Homework 3 - Maths Foundation for Machine Learning

kipngeno koech - bkoech

October 14, 2024

Question 1: Eigen Values and Eigen Vectors

Consider the matrix $A = \begin{bmatrix} 9 & 1 & -1 \\ -1 & 11 & 1 \\ -2 & 2 & 10 \end{bmatrix}$.

1. Obtain the characteristic polynomial of A. (Hint: find the determinant of $A - \lambda I$)

$$A - \lambda I = \begin{bmatrix} 9 - \lambda & 1 & -1 \\ -1 & 11 - \lambda & 1 \\ -2 & 2 & 10 - \lambda \end{bmatrix}$$

2. From the characteristic polynomial, obtain the 3 eigenvalues of $\lambda_1, \lambda_2, \lambda_3$, such that $\lambda_1 < \lambda_2 < \lambda_3$.

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

$$\det \begin{bmatrix} 9 - \lambda & 1 & -1 \\ -1 & 11 - \lambda & 1 \\ -2 & 2 & 10 - \lambda \end{bmatrix} = 0$$

$$(9 - \lambda)((11 - \lambda)(10 - \lambda) - (2 \times 1)) - 1((-1)(10 - \lambda) - (-2 \times 1)) - 1((-1)(2) - (-2)(11 - \lambda)) = 0$$

$$(9 - \lambda)((11 - \lambda)(10 - \lambda) - 2) - 1((-10 + \lambda) + 2) - 1(-2 + 22 - 2\lambda) = 0$$

$$(9 - \lambda)(110 - 21\lambda + \lambda^2 - 2) - 1(-8 + \lambda) - 1(20 - 2\lambda) = 0$$

$$(9 - \lambda)(108 - 21\lambda + \lambda^2) - 1(-8 + \lambda) - 1(20 - 2\lambda) = 0$$

$$9(108 - 21\lambda + \lambda^2) - \lambda(108 - 21\lambda + \lambda^2) + 8 - \lambda - 20 + 2\lambda = 0$$

$$972 - 189\lambda + 9\lambda^2 - 108\lambda + 21\lambda^2 - \lambda^3 + 8 - \lambda - 20 + 2\lambda = 0$$

$$-\lambda^3 + 30\lambda^2 - 296\lambda + 960 = 0$$

$$\lambda^3 - 30\lambda^2 + 296\lambda - 960 = 0$$

$$(\lambda - 10)(\lambda^2 - 20\lambda + 96) = 0$$

$$(\lambda - 10)(\lambda - 8)(\lambda - 12) = 0$$

$$\lambda_1 = 8, \lambda_2 = 10, \lambda_3 = 12$$

3. Using Gaussian elimination, find the eigenvectors of A corresponding to the eigenvalues $\lambda_1, \lambda_2, \lambda_3$.
   for eigen value $\lambda_1 = 8$

$$A - 8I = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 3 & 1 \\ -2 & 2 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & -1 \\ -1 & 3 & 1 \\ -2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Augumented matrix:

$$\begin{bmatrix} 1 & 1 & -1 & | & 0 \\ -1 & 3 & 1 & | & 0 \\ -2 & 2 & 2 & | & 0 \end{bmatrix} \xrightarrow{R_2 = R_2 + R_1} \begin{bmatrix} 1 & 1 & -1 & | & 0 \\ 0 & 4 & 0 & | & 0 \\ -2 & 2 & 2 & | & 0 \end{bmatrix}$$

$$\xrightarrow{R_3 = R_3 + 2R_1} \begin{bmatrix} 1 & 1 & -1 & | & 0 \\ 0 & 4 & 0 & | & 0 \\ 0 & 4 & 0 & | & 0 \end{bmatrix} \xrightarrow{R_3 = R_3 - R_2} \begin{bmatrix} 1 & 1 & -1 & | & 0 \\ 0 & 4 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$$\xrightarrow{R_2 = \frac{1}{4}R_2} \begin{bmatrix} 1 & 1 & -1 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \xrightarrow{R_1 = R_1 - R_2} \begin{bmatrix} 1 & 0 & -1 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$$x_1 - x_3 = 0 \Rightarrow x_1 = x_3$$

$$x_2 = 0$$

Let $x_3 = 1$, then $x_1 = 1$ and $x_2 = 0$. Therefore, the eigenvector corresponding to $\lambda_1 = 8$ is $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

for eigen value $\lambda_2 = 10$

$$A - 10I = \begin{bmatrix} -1 & 1 & -1 \\ -1 & 1 & 1 \\ -2 & 2 & 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 1 & -1 \\ -1 & 1 & 1 \\ -2 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Augmented matrix:

$$\begin{bmatrix} -1 & 1 & -1 & | & 0 \\ -1 & 1 & 1 & | & 0 \\ -2 & 2 & 0 & | & 0 \end{bmatrix} \xrightarrow{R_1 = -R_1} \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ -1 & 1 & 1 & | & 0 \\ -2 & 2 & 0 & | & 0 \end{bmatrix}$$

$$\xrightarrow{R_2 = R_2 + R_1} \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & 0 & 2 & | & 0 \\ -2 & 2 & 0 & | & 0 \end{bmatrix} \xrightarrow{R_3 = R_3 + 2R_1} \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & 0 & 2 & | & 0 \\ 0 & 0 & 2 & | & 0 \end{bmatrix}$$

$$\xrightarrow{R_3 = R_3 - R_2} \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & 0 & 2 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \xrightarrow{R_2 = \frac{1}{2}R_2} \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$$\xrightarrow{R_1 = R_1 + R_2} \begin{bmatrix} 1 & -1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$$x_1 - x_2 = 0, x_3 = 0 \Rightarrow x_1 = x_2$$

Let $x_2 = 1$, then $x_1 = 1$ and $x_3 = 0$. Therefore, the eigenvector corresponding to $\lambda_2 = 10$ is $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

for eigen value $\lambda_3 = 12$

$$A - 12I = \begin{bmatrix} -3 & 1 & -1 \\ -1 & -1 & 1 \\ -2 & 2 & -2 \end{bmatrix}$$

$$\begin{bmatrix} -3 & 1 & -1 \\ -1 & -1 & 1 \\ -2 & 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -3 & 1 & -1 & | & 0 \\ -1 & -1 & 1 & | & 0 \\ -2 & 2 & -2 & | & 0 \end{bmatrix} \xrightarrow{R_1 = -\frac{1}{3}R_1} \begin{bmatrix} 1 & -\frac{1}{3} & \frac{1}{3} & | & 0 \\ -1 & -1 & 1 & | & 0 \\ -2 & 2 & -2 & | & 0 \end{bmatrix}$$

$$\xrightarrow{R_2 = R_2 + R_1} \begin{bmatrix} 1 & -\frac{1}{3} & \frac{1}{3} & | & 0 \\ 0 & -\frac{4}{3} & \frac{4}{3} & | & 0 \\ -2 & 2 & -2 & | & 0 \end{bmatrix} \xrightarrow{R_3 = R_3 + 2R_1} \begin{bmatrix} 1 & -\frac{1}{3} & \frac{1}{3} & | & 0 \\ 0 & -\frac{4}{3} & \frac{4}{3} & | & 0 \\ 0 & \frac{4}{3} & -\frac{4}{3} & | & 0 \end{bmatrix}$$

$$\xrightarrow{R_3 = R_3 + R_2} \begin{bmatrix} 1 & -\frac{1}{3} & \frac{1}{3} & | & 0 \\ 0 & -\frac{4}{3} & \frac{4}{3} & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \xrightarrow{R_2 = -\frac{3}{4}R_2} \begin{bmatrix} 1 & -\frac{1}{3} & \frac{1}{3} & | & 0 \\ 0 & 1 & -1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$$\xrightarrow{R_1 = R_1 + \frac{1}{3}R_2} \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & -1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$$x_1 = 0, x_2 = x_3$$

Let $x_2 = 1$, then $x_1 = 0$ and $x_3 = 1$. Therefore, the eigenvector corresponding to $\lambda_3 = 12$ is $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

4. (in the notebook)

5. Compare the eigen vectors computed using np.linalg.eig and that obtained using Gaussian elimination in (3) above. Are they the same? If they are the same, explain why this is the case. Otherwise, explain why they are different
   eigen vectors using np.linalg.eig:

$$\begin{pmatrix} 9.00258517e-16 & -7.07106781e-01 & 7.07106781e-01 \\ -7.07106781e-01 & -3.36518470e-16 & -7.07106781e-01 \\ -7.07106781e-01 & -7.07106781e-01 & -5.62025848e-15 \end{pmatrix}$$

eigen vectors using Gaussian elimination:

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

The eigen vectors are different. This is because the eigen vectors obtained using np.linalg.eig normalizes the eigen vectors to have a norm of 1. The eigen vectors obtained using Gaussian elimination are not normalized.

6. Now, given a diagonal matrix D, whose diagonal entries are the eigenvalues of A, and a matrix P whose columns are the eigenvectors of A, use the eigen values you obtained in (2) and the eigenvectors you obtained in (3) to show that A as PDP^{-1}.

$$D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 12 \end{bmatrix}, P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, A = \begin{bmatrix} 9 & 1 & -1 \\ -1 & 11 & 1 \\ -2 & 2 & 10 \end{bmatrix}$$

$$PDP^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 8 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 12 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}^{-1}$$

$$P^{-1} = \begin{pmatrix} 0.5 & -0.5 & 0.5 \\ 0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & 0.5 \end{pmatrix}$$

$$PDP^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 8 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 12 \end{bmatrix} \begin{bmatrix} 0.5 & -0.5 & 0.5 \\ 0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & 0.5 \end{bmatrix}$$

$$DP^{-1} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 12 \end{bmatrix} \begin{bmatrix} 0.5 & -0.5 & 0.5 \\ 0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 4 & -4 & 4 \\ 5 & 5 & -5 \\ -6 & 6 & 6 \end{bmatrix}$$

$$P(DP^{-1}) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & -4 & 4 \\ 5 & 5 & -5 \\ -6 & 6 & 6 \end{bmatrix} = \begin{bmatrix} 9 & 1 & -1 \\ -1 & 11 & 1 \\ -2 & 2 & 10 \end{bmatrix}$$

Therefore, $A = PDP^{-1}$.

7. Calculate the determinant of A, $det(A)$:

$$\det(A) = 9((11 \times 10) - (2 \times 1)) - 1((-1 \times 10) - (-2 \times 1)) - 1((-1 \times 2) - (11 \times -2))$$

$$= 9(110 - 2) - 1(-10 + 2) - 1(-2 + 22) = 9(108) - 1(-8) - 1(20) = 972 + 8 - 20 = \mathbf{960}$$

8. Calculate the trace of A, $tr(A)$:

$$tr(A) = 9 + 11 + 10 = \mathbf{30}$$

9. confirm the following holds:
$det(A) = \lambda_1 \lambda_2 \lambda_3$

$$960 = 8 \times 10 \times 12 = 960$$

$tr(A) = \lambda_1 + \lambda_2 + \lambda_3$

$$30 = 8 + 10 + 12 = 30$$

Question 2: Image Compression using SVD

5. 5. Assuming 1 byte of memory is required for one element of the image matrix, determine the top k rank-1 matrices required to represent the image if a compression ratio of 13:1 is desired. What is the value of k? Add up the rank-1 matrices for the top K singular values and visaulize your results. Remember to show your workings

$$\text{Compression ratio} = \frac{\text{Original size}}{\text{Compressed size}} = 13$$

Original size:

$$\text{Original size} = \text{Number of elements in the image} = 902 \times 602 = \textbf{543004} \text{ bytes}$$

$$\text{Compressed size} = \text{Number of elements in the top k rank-1 matrices} = k \times 902 + k \times 602 = 1504k \text{ bytes}$$

$$\frac{543004}{1504k} = 13 = \frac{543004}{13} = 1504k$$

$$k = \frac{543004}{13 \times 1504} = \textbf{28}$$