sigma_j \neq 0$ or $A(\frac{\mathbf{v}_j}{\sigma_j}) = \mathbf{u_j}$. Since, $A\mathbf{x} = \mathbf{u}_j(\mathbf{x} = \frac{\mathbf{v}_j}{\sigma_j})$, $\mathbf{u}_j \in \mathcal{C}(A)$. $\therefore$ as the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots \mathbf{u}_r$ are all orthonormal and $\in \mathcal{C}(A)$, they form a basis for $\mathcal{C}(A)$, i.e. basis for $\mathcal{C}(A) = \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_r\end{bmatrix}$
2. *if* $j > r$: This implies $\sigma_j = 0$ and hence $A\mathbf{v_j} = 0 \cdot \mathbf{u}_j = 0$. Since $A\mathbf{v}_j = 0, \mathbf{v}_j \in \mathcal{N}(A)$. $\therefore$ as the vectors $\mathbf{v}_{r+1}, \mathbf{v}_{r+2} \dots \mathbf{v}_n$ are all orthonormal and $\in \mathcal{N}(A)$, they form a basis for $\mathcal{N}(A)$, i.e. the basis for $\mathcal{N}(A) = \begin{bmatrix}\mathbf{v}_{r+1} & \mathbf{v}_{r+2} & \dots & \mathbf{v}_n\end{bmatrix}$.

Let's analyse $A^T\mathbf{u}_j$ ($\mathbf{u}_j \in U$). $\therefore A^T\mathbf{u}_j = (U\Sigma V^T)^T\mathbf{u}_j = V\Sigma^T U^T\mathbf{u}_j = V\Sigma U^T\mathbf{u}_j = V\Sigma\mathbf{e}_j$ (where $\mathbf{e}_j$ is the j^{th} std basis). Since multiplying by $\mathbf{e}_j$ selects the j^{th} column of a matrix, we have $V\Sigma\mathbf{e}_j = V(\sigma_j\mathbf{e}_j) = \sigma_j V\mathbf{e}_j = \sigma_j\mathbf{v}_j$. $\therefore A\mathbf{u}_j = \sigma_j\mathbf{v}_j$. Now two cases arise:

1. *if* $j \leq r$: This implies $\sigma_j \neq 0$ or $A(\frac{\mathbf{u}_j}{\sigma_j}) = \mathbf{v}_j$. Since $A^T\mathbf{x} = \mathbf{v}_j$ $(\mathbf{x} = \frac{\mathbf{u}_j}{\sigma_j})$, $\mathbf{v}_j \in C(A^T)$. $\therefore$ as the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_r$ are all orthonormal and $\in \mathcal{C}(A^T)$, they form a basis for $\mathcal{C}(A^T)$ or $\mathcal{R}(A)$, i.e. basis for $\mathcal{R}(A) = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_r\end{bmatrix}$.
2. *if* $j > r$: This implies $\sigma_j = 0$ and hence $A\mathbf{u_j} = 0 \cdot \mathbf{v}_j = 0$. Since $A\mathbf{u}_j = 0, \mathbf{u}_j \in \mathcal{N}(A^T)$. $\therefore$ as the vectors $\mathbf{u}_{r+1}, \mathbf{u}_{r+2} \dots \mathbf{u}_n$ are all orthonormal and $\in \mathcal{N}(A^T)$, they form a basis for $\mathcal{N}(A^T)$, i.e. the basis for $\mathcal{N}(A^T) = \begin{bmatrix}\mathbf{u}_{r+1} & \mathbf{u}_{r+2} & \dots & \mathbf{u}_n\end{bmatrix}$.

...And that concludes the story of *How I Met Your Eigenvectors :-)* (Hope you enjoyed it!)

EXTRA QUESTIONS FOR PRACTICE. These questions will not be evaluated for grades, but we encourage you to solve them to gain better understanding of the concepts.

16. (0 points) Prove that for any square matrix A the eigenvectors corresponding to distinct eigenvalues are always independent.

17. (0 points) Prove the following.
    (a) The sum of the eigenvalues of a matrix is equal to its trace.
    (b) The product of the eigenvalues of a matrix is equal to its determinant.

18. (0 points) What is the relationship between the rank of a matrix and the number of non-zero eigenvalues? Explain your answer.

> **Solution:** I think the answer to this question is "The rank of a matrix is equal to the number of non-zero eigenvalues if $\cdots$"

19. (0 points) If A is a square symmetric matrix then prove that the number of positive pivots it has is the same as the number of positive eigenvalues it has.

20. (0 points) For each of the statements below state True or False with reason.
    (a) 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.
    (b) If $\mathbf{x}$ is and eigenvector of A and B then it is also an eigenvector of $A + B$
    (c) The non-zero eigenvalues of $AA^\top$ and $A^\top A$ are equal.

21. (0 points) How are PCA and SVD related? (no vague answers please, think and answer very precisely with mathematical reasoning)

22. (0 points) Fun with Objects.
    (a) In this activity, you need to find four different rank one objects and paste their photos. For e.g. the flag of Russia is a rank one flag. You can use an object of the same type only once, for e.g. you cannot use flags twice. Also avoid matrices and flag of Russia as answer.
    (b) What is the rank of a hypothetical 4 x 8 chess board?

23. (0 points) Consider the LFW dataset (Labeled Faces in the Wild).

(a) Perform PCA using this dataset and plot the first 25 eigenfaces (in a 5×5 grid)

Solution: Here is something to get you started.

```
import matplotlib.pyplot as plt
from sklearn.datasets import fetch_lfw_people
from sklearn.decomposition import PCA



# Load data
lfw_dataset = fetch_lfw_people(min_faces_per_person=100)

_, h, w = lfw_dataset.images.shape
X = lfw_dataset.data

# Compute a PCA
n_components = 100
pca = PCA(n_components=n_components, whiten=True).fit(X)

Beyond this you are on your own. Good Luck!
```

(b) Take your close-up photograph (face only) and reconstruct it using the first 25 eigenfaces :-). If due to privacy concerns, you do not want to to use your own photo then feel free to use a publicly available close-up photo (face only) of your favorite celebrity.