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.

N Kausik
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:

Since we can construct A = $S\lambda S^{-1}$, where S are its eigenvectors and λ is its eigenvalues,

If we take, $\lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ and $S = I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

eigenvalues of A are 1, 2, 3 and their corresponding eigenvectors are $(\begin{bmatrix} 1 & 0 & 0 \end{bmatrix})^\top$, $(\begin{bmatrix} 0 & 1 & 0 \end{bmatrix})^\top$ and $(\begin{bmatrix} 0 & 0 & 1 \end{bmatrix})^\top$.

Clearly, any two eigenvectors corresponding to distinct eigenvalues are independent.

$A = I \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} (I)^{-1} => \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

(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:

Since we can construct A = $S\lambda S^{-1}$, where S are its eigenvectors and λ is its eigenvalues,

If we take, $\lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Hence, $S^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$.

Eigenvalues of A have repeating value 1 and its corresponding eigenvectors are $(\begin{bmatrix} 1 & 0 & 0 \end{bmatrix})^\top$ and $(\begin{bmatrix} 0 & 0 & 2 \end{bmatrix})^\top$.

Clearly, here, the two eigenvectors corresponding to non-distinct eigenvalues are independent.

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

Hence A is not identity matrix and is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

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:

If λ is a eigenvalue of A,

$det(A - \lambda I) = 0$, and so, $(A - \lambda I)$ is not invertible. Since we know that if A is non-invertible, then $A^\top$ is non-invertible, $(A - \lambda I)^\top$ is also non-invertible.

=> $(A^\top - (\lambda I)^\top) = (A^\top - \lambda I)$ as $I^\top = I$.

=> $(A^\top - \lambda I)$ is non-invertible. Hence, λ is a eigenvalue of $A^\top$.

Therefore, if λ is a eigenvalue of A, it is a eigenvalue of $A^\top$.

If λ is a eigenvalue of $A^\top$,

$det(A^\top - \lambda I) = 0$, and so, $(A^\top - \lambda I)$ is not invertible. Since we know that if A is non-invertible, then $A^\top$ is non-invertible, $((A^\top)^\top - \lambda I)^\top$ is also non-invertible.

=> $(A - (\lambda I)^\top) = (A - \lambda I)$ as $I^\top = I$.

=> $(A - \lambda I)$ is non-invertible. Hence, λ is a eigenvalue of A.

Therefore, if λ is a eigenvalue of $A^\top$, it is a eigenvalue of A.

Therefore, 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:

Since for a triangular matrix, all elements are 0 except along its diagonal, $det(A)$ for a triangular matrix A is simply the product of its diagonal elements.
Hence, if A has eigenvalues $\lambda_1, ...\lambda_K$ (may have repeating) and if A has diagonal elements $a_1, ...a_K$, then, $det(\begin{bmatrix} a_1 - \lambda & 0 & ... & 0 \\ 0 & a_2 - \lambda & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & a_K - \lambda \end{bmatrix}) = 0$
$=> \prod_{i=1}^{K}(a_i - \lambda) = 0$
This is possible only when λ is equal to one of the diagonal elements. Hence the possible eigenvalues are simply the diagonal elements.

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 g. Therefore, we have g linearly independent vectors, $x_1, ...x_g$ such that $Ax_i = \lambda_k x_i$ where $i \in \{1, 2, ...g\}$.

Now, we add K-g more vectors to the g vectors such that the resulting K vectors are linearly independent. If we denote the extra vectors by $y_1, ...y_{K-g}$, the final vectors are used to form a matrix S such that it is $\begin{bmatrix} x_1 & x_2 & ... & x_g & y_1 & ... & y_{K-g} \end{bmatrix}$.

Hence, AS = $\begin{bmatrix} Ax_1 & Ax_2 & ... & Ax_g & Ay_1 & ... & Ay_{K-g} \end{bmatrix}$

$= \begin{bmatrix} \lambda x_1 & \lambda x_2 & ... & \lambda x_g & Ay_1 & ... & Ay_{K-g} \end{bmatrix}$.

Since we know that vectors of S form basis of whole R^K space (K independent vectors), we can write Ay_i as linear combination of columns of S.

Hence we can rewrite for each Ay_i to form it in matrix form as,

$A \begin{bmatrix} y_1 & y_2 & ... & y_{K-g} \end{bmatrix} = S \begin{bmatrix} C \\ D \end{bmatrix}$, where C is a (g)x(K-g) matrix and D is a (K-g)x(K-g) matrix and they contain the weights for the linear combination of S columns to produce each Ay_i.

Now for some scalar p,

$(A - pI)S = \begin{bmatrix} (\lambda - p)x_1 & ... & (\lambda - p)x_g & (Ay_1 - py_1) & ... & (Ay_{K-g} - py_{K-g}) \end{bmatrix}$

$= \begin{bmatrix} x_1 & ... & x_g & y_1 & ... & y_{K-g} \end{bmatrix} \begin{bmatrix} (\lambda - p)I_{gxg} & C \\ 0_{(K-g)xg} & D - pI_{(K-g)x(K-g)} \end{bmatrix}$

Now, $det((A - pI)S) = det(A - pI)det(S)$, as S has all columns independent, its determinant cant be 0. Also, $det(A - pI) = (\lambda - p)^g det(D - pI_{(K-g)x(K-g)})$.

Hence, $det(A - pI) = (\lambda - p)^g f(p)$, where $f(p)$ is a (K-g) degree polynomial in p. From this we can see that the minimum algebraic multiplicity of λ_k is g due to the degree of λ_k root for $det(A - pI) = 0$ is atleast g. It is also possible that λ_k is a root of $f(p)$ also which results in a higher algebraic multiplicity.

Hence Geometric Multiplicity $\leq$ Algebraic Multiplicity.

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

Solution:

If a matrix M is positive definite, then its eigenvalues are > 0. If we consider a non-zero eigenvector x, $Ax = \lambda x$. Hence, $x^\top Mx = x^\top \lambda x = \lambda x^\top x$. Since $\lambda > 0$ and $x^\top x > 0$, $x^\top Mx > 0$ if M is positive definite.

Since A and B are positive definite,

for a non-zero eigenvector x, $x^\top Ax > 0$ and $x^\top Bx > 0$.

Hence, for A + B, $x^\top(A+B)x = x^\top(Ax+Bx) = (x^\top Ax)+(x^\top Bx)$. Since both the terms are > 0, their sum is also > 0.

Hence, $x^\top(A+B)x > 0$ and so, A + B 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:

$R = I - 2uu^\top$, Therefore, $R^\top = (I-2uu^\top)^\top = I^\top - 2(uu^\top)^\top = I - 2(u^\top)^\top u^\top = I - 2uu^\top = R$. Since $R = R^\top$, R is a symmetric matrix.

$R^\top R = RR = (I - 2uu^\top)(I - 2uu^\top) = II - 2uu^\top I - I2uu^\top + 4uu^\top uu^\top$

Since u is a unit vector, $u^\top u = 1$, => $I - 4uu^\top + 4uu^\top = I$.

Since $R^\top R = I$, R is a orthogonal matrix.

Hence it is shown that R is a symmetric orthogonal matrix.

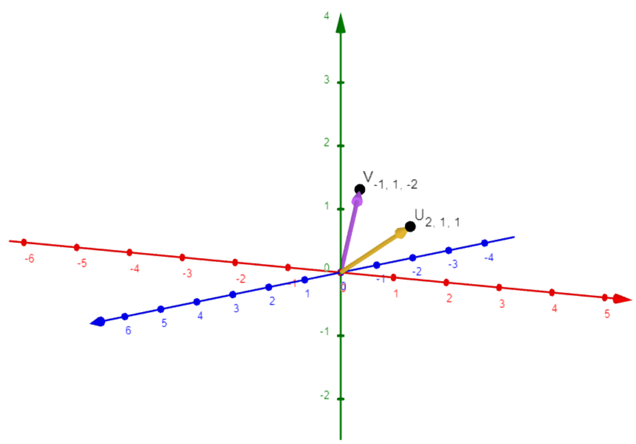
(b) Let $\mathbf{u} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix}$. Draw the line passing through this vector in geogebra (or any tool of your choice). Now take any vector in $\mathbf{R}^3$ and multiply it with the matrix R (i.e., the matrix R as defined above with $\mathbf{u} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix}$).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

Since $u = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $R = I - 2uu^{\top} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 2\frac{1}{2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$

If we consider a vector in R^3, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $Rx = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_3 \\ x_2 \\ -x_1 \end{bmatrix}$

Hence for example, if we take $x = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $Rx = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$.



From observation, the matrix swaps the 1st and 3rd components and negates them.

Hence this matrix is called a Reversal Matrix.

(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:

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

$=> det(\begin{bmatrix} -\lambda & 0 & -1 \\ 0 & 1-\lambda & 0 \\ -1 & 0 & -\lambda \end{bmatrix}) = 0 => -\lambda((1-\lambda)(-\lambda)) - (1-\lambda) = 0$

$=> \lambda^2 - \lambda^3 - 1 + \lambda = 0 => \lambda^3 - \lambda^2 - \lambda + 1 = 0$

Hence the eigenvalues for R are 1, -1 and 1. For eigenvalue = 1,

$(R - I)x = 0 => \begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{bmatrix} x = 0$, we get eigenvectors as all linear combinations of $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

For eigenvalue = -1,

$(R - I)x = 0 => \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix} x = 0$, we get eigenvector as $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

Hence eigenvalues are 1, -1 and eigenvectors are $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

(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:

Since $R = I - 2uu^\top$,

Let us consider, $Ru = (I - 2uu^\top)u => Ru = u - 2uu^\top u$, as u is a unit vector, $u^\top u = 1$, $=> Ru = u - 2u => Ru = -1(u)$.

Hence, -1 is always a eigenvalue of R.

Similarly, let us consider a vector v such that it is orthogonal to u. Therefore, $u^\top v = 0$.

Now, $Rv = (I - 2uu^\top)v => Rv = v - 2u(u^\top v) => Rv = v - 0 => Rv = v$.

Hence, 1 is always a eigenvalue of R.

Hence for any such matrix R, 1 and -1 will always be eigenvalues.

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:

   Since $Qx = \lambda x$, squaring both sides, $||Qx||^2 = ||\lambda x||^2 \Rightarrow (Qx)^\top (Qx) = \lambda^2 ||x||^2$
   $\Rightarrow x^\top Q^\top Q x = \lambda^2 ||x||^2 \Rightarrow x^\top x = \lambda^2 ||x||^2$ (as Q is orthogonal, $Q^\top Q = I$).
   $\Rightarrow ||x||^2 = \lambda^2 ||x||^2 \Rightarrow \lambda^2 = 1$.
   Hence shown.

   (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:

   Not all orthogonal matrices are diagonalisable.
   Example, if we take $Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, here, $QQ^\top = Q^\top Q = I$ and so Q is a orthogonal matrix with real entries.
   Now, $det(Q - \lambda I) = 0 \Rightarrow det(\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix}) = 0 \Rightarrow \lambda^2 + 1 = 0 \Rightarrow \lambda^2 = -1$.
   Hence there are no real eigenvalues and hence given orthogonal matrix is not diagonalisable.
   Hence for Q to be diagonalisable, it must be a symmetric matrix. Then by spectral theorem it will be diagonalisable. I.e. $Q^2 = I$ must be satisfied to be diagonalisable.

8. ($1\frac{1}{2}$ 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:

For any rank 1 matrix A, we can write all columns in the matrix as a multiple of the first column. Since it has dependent columns, $det(A) = 0 => det(A - 0I) = 0$. Hence, 0 is a eigenvalue.

Also, since $A = uv^\top$, $Au = (uv^\top)u = u(v^\top u)$. Since $v^\top u$ is a scalar, $=> Au = (v^\top u)u = \lambda u$. Hence, $v^\top u$ is another eigenvalue.

Therefore, any rank 1 matrix has eigenvalues 0 and $v^\top u$.

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

Solution:

Since $v^\top u$ is eigenvalue corresponding to the eigenvector u, all other eigenvectors will have eigenvalue 0 since these are the only possible eigenvalues. Hence, if the matrix A is a nxn matrix, then eigenvalue 0 repeats (n-1) times.

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

Solution:

From the derivation, eigenvector corresponding to eigenvalue $v^\top u$ is u.

For eigenvalue 0, the eigenvectors are, $Ax = 0x => Ax = 0$. I.e. all vectors formed by linear combinations of basis vectors of null space of A.

Since $A = uv^\top$, A has rows as multiples of $v^\top$. Hence its row reduced echelon form will have all rows other than the first row as 0. Hence, eigenvectors for eigenvalue 0 will be any vector x satisfying $v^\top x = 0$.

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

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

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

Markov matrix is any matrix with non-negative values and having the sum along every column as 1. For 1 to be a eigenvalue for a markov matrix M, $det(M-I)$ must be 0.

$$M - I = \begin{bmatrix} m_{11}-1 & m_{12} & ... & m_{1n} \\ m_{21} & m_{22}-1 & ... & m_{2n} \\ ... & ... & ... & ... \\ m_{n1} & m_{n2} & ... & m_{nn}-1 \end{bmatrix}$$

Since we know from properties of determinants that adding a row to another row does not change the value of determinant, if we add all the rows with the first row, we get,

$$det(M-I) = det(\begin{bmatrix} (\sum_{i=1}^{n} m_{i1}) - 1 & (\sum_{i=1}^{n} m_{i2}) - 1 & ... & (\sum_{i=1}^{n} m_{in}) - 1 \\ m_{21} & m_{22}-1 & ... & m_{2n} \\ ... & ... & ... & ... \\ m_{n1} & m_{n2} & ... & m_{nn}-1 \end{bmatrix})$$

Since we know that sum of elements of any column is 1, => $det(M-I) = det(\begin{bmatrix} 1-1 & 1-1 & ... & 1-1 \\ m_{21} & m_{22}-1 & ... & m_{2n} \\ ... & ... & ... & ... \\ m_{n1} & m_{n2} & ... & m_{nn}-1 \end{bmatrix}) = det(\begin{bmatrix} 0 & 0 & ... & 0 \\ m_{21} & m_{22}-1 & ... & m_{2n} \\ ... & ... & ... & ... \\ m_{n1} & m_{n2} & ... & m_{nn}-1 \end{bmatrix}) = 0$.

Since $det(M-1I) = 0$, 1 is a eigenvalue of any markov matrix.

Proof (part 2): all other eigenvalues are less than 1 (If you have a simpler way of proving this instead of proving it in two parts then feel free to do so but your proof should convince me about both these parts.)

Let us assume that there exists a $\lambda > 1$ and non-zero vector x such that $Ax = \lambda x$. Let x_i be the largest element of x. As any multiple of x will also satisfy $Ax = \lambda x$, we can take that $x_i > 0$. Since the rows of A are not negative and their sum is 1, every element in λx is a linear combination of the elements of x and all elements in λx must be $\leq x_i$. But as we assumed that $\lambda > 1$, $\lambda x_i > x_i$.

Hence we get a contradiction. Hence our assumption that there exist a $\lambda > 1$ is wrong.

Hence all other eigenvalues other than 1 are lesser than 1.

(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:

For k to be an eigenvalue for the matrix A, $det(A-kI)=0$ should be satisfied.
$det(A-kI) = det(\begin{bmatrix} a-k & b \\ c & d-k \end{bmatrix}) = (a-k)(d-k)-bc = ad-ka-kd+k^2-bc$
Since $a+b=c+d=k => c=k-d,\ b=k-a,$
$=> ad-ka-kd+k^2-(k-a)(k-d) = ad-ka-kd+k^2-k^2+ka+kd-ad=0$
Since $det(A-kI)=0$, k is an eigenvalue for the matrix.

(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:

YES
For the result to hold for any nxn matrix, we need $det(A-kI)=0$ where k is the sum of all elements along any row in matrix A.

$$A-kI = \begin{bmatrix} a_{11}-k & a_{12} & ... & a_{1n} \\ a_{21} & a_{22}-k & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn}-k \end{bmatrix}$$

Since we know from properties of determinants that adding a column to another column does not change the value of determinant, if we add all the columns with the first column, we get,

$$det(A-kI) = det(\begin{bmatrix} (\sum_{i=1}^{n} a_{1i})-k & a_{12} & ... & a_{1n} \\ (\sum_{i=1}^{n} a_{2i})-k & a_{22}-k & ... & a_{2n} \\ ... & ... & ... & ... \\ (\sum_{i=1}^{n} a_{ni})-k & a_{n2} & ... & a_{nn}-k \end{bmatrix})$$

Since we know that sum of elements of any row is 1, $=> det(A-kI) = det(\begin{bmatrix} k-k & a_{12} & ... & a_{1n} \\ k-k & a_{22}-k & ... & a_{2n} \\ ... & ... & ... & ... \\ k-k & a_{n2} & ... & a_{nn}-k \end{bmatrix}) = det(\begin{bmatrix} 0 & a_{12} & ... & a_{1n} \\ 0 & a_{22}-k & ... & a_{2n} \\ ... & ... & ... & ... \\ 0 & a_{n2} & ... & a_{nn}-k \end{bmatrix}).$
Since there is a zero column, determinant is 0.
Since $det(A-kI)=0$, k is a eigenvalue of matrix A.

(d) What is the corresponding eigenvector?

Solution:

Corresponding eigenvector for eigenvalue k is,

$$(A - kI)x = 0 => \begin{bmatrix} a_{11}-k & a_{12} & ... & a_{1n} \\ a_{21} & a_{22}-k & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn}-k \end{bmatrix} x = 0$$

If we take all values in x as equal, LHS becomes a summation over each row = k - k = 0 = RHS.

Hence the eigenvector can be any multiple of $\begin{bmatrix} 1 \\ ... \\ 1 \end{bmatrix}$ of size nx1.

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:

FALSE

Since i is a eigenvalue for A, $Au = iu => A^{-1}u = \frac{1}{i}u = -iu$. Hence eigenvalue of A^{-1} is -i.

(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

Since for any polynomial equation, $a_n x^n + a_{n-1}x^{n-1} + ... + a_0 = 0$, sum of its roots is given by $-\frac{a_{n-1}}{a_n}$,

Here, Sum of roots = Sum of eigenvalues = $-\frac{0}{1} = 0$.

(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:

TRUE

Let $u = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$, $v = \begin{bmatrix} 16 \\ -12 \\ 8 \end{bmatrix} = 4u$.

Since u is a eigenvector, $Au = \lambda u$, $Av = A(4u) = 4(Au) = 4(\lambda u) = \lambda(4u) = \lambda v$.

Hence as $Av = \lambda v$, it is also a eigenvector of A.

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

Solution:

TRUE

Since A is real-symmetric matrix, it is diagonalisable. $AS = S\lambda$ where S is a matrix with columns as the eigenvectors of A and λ is a diagonal matrix with eigenvalues as the diagonal elements.

Since S is invertible for symmetric matrix A, A has all eigenvectors independent.

Hence for any eigenvalue we will get corresponding eigenvectors equal to number of occurences of the eigenvalue (1 + number of repetitions).

Hence algebraic multiplicity = geometric multiplicity.

(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

Let $Ax = \lambda_A x$ and $Bx = \lambda_B x$.

For AB, $ABx = A(Bx) = A(\lambda_B x) = \lambda_B(Ax) = \lambda_A\lambda_B x = \lambda x$.

Hence, $(AB)x = \lambda x$ and so, x is a eigenvector to AB.

For BA, $BAx = B(Ax) = B(\lambda_A x) = \lambda_A(Bx) = \lambda_A\lambda_B x = \lambda x$.

Hence, $(BA)x = \lambda x$ and so, x is a eigenvector to BA.

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

Solution:

TRUE

Let $Ax = \lambda_A x$ and $Bx = \lambda_B x$.

$(A + B)x = Ax + Bx = \lambda_A x + \lambda_B x = (\lambda_A + \lambda_B)x = \lambda x$.

As $(A + B)x = \lambda x$, x is a eigenvector to A + B.

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

Solution:

TRUE

$AA^\top x = \lambda x => A^\top(AA^\top) = A^\top(\lambda x) => A^\top A(A^\top x) = \lambda(A^\top x)$. If we take $A^\top x = y$, $A^\top Ay = \lambda y$.

I.e. every non-zero eigenvalue in $AA^\top$ is also an eigenvalue for $A^\top A$.

Similarly, $A^\top Ax = \lambda x => A(A^\top A) = A(\lambda x) => AA^\top(Ax) = \lambda(Ax)$. If we take $Ax = z$, $AA^\top z = \lambda z$.

I.e. every non-zero eigenvalue in $A^\top A$ is also an eigenvalue for $AA^\top$.

Combining both, we see that non-zero eigenvalues of $AA^\top$ and $A^\top A$ are equal.

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

Solution:

FALSE

Suppose A is a 3x2 matrix, then $AA^\top$ will be a 3x3 matrix and $A^\top A$ will be a 2x2 matrix.

Hence eigenvectors of $AA^\top$ will be 3x1 and eigenvectors of $A^\top A$ will be 2x1. Hence they cant be the same.

Hence there exists matrix A where $AA^\top$ and $A^\top A$ have different eigenvectors.

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 the vector x represented in Basis 2 be $\begin{bmatrix} c \\ d \end{bmatrix}$. Therefore, $au_{11} + bu_{12} = cu_{21} + du_{22}$, where u_{11} and u_{12} correspond to basis vectors of basis 1 and u_{21} and u_{22} correspond to basis vectors of basis 2.

$=> \frac{1}{5}\begin{bmatrix} 3a+4b \\ 4a-3b \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 3c-4d \\ -4c-3d \end{bmatrix}$

$=> 3c - 4d = 3a + 4b => c = \frac{4d+(3a+4b)}{3}$

$=> 4c + 3d = -4a + 3b => 3d + \frac{16d+(12a+16b)}{3} = -4a + 3b$

$=> 9d + 16d + 12a + 16b = -12a + 9b => 25d = -24a - 7b => d = -\frac{(24a+7b)}{25}$

$=> c = -\frac{(7a-24b)}{25}$

The vector in basis 2 will be represented as $\begin{bmatrix} -\frac{(7a-24b)}{25} \\ -\frac{(24a+7b)}{25} \end{bmatrix} = \frac{-1}{25}\begin{bmatrix} 7a-24b \\ 24a+7b \end{bmatrix}$

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:

To prove, if $u \cdot v = T(u) \cdot T(v)$ then T represents a orthonormal basis.

Let the transformation T be represented using a matrix A. Therefore, given that $u^\top v = (Au)^\top (Av) \Rightarrow u^\top v = u^\top (A^\top A)v$, since u and v are vectors in the standard basis, $\Rightarrow A^\top A = I$, which means that A is a orthogonal matrix and hence T represents a orthonormal basis as all column vectors in A which correspond to basis vectors for T are orthonormal.

To prove, if T represents a orthonormal basis then $u \cdot v = T(u) \cdot T(v)$.

Let the transformation T be represented using a matrix A.

Since T represents a orthonormal basis, the basis vectors of T are orthonormal. Since these basis vectors correspond to the columns of A, A will be a orthogonal matrix. Hence, $A^\top A = I$.

Therefore, $LHS = u \cdot v = u^\top v = u^\top (I)v = u^\top (A^\top A)v = (u^\top A^\top)(Av) = (Au)^\top (Av) = (Au) \cdot (Av) = T(u) \cdot T(v) = RHS$.

Hence if T represents a orthonormal basis then $u \cdot v = T(u) \cdot T(v)$.

Since both are proved, $u \cdot v = T(u) \cdot T(v)$ if and only if T represents a orthonormal basis.

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:

In SVD, $A = U \sum V^\top$ where A is a mxn matrix, U is a mxm matrix, $\sum$ is a mxn matrix and V is a nxn matrix.

In SVD, V is the eigenvectors of $A^\top A$.

In PCA, principal directions are the eigenvectors of $A^\top A$ and hence it is same as V from SVD of A.

Also, in PCA, pricipal components are formed using $XV = U \sum V^\top V = U \sum$ as V is a orthogonal matrix (matrix of eigenvectors for symmetric matrix $A^\top A$ is orthogonal).

Hence, principal components in PCA is given by $U \sum$ where U is a mxm matrix and $\sum$ is a mxn matrix from SVD of A and the principal components forms a mxn matrix.

Also, principal directions in PCA is given by V where V is a nxn matrix from SVD of A.

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:

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

For eigenvalues, $det(A^\top A - \lambda I) = 0 \Rightarrow (41 - \lambda)^2 - 81 = 0$

$=> (41 - \lambda)^2 = 81$

$=> (41 - \lambda) = 9 => \lambda = 32$ or $=> (41 - \lambda) = -9 => \lambda = 50$

Eigenvalues are 32 and 50.

For eigenvectors for eigenvalue 32, $(A^\top A - 32I)x = 0 => \begin{bmatrix} 9 & 9 \\ 9 & 9 \end{bmatrix} x = 0$

$=> x = a \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

For eigenvectors for eigenvalue 50, $(A^\top A - 50I)x = 0 => \begin{bmatrix} -9 & 9 \\ 9 & -9 \end{bmatrix} x = 0$

$=> x = a \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Hence eigenvalues are 50 and 32 and corresponding eigenvectors are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ respectively.

Since $\sum^\top \sum = \begin{bmatrix} 50 & 0 \\ 0 & 32 \end{bmatrix} => \sum^2 = \begin{bmatrix} 50 & 0 \\ 0 & 32 \end{bmatrix} => \sum = \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix}$

Hence, $\sum = \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix}$ and $V = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

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

Solution:

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

For eigenvalues, $det(AA^\top - \lambda I) = 0 => (50 - \lambda)(32 - \lambda) - 0 = 0$

$=> (50 - \lambda)(32 - \lambda) = 0 => \lambda = 50, 32.$

Eigenvectors for 50, $(AA^\top - 50I)x = 0 => \begin{bmatrix} 0 & 0 \\ 0 & -18 \end{bmatrix} x = 0$

$=> x = a \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Eigenvectors for 32, $(AA^\top - 32I)x = 0 => \begin{bmatrix} 18 & 0 \\ 0 & 0 \end{bmatrix} x = 0$

$=> x = a \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Hence eigenvalues are 50 and 32 and corresponding eigenvectors are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ respectively.

Since $\sum^\top \sum = \begin{bmatrix} 50 & 0 \\ 0 & 32 \end{bmatrix} => \sum^2 = \begin{bmatrix} 50 & 0 \\ 0 & 32 \end{bmatrix} => \sum = \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix}$

Hence, $\sum = \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix}$ and $U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

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

Solution: 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)

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

We did not get back A since we just took the eigenvectors of $AA^\top$ and $A^\top A$ for U and V, without normalising them.

Hence if we normalise the eigenvectors for $AA^\top$, then U is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

and V is $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

Now, $U \sum V^\top = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5\sqrt{2} & 0 \\ 0 & 4\sqrt{2} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 5\sqrt{2} & 5\sqrt{2} \\ -4\sqrt{2} & 4\sqrt{2} \end{bmatrix} = \begin{bmatrix} 5 & 5 \\ -4 & 4 \end{bmatrix}$
$= A$.

Now we get back A by taking normalised eigenvectors for U and V.

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:

Let A be a mxn matrix with rank r.

Let $u_1, ...u_r$ be a orthonormal basis for $C(A)$.

Let $u_{r+1}, ...u_m$ be a orthonormal basis for $N(A^\top)$.

Let $v_1, ...v_r$ be a orthonormal basis for $C(A^\top)$.

Let $v_{r+1}, ...v_n$ be a orthonormal basis for $N(A)$.

Now the vectors, $u_1, ...u_r, u_{r+1}, ...u_m$ are orthonormal as every 2 vectors within $u_1, ...u_r$ and $u_{r+1}, ...u_m$ are orthonormal (given) and a vector from first and another vector from second are orthogonal with each other since we know that $N(A^\top) \perp C(A)$.

Similarly as $N(A) \perp C(A^\top)$, $v_1, ...v_r, v_{r+1}, ...v_n$ are orthonormal.

Now we want, $Av_i = \lambda_i u_i$ for all $i \leq r$.

Hence, $A \begin{bmatrix} v_1 & ... & v_r \end{bmatrix} = \begin{bmatrix} u_1 & ... & u_r \end{bmatrix} \begin{bmatrix} \lambda_1 & ... & 0 \\ ... & ... & ... \\ 0 & ... & \lambda_r \end{bmatrix}$

Now, since $v_{r+1}...v_n \in N(A)$, for any vector v from them, $Av = 0$.

Using this fact, we can expand the formula to include the remaining vectors as,

$$A \begin{bmatrix} v_1 & ... & v_r & v_{r+1} & ... & v_n \end{bmatrix} = \begin{bmatrix} u_1 & ... & u_r & u_{r+1} & ... & u_m \end{bmatrix} \begin{bmatrix} \lambda_1 & ... & 0 & 0 & ... & 0 \\ ... & ... & ... & 0 & ... & 0 \\ 0 & ... & \lambda_r & 0 & ... & 0 \\ 0 & ... & 0 & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & 0 & ... & 0 \end{bmatrix}$$

Here we have appended (n-r) 0 columns and (m-r) 0 rows to the matrix with λs.

This equation is still satisfied as in LHS since the extra vectors belong to null space of A, Av just gives 0 for each of them. Correspondingly, in RHS, since the λ matrix has (m-r) 0 rows, the weights for the extra $u_{r+1}...u_m$ vectors are 0 and hence dont affect the equation. Similary since it has (n-r) 0 columns, the last (n-r) columns in RHS finally will be 0s which matches with LHS.

Hence, LHS = RHS here.

Since the λ matrix is a diagonal matrix, this is of our required form in SVD which is, $AV = U\sum$.

Hence, it is clear that V contains the basis vectors for $C(A^\top)$ and $N(A)$ while U contains the basis vectors for $C(A)$ and $N(A^\top)$. So, U and V contain the basis vectors of the 4 fundamental subspaces of A.

...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.