Eigen Decomposition

Mathematical basis for powerful machine learning technique of Principal Component Analysis (PCA)

Matrix $A_{n \times n}$ has eigenvector $\underline{u} \neq \underline{0}$ with eigenvalue λ if:

$$A \underline{u} = \lambda \underline{u}$$

$$\Rightarrow \quad (A - \lambda I)\underline{u} = 0$$

$\Rightarrow \quad A - \lambda I$ is not invertible (since $\underline{u} \neq \underline{0}$)

$$\Rightarrow \quad \underbrace{\det(A - \lambda I)}_{\text{polynomial in } \lambda \text{ - characteristic polynomial}} = 0 \quad \text{— } (*)$$

Solving for λ that satisfies $\det(A - \lambda I) = 0$, yields all eigenvalues of A. Eigenvector for eigenvalue λ obtained by solving $(A - \lambda I)\underline{u} = 0$

Example

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

$$A - \lambda I = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5-\lambda & 1 \\ 1 & 5-\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} 5-\lambda & 1 \\ 1 & 5-\lambda \end{vmatrix} = (5-\lambda)^2 - 1$$

$$= \lambda^2 - 10\lambda + 24 = (\lambda - 6)(\lambda - 4)$$

$$\det(A - \lambda I) = 0 \Rightarrow \lambda_1 = 6, \ \lambda_2 = 4$$

Finding eigenvectors $\underline{u}_j$, $j = 1, 2$

Solve for $\underline{u}_j$ that satisfies $(A - \lambda_j I)\,\underline{u}_j = \underline{0}$, $j = 1, 2$

$$A - \lambda_j I = \begin{bmatrix} 5-\lambda_j & 1 \\ 1 & 5-\lambda_j \end{bmatrix}$$

$$\lambda_1 = 6 : \quad \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \quad u_{11} - u_{12} = 0 \Rightarrow u_{11} = u_{12}$$

$$\Rightarrow \underline{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ or any scaled version}$$

Unit norm $\underline{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ or $\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ -1 \end{bmatrix}$

$$\lambda_1 = 4 : \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_{21} \\ u_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \quad u_{21} + u_{22} = 0 \Rightarrow u_{21} = -u_{22}$$

Unit norm $\underline{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ or $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Note that no matter how $\underline{u}_1$ and $\underline{u}_2$ are scaled,

$\underline{u}_1 \cdot \underline{u}_2 = \underline{u}_1^T \underline{u}_2 = 0$ — orthogonal

---

Repeated Eigenvalues

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

$$\det(A - \lambda I) = (1-\lambda)\left((2-\lambda)^2 - 1\right)$$
$$= (1-\lambda)(\lambda^2 - 4\lambda + 3) = (1-\lambda)(\lambda-1)(\lambda-3)$$

$$\det(A - \lambda I) = 0 \Rightarrow \lambda_1 = 3, \ \lambda_2 = 1, \ \lambda_3 = 1$$

$\lambda_1 = 3 : \; (A - 3I)\underline{u}_1 = \underline{0} \; \Rightarrow \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\Rightarrow \; u_{11} = 0, \quad u_{12} = -u_{13}$

$\Rightarrow$ unit norm $\underline{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$ or $\frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$

For $\lambda_2 = \lambda_3 = 1$,

$(A - I)\underline{u}_j = 0, \; j = 2, 3. \; \Rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} u_{j1} \\ u_{j2} \\ u_{j3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow u_{j2} = u_{j3}$

$\underline{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ and $\underline{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ unit norm eigenvectors

Note: $\underline{u}_2$ and $\underline{u}_3$ chosen to be orthogonal, and $\underline{u}_1, \underline{u}_2, \underline{u}_3$ are mutually orthogonal

---

Key Result For symmetric matrix A, $A^T = A$

1. Eigenvalues and eigenvectors are real-valued
2. Eigenvectors for distinct eigenvalues are orthogonal

   Proof Let $\underline{u}_1, \underline{u}_2$ be eigenvectors corresponding to eigenvalues λ_1, λ_2, $\lambda_1 \neq \lambda_2$. Then

   $$A\underline{u}_1 = \lambda_1 \underline{u}_1, \quad A\underline{u}_2 = \lambda_2 \underline{u}_2$$

   $$\lambda_1 \underline{u}_1^T \underline{u}_2 = (A\underline{u}_1)^T \underline{u}_2 = \underline{u}_1^T A \underline{u}_2 = \lambda_2 \underline{u}_1^T \underline{u}_2$$

   $$\Rightarrow (\lambda_1 - \lambda_2)\underline{u}_1^T \underline{u}_2 = 0 \Rightarrow \underline{u}_1^T \underline{u}_2 = 0$$

3. If λ is an eigenvalue of multiplicity k, then we can find k mutually orthogonal eigenvectors for that eigenvalue

Eigen (Spectral) Decomposition of Matrix

For symmetric matrix $A_{n\times n}$, $\lambda_1, \lambda_2, \ldots, \lambda_n$ are real and we can find $\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_n$ that are mutually orthogonal. Normalize $\underline{u}_j$ to unit norm. Then

$$\underline{u}_i^T \underline{u}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad \text{orthonormal}$$

$$\text{Define } U_{n\times n} = \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_n \end{bmatrix} \Rightarrow U^T = \begin{bmatrix} \underline{u}_1^T \\ \underline{u}_2^T \\ \vdots \\ \underline{u}_n^T \end{bmatrix}_{n\times n}$$

Then $U^T U = I_{n\times n} \Rightarrow U^{-1} = U^T$

i.e., U is an orthogonal or unitary matrix

$$\text{Define } D = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{bmatrix} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

$$A \underline{u}_i = \lambda_i \underline{u}_i, \quad i = 1, 2, \ldots, n$$

$$\Rightarrow A \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_n \end{bmatrix} = \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

$$\Rightarrow A U = U D \Rightarrow A U U^T = U D U^T \Rightarrow A = U D U^T$$

$$A = U D U^T$$

$$= \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \begin{bmatrix} \underline{u}_1^T \\ \underline{u}_2^T \\ \vdots \\ \underline{u}_n^T \end{bmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i \underline{u}_i \underline{u}_i^T.$$

Example $\quad A = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}, \quad \lambda_1 = 6, \quad \lambda_2 = 4$

$$\underline{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \underline{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad U^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$UU^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix}$$

$$\begin{aligned} UDU^T &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 6 & 6 \\ 4 & -4 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 10 & 2 \\ 2 & 10 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix} \end{aligned}$$

Also, $\lambda_1 \underline{u}_1 \underline{u}_1^T + \lambda_2 \underline{u}_2 \underline{u}_2^T$

$$\begin{aligned} &= 6 \cdot \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} + 4 \cdot \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix} \\ &= 3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + 2\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix} \end{aligned}$$