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.

Kartik Bharadwaj
Name and Signature

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

Solution: Yes

Eigenstory: Special Properties

2. (1 point) Prove that for any square matrix A the eigenvectors corresponding to distinct eigenvalues are always independent.

Solution: Given an $n \times n$ matrix A, let there be r distinct eigenvalues having corresponding eigenvectors. Then, for $c_1 = c_2 = \cdots = c_r = 0$, we would like to prove:

$$c_1x_1 + c_2x_2 + \cdots + c_rx_r = 0 \tag{1}$$

By definition of eigenvalues and eigenvectors, we know that $Ax_i = \lambda_i x_i$, for $i \in [1, r]$. Multiplying (1) by A we get,

$$\begin{aligned} A(c_1x_1 + c_2x_2 + \cdots + c_rx_r) &= 0 \\ c_1Ax_1 + c_2Ax_2 + \cdots + c_rAx_r &= 0 \\ c_1\lambda_1x_1 + c_2\lambda_2x_2 + \cdots + c_r\lambda_rx_r &= 0 \end{aligned} \tag{2}$$

Now let's multiply (1) by λ_1.

$$\begin{aligned} \lambda_1(c_1x_1 + c_2x_2 + \cdots + c_rx_r) &= 0 \\ c_1\lambda_1x_1 + c_2\lambda_1x_2 + \cdots + c_r\lambda_1x_r &= 0 \end{aligned} \tag{3}$$

Subtracting (2) and (3), we get:

$$c_2(\lambda_2 - \lambda_1)x_2 + c_3(\lambda_3 - \lambda_1)x_3 + \cdots + c_r(\lambda_r - \lambda_1)x_r = 0$$

Since $x_1, x_2, \cdots, x_r$ are non-zero eigen vectors whose eigen values are distinct, $\lambda_i \neq \lambda_j$. This means each term should be equal to 0. In other words, $c_2 = \cdots = c_r = 0$. Similarly, by multiplying different λ_i's to (1), we get $c_1 = c_2 = \cdots = c_r = 0$, thereby, making the eigenvectors independent.

3. (2 points) Prove the following.

(a) The sum of the eigenvalues of a matrix is equal to its trace.

Solution: The characteristic polynomial for a $n \times n$ matrix is:

$$\begin{aligned} det(A - \lambda I_n) &= (-1)^n \big(\lambda^n - (\mathrm{tr}A)\,\lambda^{n-1} + \cdots + (-1)^n \det A\big) &(1)\\ &= (-1)^n (\lambda - \lambda_1)\ldots(\lambda - \lambda_n) &(2)\end{aligned}$$

By comparing coefficients of (1) and (2),

$$\mathrm{tr}A = \lambda_1 + \cdots + \lambda_n$$

(b) The product of the eigenvalues of a matrix is equal to its determinant.

Solution: For a $n \times n$, we know that eigenvalues are the roots of the characteristic polynomial, $\det(A - \lambda I_n)$.

$$\begin{aligned} \det(A - \lambda I_n) &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{m2} & \cdots & a_{nn} - \lambda \end{vmatrix} \\ &= \prod_{i=1}^{n} (\lambda_i - \lambda) \end{aligned}$$

Setting $\lambda = 0$, we see that $\det(A) = \prod_{i=1}^{n} \lambda_i$.

4. (2 points) What is the relationship between the rank of a matrix and the number of non-zero eigenvalues? Explain your answer.

Solution: The rank of a matrix is equal to the number of non-zero eigenvalues if the algebraic multiplicity (AM) is equal to geometric multiplicity (GM) for every eigenvalue. Also, rank of a matrix is equal to the number of non-zero eigenvalues if the matrix is diagonalizable.

5. (1 point) 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.

Solution: A square symmetric matrix A can be diagonalized as $A = Q\Lambda Q^T$, where Q is an orthogonal matrix and Λ are the eigenvalues of matrix A. Also, LDU decomposition theorem, gives us:

$$
\begin{aligned}
A &= LDU \\
A &= LDL^T && (\because A \text{ is symmetric, } U = L^T)
\end{aligned}
$$

We can see that Λ and D are congruent in the sense that A has the same number of positive eigenvalues in Λ as D. But the eigenvalues of D are just the diagonal entries or pivots of A. Therefore, for a square symmetric matrix, number of positive pivots it has is the same as the number of positive eigenvalues.

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: For matrix R to be symmetric, $R = R^T$.

$$
\begin{aligned}
R^T &= (I - 2\mathbf{u}\mathbf{u}^T)^T \\
&= I^T - 2(\mathbf{u}^T)^T\mathbf{u}^T && (\because (A+B)^T = A^T + B^T) \\
&= I - 2\mathbf{u}\mathbf{u}^T && (\because (A^T)^T = A \text{ and } I^T = I) \\
&= R
\end{aligned}
$$

For matrix R to be orthogonal, $RR^T = R^TR = I$. Since R is symmetric, we have to prove $R^2 = I$ for R to be an orthogonal matrix.

$$
\begin{aligned}
R^2 &= (I - 2\mathbf{u}\mathbf{u}^T)^2 \\
&= I^2 + 4(\mathbf{u}\mathbf{u}^T)^2 - 4I\mathbf{u}\mathbf{u}^T \\
&= I^2 + 4(\mathbf{u}\mathbf{u}^T\mathbf{u}\mathbf{u}^T) - 4\mathbf{u}\mathbf{u}^T \\
&= I + 4(\mathbf{u}\mathbf{u}^T) - 4\mathbf{u}\mathbf{u}^T && (\because \mathbf{u} \text{ is a unit vector, } u^Tu = 1) \\
&= I
\end{aligned}
$$

Hence, R is symmetric and orthogonal.

(b) Let $\mathbf{u} = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\1\end{bmatrix}$. 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 $\mathbf{u} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$). 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: R is a reflection matrix which reflects any incoming vector about hyper-plane $H = \{a | u^T a = 0\}$. H contains vectors a orthogonal to u.

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

Solution: With $\mathbf{u} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, our matrix $R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$.

To calculate eigenvalues,

$$\begin{aligned} det(R - \lambda I) &= 0 \\ det(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}) &= 0 \\ det(\begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & -1 & -\lambda \end{bmatrix}) &= 0 \\ (1-\lambda)(\lambda^2 - 1) &= 0 \end{aligned} \tag{1}$$

From the above equation, $\lambda_1 = 1$, $\lambda_2 = 1$, and $\lambda_3 = -1$. Eigenvectors for λ_1 and λ_2 will be the same. To get an eigenvector,

$$(R - \lambda_1 I)\mathbf{x} = 0$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix} \mathbf{x} = 0$$

Here, $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$ are eigenvectors for $\lambda_1 = \lambda_2 = 1$.

Eigenvector for λ_3 is:

$$(R - \lambda_3 I)\mathbf{x} = 0$$

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \mathbf{x} = 0$$

Here, the first two columns are independent and the third column is a scalar multiple of the second column. Therefore, the RREF of the above matrix is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix} \mathbf{x} = 0$$

Here, $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ is an eigen vector for λ_3.

(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: Given $\mathbf{u}^T\mathbf{u} = 1$,

$$\begin{aligned} R &= I - 2\mathbf{u}\mathbf{u}^\top \\ R\mathbf{u} &= (I - 2\mathbf{u}\mathbf{u}^\top)\mathbf{u} && \text{(Multiplying } \mathbf{u} \text{ on both sides)} \\ R\mathbf{u} &= I\mathbf{u} - 2\mathbf{u}\mathbf{u}^\top\mathbf{u} \\ R\mathbf{u} &= \mathbf{u} - 2\mathbf{u} && (\mathbf{u} \text{ is an unit vector}) \\ R\mathbf{u} &= -\mathbf{u} \end{aligned}$$

Therefore, $\mathbf{u}$ is an eigen vector of R with eigenvalue equal to -1. Similarly, let $\mathbf{v}$ be orthogonal to $\mathbf{u}$. Then,

$$\begin{aligned} R &= I - 2\mathbf{u}\mathbf{u}^\top \\ R\mathbf{v} &= (I - 2\mathbf{u}\mathbf{u}^\top)\mathbf{v} && \text{(Multiplying } \mathbf{v} \text{ on both sides)} \\ R\mathbf{v} &= I\mathbf{v} - 2\mathbf{u}\mathbf{u}^\top\mathbf{v} \\ R\mathbf{v} &= \mathbf{v} && (\because \mathbf{v}^T\mathbf{u} = 0) \end{aligned}$$

Therefore, $\mathbf{v}$ is an eigen vector of R with eigenvalue equal to 1. In essence, all vectors orthogonal to $\mathbf{u}$ have eigenvalue equal to 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 $|\lambda| = 1$

Solution: True. Let Q be a $n \times n$ orthogonal matrix with $\mathbf{x}$ as an eigenvector and λ its eigenvalue. Then,

$$Q\mathbf{x} = \lambda\mathbf{x}$$

Taking 2-norm on both sides,

$$\begin{aligned}
\|Q\mathbf{x}\|^2 &= \|\lambda\mathbf{x}\|^2 = |\lambda|^2\|\mathbf{x}\|^2 \\
(Q\mathbf{x})^T Q\mathbf{x} &= |\lambda|^2\|\mathbf{x}\|^2 \\
\mathbf{x}^T Q^T Q\mathbf{x} &= |\lambda|^2\|\mathbf{x}\|^2 \\
\mathbf{x}^T\mathbf{x} &= |\lambda|^2\|\mathbf{x}\|^2 && (\because Q^TQ = QQ^T = I) \\
\|\mathbf{x}\|^2 &= |\lambda|^2\|\mathbf{x}\|^2 \\
|\lambda|^2 &= 1 && (\because \mathbf{x} \text{ is a non-zero eigen vector}) \\
|\lambda| &= 1
\end{aligned}$$

(b) The eigenvectors of Q are orthogonal

Solution: True. We know that the eigenvalue of Q is $|\lambda| = 1$. Therefore, for $\lambda = 1$, $Q\mathbf{x} = \mathbf{x}$, and for $\lambda = -1$, $Q\mathbf{y} = -\mathbf{y}$. Hence,

$$\begin{aligned}
\mathbf{x}^T\mathbf{y} &= (Q\mathbf{x})^T(-Q\mathbf{y}) \\
\mathbf{x}^T\mathbf{y} &= -\mathbf{x}^TQ^TQ\mathbf{y} && (\because (AB)^T = B^TA^T) \\
\mathbf{x}^T\mathbf{y} &= -\mathbf{x}^T\mathbf{y} && (\because Q^TQ = I) \\
2\mathbf{x}^T\mathbf{y} &= 0 \\
\mathbf{x}^T\mathbf{y} &= 0
\end{aligned}$$

Hence, eigenvectors of Q are orthogonal.

(c) Q is always diagonalizable.

Solution: True. Let's assume that a real $n \times n$ orthogonal matrix Q is not diagonalizable. However, we know Q^TQ and QQ^T are real symmetric matrices which are indeed diagonalizable. The SVD of $Q^TQ = V\Sigma V^T$ and $QQ^T = U\Sigma U^T$. Hence, our SVD of $Q = U\Sigma V^T$.

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: For rank 1 matrix $\mathbf{u}\mathbf{v}^T$,

$$\begin{aligned}
\mathbf{u}\mathbf{v}^T x &= \lambda x \\
\mathbf{v}^T\mathbf{u}(\mathbf{v}^T x) &= \lambda(\mathbf{v}^T x) && \text{(Multiplying by } v^T\text{)} \\
\lambda &= \mathbf{v}^T\mathbf{u} && (\because \mathbf{v}^T x \text{ is a scalar})
\end{aligned}$$

Since rank 1 matrix $\mathbf{u}\mathbf{v}^T$ will have dependent rows, λ will also be equal to 0.

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

Solution: Since $\mathbf{u}\mathbf{v}^T$ is a rank one matrix, there will be $(n-1)$ eigenvalues equal to 0. However, if $\mathbf{u}$ and $\mathbf{v}$ are unit orthogonal vectors, then, there will be n eigenvalues equal to 0.

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

Solution: For a rank one matrix $\mathbf{u}\mathbf{v}^T$, column space is a line. One of the eigenvectors x will be along the line with eigenvalue $\mathbf{v}^T\mathbf{u}$. Another set of eigenvectors is when $\mathbf{u}\mathbf{v}^T y = 0$.

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

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

Solution: For a $n \times n$ Markov matrix M, we know that $M\mathbf{x} = \lambda\mathbf{x}$. Let there be a x_j such that $|x_j| \geq |x_i|$, for $1 \leq i \leq n$. Then, the j^{th} row will look like:

$$\begin{aligned}
\sum_{i=1}^{n} m_{ji} x_i &= \lambda x_j \\
|\lambda x_j| &= |\sum_{i=1}^{n} m_{ji} x_i| \\
|\lambda| \cdot |x_j| &\leq \sum_{i=1}^{n} m_{ji} |x_i| \\
|\lambda| \cdot |x_j| &\leq \sum_{i=1}^{n} m_{ji} |x_i| \\
|\lambda| \cdot |x_j| &\leq \sum_{i=1}^{n} m_{ji} |x_j| \\
|\lambda| \cdot |x_j| &\leq |x_j| \\
|\lambda| &\leq 1
\end{aligned}$$

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

Solution: Let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ matrix. Also, let assume that $a+b = c+d = k$. Then,

$$\begin{aligned}
\begin{bmatrix} a+b \\ c+d \end{bmatrix} - k \begin{bmatrix} 1+0 \\ 0+1 \end{bmatrix} &= 0 \\
det(M - kI) &= 0 \\
(M - kI) \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} &= 0 \\
M \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} &= kI \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\
M \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} &= k \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}
\end{aligned}$$

Here, $\lambda = k$ and its eigenvector $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$.

(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 M be an $n \times n$ matrix. Also, let assume that $\sum_{j=1}^{n} m_{ij} = d$ and

$\sum_{j=1}^{n} I_{ij} = 1$. Then,

$$\sum_{j=1}^{n} m_{ij} - d\sum_{j=1}^{n} I_{ij} = 0$$

$$\sum_{j=1}^{n} m_{ij} - dI_{ij} = 0$$

$$(M - dI)\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = 0$$

$$M\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = dI\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$$

$$M\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} = d\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$$

Here, $\lambda = d$ and its eigenvector $\mathbf{x} = \begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$.

(d) What is the corresponding eigenvector?

Solution: Invalid question.

Eigenstory: Special Relations

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

(a) The eigenvalues of A^T are **always** the same as that of A.

Solution: True. Let the characteristic polynomial of A be equal to $det(A - \lambda I_n)$.

Now, let the characteristic polynomial of A^T be:

$$\begin{aligned}
&= det(A^T - \lambda I_n) \\
&= det(A^T - \lambda I_n^T) \\
&= det((A - \lambda I_n)^T) && (\because \det(A+B)^T = \det(A^T) + \det(B^T)) \\
&= det(A - \lambda I_n) && (\because \det A^T = \det A)
\end{aligned}$$

Eigenvalues of A^T and A will be same since they have the same characteristic polynomial.

(b) The eigenvectors of A^T are **always** the same as that of A

Solution: False. For matrix A, we have,

$$(A - \lambda I_n)\mathbf{u} = 0$$

This means $\mathbf{u} \in \mathcal{N}(A)$. Similarly, for matrix A^T, we have,

$$(A^T - \lambda I_n)\mathbf{v} = 0 \tag{1}$$

Since $det(A^T - \lambda I_n) = det(A^T - \lambda I_n^T) = det((A - \lambda I_n)^T)$, (1) becomes,

$$(A - \lambda I_n)^T\mathbf{v} = 0$$

This means $\mathbf{v} \in \mathcal{N}(A^T)$ which may not be the same as $\mathbf{u}$. The eigenvectors of A and A^T aren't the always the same.

(c) The eigenvalues of A^{-1} are **always** the reciprocal of the eigenvalues of A.

Solution: True. Assuming A^{-1} exist, we know that for matrix A,

$$\begin{aligned}
A\mathbf{x} &= \lambda\mathbf{x} \\
A^{-1}A\mathbf{x} &= \lambda A^{-1}\mathbf{x} && \text{(Multiplying by } A^{-1} \text{ on both sides.)} \\
\mathbf{x} &= \lambda A^{-1}\mathbf{x} && (\because A^{-1}A = I) \\
A^{-1}\mathbf{x} &= \frac{1}{\lambda}\mathbf{x}
\end{aligned}$$

(d) The eigenvectors of A^{-1} are **always** the same as the eigenvectors of A.

Solution: True. Eigenvector $\mathbf{x}$ will be the same for both A and A^{-1} as followed from part (c) above.

(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 $A\mathbf{x} = \lambda_1\mathbf{x}$ and $B\mathbf{x} = \lambda_2\mathbf{x}$. Now, for matrix AB,

$$
\begin{aligned}
B\mathbf{x} &= \lambda_2\mathbf{x} \\
AB\mathbf{x} &= \lambda_2 A\mathbf{x} && \text{(Multiplying by matrix } A \text{ on both sides)} \\
AB\mathbf{x} &= \lambda_2\lambda_1\mathbf{x}. && (\because A\mathbf{x} = \lambda_1\mathbf{x})
\end{aligned}
$$

Here, we see that $\mathbf{x}$ is an eigenvector of matrix AB. Similarly, for matrix BA,

$$
\begin{aligned}
A\mathbf{x} &= \lambda_1\mathbf{x} \\
BA\mathbf{x} &= \lambda_1 B\mathbf{x} && \text{(Multiplying by matrix } B \text{ on both sides)} \\
BA\mathbf{x} &= \lambda_1\lambda_2\mathbf{x}. && (\because B\mathbf{x} = \lambda_2\mathbf{x})
\end{aligned}
$$

Here, we see that $\mathbf{x}$ is an eigenvector of matrix BA.

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

Solution: True. Let $A\mathbf{x} = \lambda_1\mathbf{x}$ and $B\mathbf{x} = \lambda_2\mathbf{x}$. Then, $(A + B)\mathbf{x} = (\lambda_1 + \lambda_2)\mathbf{x}$. Hence, $\mathbf{x}$ is also an eigenvector of $A + B$.

(g) If λ is an eigenvalue of A then $\lambda + k$ is an eigenvalue of $A + kI$.

Solution: True. For matrix $A + kI$,

$$
\begin{aligned}
&= (A + kI)\mathbf{x} \\
&= A\mathbf{x} + kI\mathbf{x} \\
&= \lambda I\mathbf{x} + kI\mathbf{x} && (\because A\mathbf{x} = \lambda\mathbf{x}) \\
&= (\lambda + k)I\mathbf{x} \\
&= (\lambda + k)\mathbf{x}
\end{aligned}
$$

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

Solution: True. For matrix $A^T A$,

$$
\begin{aligned}
A^T A\mathbf{x} &= \lambda\mathbf{x} \\
AA^T A\mathbf{x} &= \lambda A\mathbf{x} && \text{(Multiplying by matrix } A \text{ on both sides)} \\
AA^T(A\mathbf{x}) &= \lambda(A\mathbf{x}) \\
AA^T y &= \lambda y && \text{(Assigning } Ax = y\text{)}
\end{aligned}
$$

Eigenstory: Change of basis

11. (2 points) Consider the following two basis. Basis 1: $\mathbf{u_1} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \mathbf{u_2} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$, and Basis 2: $\mathbf{u_1} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}, \mathbf{u_2} = \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\-1\end{bmatrix}$. 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: Change of basis says that:

$$\begin{aligned}
Basis_2\mathbf{y} &= Basis_1\mathbf{x} \\
\mathbf{y} &= Basis_2^{-1}Basis_1\mathbf{x} \\
\mathbf{y} &= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\-1 & -1\end{bmatrix}^{-1}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix} \\
\mathbf{y} &= \begin{bmatrix}\frac{1}{2} & \frac{-1}{2}\\ \frac{-1}{2} & \frac{-1}{2}\end{bmatrix}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix} \\
\mathbf{y} &= \begin{bmatrix}b\\-a\end{bmatrix}
\end{aligned}$$

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 Q be the orthonormal basis represented by T. Let $T(\mathbf{u}) = \mathbf{u}_Q$ and $T(\mathbf{v}) = \mathbf{v}_Q$. Let matrix B be the standard basis. Therefore, change of basis says that, $Q\mathbf{u}_Q = B\mathbf{u}$ and $Q\mathbf{v}_Q = B\mathbf{v}$.

$$\begin{aligned}
\mathbf{u}^T\mathbf{v} &= (B^{-1}Q\mathbf{u}_Q)^T(B^{-1}Q\mathbf{v}_Q) \\
&= \mathbf{u}_Q^TQ^T(B^{-1})^TB^{-1}Q\mathbf{v}_Q \\
&= \mathbf{u}_Q^TQ^TQ\mathbf{v}_Q && (\because B \text{ is a standard basis, } (B^{-1})^TB^{-1} = I) \\
&= \mathbf{u}_Q^T\mathbf{v}_Q && (\because Q \text{ is an orthonormal matrix, } Q^TQ = I) \\
&= T(\mathbf{u})^TT(\mathbf{v})
\end{aligned}$$

Hence, $\mathbf{u} \cdot \mathbf{v} = T(\mathbf{u}) \cdot T(\mathbf{v})$.

Eigenstory: PCA and SVD

13. (1 point) How are PCA and SVD related? (no vague answers please, think and answer very precisely with mathematical reasoning)

Solution: Let matrix X be a $m \times n$ data matrix where m is the number of datapoints and n is the number of features. Assume our data matrix is centered. Let the covariance matrix of X be $C = \frac{X^T X}{m-1}$. Since C is a symmetric matrix, diagonalization of C is:

$$C = VDV^T$$

Columns of matrix V contains the principal components (or eigenvectors) of matrix C. D is a diagonal matrix containing variances (or eigenvalues) of each eigenvector of C. PCA takes the data matrix X, and performs change of basis from standard basis into the new basis as defined by eigenvector matrix V and eigenvalue matrix D.

Now, let the SVD of $X = U\Sigma V^T$. Then,

$$\begin{aligned} C &= X^T X \\ &= (U\Sigma V^T)^T (U\Sigma V^T) \\ &= (V\Sigma U^T U\Sigma V^T) \\ &= (V\Sigma^T \Sigma V^T) \\ &= VDV^T && \text{(where } D = \Sigma^T\Sigma) \end{aligned}$$

This means that the matrix V computed from SVD of X are the principal components of C and singular values of X, i.e. Σ, become the eigenvalues/variances in matrix D.

In essence, SVD decomposes a matrix into special matrices (or new eigen basis). PCA takes this new basis and manipulates the data matrix with ease.

14. (1½ points) Consider the matrix $\begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$

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

Solution: Let $B = A^T A$.

$$B = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}^T \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$$
$$B = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix}$$

Finding the eigenvalues of B,

$$\begin{aligned} det(B - \lambda I) &= 0 \\ \begin{vmatrix} 25-\lambda & 7 \\ 7 & 25-\lambda \end{vmatrix} &= 0 \\ (25-\lambda)^2 - 49 &= 0 \\ (25-\lambda)^2 &= 49 \\ (25-\lambda) &= \pm 7 \end{aligned}$$

Our eigenvalues are $\lambda_1 = 32$ and $\lambda_2 = 18$. Eigenvector for λ_1 is,

$$\begin{aligned} (B - \lambda_1 I)\mathbf{x}_1 &= 0 \\ \begin{bmatrix} 25-\lambda_1 & 7 \\ 7 & 25-\lambda_1 \end{bmatrix} \mathbf{x}_1 &= 0 \\ \begin{bmatrix} -7 & 7 \\ 7 & -7 \end{bmatrix} \mathbf{x}_1 &= 0 \end{aligned}$$

In the above equation, we can see that $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as the eigenvector for λ_1. By converting $\mathbf{x}_1$ into a unit length vector, $\mathbf{x}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Similarly, Eigenvector for λ_2 is,

$$\begin{aligned} (B - \lambda_2 I)\mathbf{x}_2 &= 0 \\ \begin{bmatrix} 25-\lambda_2 & 7 \\ 7 & 25-\lambda_2 \end{bmatrix} \mathbf{x}_2 &= 0 \\ \begin{bmatrix} 7 & 7 \\ 7 & 7 \end{bmatrix} \mathbf{x}_2 &= 0 \end{aligned}$$

In the above equation, we can see that $\mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ as the eigenvector for λ_2. By converting $\mathbf{x}_2$ into a unit length vector, $\mathbf{x}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Therefore, our $\Sigma = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$ and $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

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

Solution: Let $C = AA^T$.

$$C = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}^T$$
$$C = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$$

Since eigenvalues of a diagonal matrix are the diagonals, $\lambda_1 = 32$ and $\lambda_2 = 18$. Eigenvector for λ_1 is,

$$(C - \lambda_1 I)\mathbf{x} = 0$$
$$\begin{bmatrix} 32 - \lambda_1 & 0 \\ 0 & 18 - \lambda_1 \end{bmatrix} \mathbf{x} = 0$$
$$\begin{bmatrix} 0 & 0 \\ 0 & -14 \end{bmatrix} \mathbf{x} = 0$$

In the above equation, we can see that $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as the eigenvector for λ_1.
Similarly, Eigenvector for λ_2 is,

$$(C - \lambda_2 I)\mathbf{x} = 0$$
$$\begin{bmatrix} 32 - \lambda_2 & 0 \\ 0 & 18 - \lambda_2 \end{bmatrix} \mathbf{x} = 0$$
$$\begin{bmatrix} 14 & 0 \\ 0 & 0 \end{bmatrix} \mathbf{x} = 0$$

In the above equation, we can see that $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as the eigenvector for λ_2.
Therefore, our $\Sigma = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$ 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:

$$
\begin{aligned}
U\Sigma V^\top &= \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}32 & 0\\ 0 & 18\end{bmatrix}^{\frac{1}{2}}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}^T\\
&= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\sqrt{32} & 0\\ 0 & \sqrt{18}\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}\\
&= \begin{bmatrix}4 & 4\\ -3 & 3\end{bmatrix}
\end{aligned}
$$

Hence, $A = U\Sigma V^\top$.

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's say we have a $m \times n$ matrix A whose SVD is $A = U_1\ \Sigma_r\ V_1^T$ where $U_1 = U_{1\to r}$, an $m \times r$ matrix, $V_1 = V_{1\to r}$, an $n \times r$ matrix, and $1 \to r$ represents columns 1 to r.

For column space of A,
Let's assume that $\mathbf{b} \in \mathcal{C}(\mathcal{A})$. Then,

$$\mathbf{b} = A\mathbf{x} = U_1\Sigma_r V_1^T\mathbf{x} = U_1\mathbf{x}^*$$

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

$$\mathbf{b} = U_1\mathbf{y} = AV_1\Sigma_r^{-1}\mathbf{y} = A\mathbf{x}$$

Thus, $\mathbf{b} \in \mathcal{C}(\mathcal{A})$. Therefore, $\mathcal{C}(\mathcal{A}) = \mathcal{C}(\mathcal{U})_1$, or U_1 is an orthonormal basis for $\mathcal{C}(\mathcal{A})$.

For row space of A,
Let's assume that $\mathbf{b} \in \mathcal{C}(\mathcal{A}^T)$. Then,

$$\mathbf{b} = A^T\mathbf{y} = V_1\Sigma_r U_1^T\mathbf{y} = V_1\mathbf{y}^*$$

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

$$\mathbf{b} = V_1\mathbf{z} = A^T U_1\Sigma_r^{-1}\mathbf{z} = A^T\mathbf{y}$$

Thus, $\mathbf{b} \in \mathcal{C}(\mathcal{A}^{\mathcal{T}})$. Therefore, $\mathcal{C}(\mathcal{A}^{\mathcal{T}}) = \mathcal{C}(\mathcal{V})_1$, or V_1 is an orthonormal basis for $\mathcal{C}(\mathcal{A}^{\mathcal{T}})$.

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

Our Σ matrix has now changed. It is,

$$\begin{aligned}
\Sigma &= \begin{bmatrix} \Sigma_{r\times r} & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{bmatrix} \\
\begin{bmatrix} \Sigma_{r\times r} & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{bmatrix} &= \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} A \begin{bmatrix} V_1 & V_2 \end{bmatrix} \\
\begin{bmatrix} \Sigma_{r\times r} & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{bmatrix} &= \begin{bmatrix} U_1^T A V_1 & U_1^T A V_2 \\ U_2^T A V_1 & U_2^T A V_2 \end{bmatrix} \qquad (1) \\
\implies \begin{bmatrix} U_1^T A V_2 \\ U_2^T A V_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\implies \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} A V_2 &= 0 \\
\implies U^T A V_2 &= 0 \\
\implies A V_2 &= 0
\end{aligned}$$

Since there are $n - r$ columns in V_2 which is equal to $dim(\mathcal{N}(\mathcal{A}))$, and $V_2 \in \mathcal{N}(\mathcal{A})$, orthonormal columns of V_2 for a basis for $\mathcal{N}(\mathcal{A})$.

For left null space of A,

From (1),

$$\begin{aligned}
\begin{bmatrix} U_2^T A V_1 & U_2^T A V_2 \end{bmatrix} &= \begin{bmatrix} 0 & 0 \end{bmatrix} \\
\implies U_2^T A \begin{bmatrix} V_1^T & V_2^T \end{bmatrix} &= 0 \\
\implies U_2^T A V &= 0 \\
\implies U_2^T A &= 0
\end{aligned}$$

Since there are $m - r$ columns in U_2 which is equal to $dim(\mathcal{N}(\mathcal{A}^{\mathcal{T}}))$, and $U_2 \in \mathcal{N}(\mathcal{A}^{\mathcal{T}})$, orthonormal columns of U_2 for a basis for $\mathcal{N}(\mathcal{A}^{\mathcal{T}})$. Therefore,

1. U_1 forms an orthonormal basis for $\mathcal{C}(\mathcal{A})$
2. V_1 forms an orthonormal basis for $\mathcal{C}(\mathcal{A}^{\mathcal{T}})$
3. U_2 forms an orthonormal basis for $\mathcal{N}(\mathcal{A})$
4. V_2 forms an orthonormal basis for $\mathcal{N}(\mathcal{A}^{\mathcal{T}})$

16. (2 points) Fun with flags.

(a) Browse through the flags of all countries and paste 5 rank one flags below.

Solution:     
Indonesia, Monaco, Poland, Rome, Ukraine

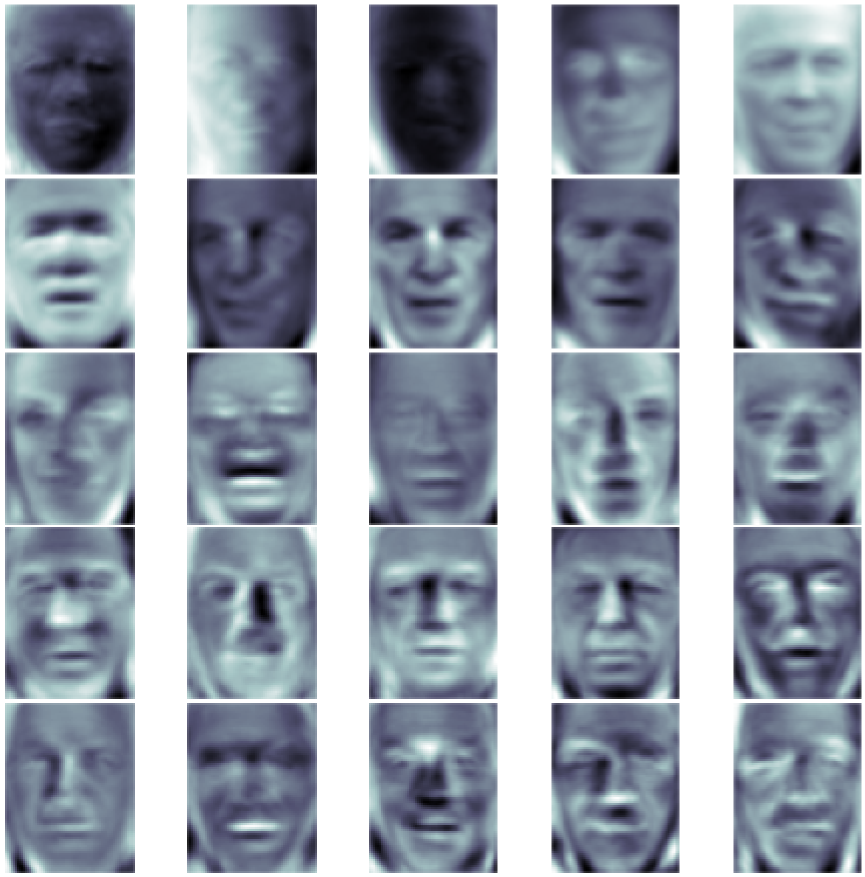
(b) What is the rank of the flag of Greece?

Solution: 3

17. (2 points) Consider the LFW dataset (Labeled Faces in the Wild).

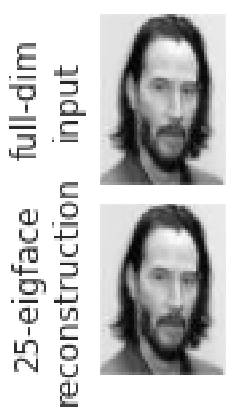
(a) Perform PCA using this dataset and plot the first 25 eigenfaces (in a 5×5 grid)

Solution:



(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.

Solution:



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