Module 6.3 : Eigenvalue Decomposition

Before proceeding let's do a quick recap of eigenvalue decomposition.

- Let $u_1, u_2, \ldots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues.

- Let $u_1, u_2, \ldots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues.
- Consider a matrix U whose columns are $u_1, u_2, \ldots, u_n$.

- Let $u_1, u_2, \ldots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues.
- Consider a matrix U whose columns are $u_1, u_2, \ldots, u_n$.
- Now

$$AU =$$

- Let $u_1, u_2, \ldots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues.
- Consider a matrix U whose columns are $u_1, u_2, \ldots, u_n$.
- Now

$$AU = A \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \ldots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

- Let $u_1, u_2, \ldots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues.
- Consider a matrix U whose columns are $u_1, u_2, \ldots, u_n$.
- Now

$$AU = A \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \ldots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ Au_1 & Au_2 & \ldots & Au_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

- Let $u_1, u_2, \ldots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues.
- Consider a matrix U whose columns are $u_1, u_2, \ldots, u_n$.
- Now

$$AU = A\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \ldots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ Au_1 & Au_2 & \ldots & Au_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \lambda_1 u_1 & \lambda_2 u_2 & \ldots & \lambda_n u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

- Let $u_1, u_2, \ldots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues.
- Consider a matrix U whose columns are $u_1, u_2, \ldots, u_n$.
- Now

$$AU = A \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \ldots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ Au_1 & Au_2 & \ldots & Au_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \lambda_1 u_1 & \lambda_2 u_2 & \ldots & \lambda_n u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \ldots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

- Let $u_1, u_2, \ldots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues.
- Consider a matrix U whose columns are $u_1, u_2, \ldots, u_n$.
- Now

$$
\begin{aligned}
AU = A \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \ldots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} &= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ Au_1 & Au_2 & \ldots & Au_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \\
&= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \lambda_1 u_1 & \lambda_2 u_2 & \ldots & \lambda_n u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \\
&= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \ldots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}
\end{aligned}
$$

- Let $u_1, u_2, \dots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues.
- Consider a matrix U whose columns are $u_1, u_2, \dots, u_n$.
- Now

$$AU = A\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \dots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ Au_1 & Au_2 & \dots & Au_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \lambda_1 u_1 & \lambda_2 u_2 & \dots & \lambda_n u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \dots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} = U\Lambda$$

- Let $u_1, u_2, \ldots, u_n$ be the eigenvectors of a matrix A and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues.
- Consider a matrix U whose columns are $u_1, u_2, \ldots, u_n$.
- Now

$$AU = A \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \ldots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ Au_1 & Au_2 & \ldots & Au_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \lambda_1 u_1 & \lambda_2 u_2 & \ldots & \lambda_n u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \ldots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix} = U\Lambda$$

- where Λ is a diagonal matrix whose diagonal elements are the eigenvalues of A.

$$AU = U\Lambda$$

$$AU = U\Lambda$$

- If U^{-1} exists, then we can write,

$$A = U\Lambda U^{-1} \quad \text{[eigenvalue decomposition]}$$

$$U^{-1}AU = \Lambda \quad \text{[diagonalization of A]}$$

$$AU = U\Lambda$$

- If U^{-1} exists, then we can write,

$$A = U\Lambda U^{-1} \qquad \text{[eigenvalue decomposition]}$$
$$U^{-1}AU = \Lambda \qquad \text{[diagonalization of A]}$$

- Under what conditions would U^{-1} exist?

$$AU = U\Lambda$$

- If U^{-1} exists, then we can write,

$$A = U\Lambda U^{-1} \quad \text{[eigenvalue decomposition]}$$
$$U^{-1}AU = \Lambda \quad \text{[diagonalization of A]}$$

- Under what conditions would U^{-1} exist?
  - If the columns of U are linearly independent [**See proof here**]

$$AU = U\Lambda$$

- If U^{-1} exists, then we can write,

$$A = U\Lambda U^{-1} \quad \text{[eigenvalue decomposition]}$$
$$U^{-1}AU = \Lambda \quad \text{[diagonalization of A]}$$

- Under what conditions would U^{-1} exist?
  - If the columns of U are linearly independent [**See proof here**]
  - *i.e.* if A has n linearly independent eigenvectors.

$$AU = U\Lambda$$

- If U^{-1} exists, then we can write,

$$A = U\Lambda U^{-1} \quad \text{[eigenvalue decomposition]}$$
$$U^{-1}AU = \Lambda \quad \text{[diagonalization of A]}$$

- Under what conditions would U^{-1} exist?
  - If the columns of U are linearly independent [**See proof here**]
  - *i.e.* if A has n linearly independent eigenvectors.
  - *i.e.* if A has n distinct eigenvalues [**sufficient condition, proof : Slide 19 Theorem 1**]

- If A is symmetric then the situation is even more convenient.

- If A is symmetric then the situation is even more convenient.
- The eigenvectors are orthogonal [**proof : Slide 19 Theorem 2**]

- If A is symmetric then the situation is even more convenient.
- The eigenvectors are orthogonal [**proof : Slide 19 Theorem 2**]
- Further let's assume, that the eigenvectors have been normalized [$u_i^T u_i = 1$]

$$Q = U^T U = \begin{bmatrix} \leftarrow u_1 \rightarrow \\ \leftarrow u_2 \rightarrow \\ \cdots \\ \leftarrow u_n \rightarrow \end{bmatrix} \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \dots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

- If A is symmetric then the situation is even more convenient.
- The eigenvectors are orthogonal **[proof : Slide 19 Theorem 2]**
- Further let's assume, that the eigenvectors have been normalized [$u_i^T u_i = 1$]

$$Q = U^T U = \begin{bmatrix} \leftarrow u_1 \rightarrow \\ \leftarrow u_2 \rightarrow \\ \dots \\ \leftarrow u_n \rightarrow \end{bmatrix} \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \dots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

- Each cell of the matrix, Q_{ij} is given by $u_i^T u_j$

$$\begin{aligned} Q_{ij} = u_i^T u_j \quad &= 0 \text{ if } i \neq j \\ &= 1 \text{ if } i = j \end{aligned}$$

$$\therefore U^T U = \mathbb{I} \text{ (the identity matrix)}$$

- If A is symmetric then the situation is even more convenient.
- The eigenvectors are orthogonal **[proof : Slide 19 Theorem 2]**
- Further let's assume, that the eigenvectors have been normalized [$u_i^T u_i = 1$]

$$Q = U^T U = \begin{bmatrix} \leftarrow u_1 \rightarrow \\ \leftarrow u_2 \rightarrow \\ \dots \\ \leftarrow u_n \rightarrow \end{bmatrix} \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ u_1 & u_2 & \dots & u_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

- Each cell of the matrix, Q_{ij} is given by $u_i^T u_j$

$$\begin{aligned} Q_{ij} = u_i^T u_j \quad &= 0 \text{ if } i \neq j \\ &= 1 \text{ if } i = j \end{aligned}$$

$$\therefore U^T U = \mathbb{I} \text{ (the identity matrix)}$$

- U^T is the inverse of U (very convenient to calculate)

Something to think about

- Given the EVD, $A = U\Sigma U^T$,
  what can you say about the sequence $x_0, Ax_0, A^2x_0, \ldots$ in terms of the eigen values of A.
  (Hint: You should arrive at the same conclusion we saw earlier)

Theorem (one more important property of eigenvectors)

If A is a square symmetric $N \times N$ matrix, then the solution to the following optimization problem is given by the eigenvector corresponding to the largest eigenvalue of A.

$$\max_x \; x^T A x$$
$$\text{s.t } \|x\| = 1$$

and the solution to

$$\min_x \; x^T A x$$
$$\text{s.t } \|x\| = 1$$

is given by the eigenvector corresponding to the smallest eigenvalue of A.

Proof: Next slide.

- This is a constrained optimization problem that can be solved using Lagrange Multipliers:

$$L = x^T A x - \lambda(x^T x - 1)$$
$$\frac{\partial L}{\partial x} = 2Ax - \lambda(2x) = 0 => Ax = \lambda x$$

- This is a constrained optimization problem that can be solved using Lagrange Multipliers:

$$L = x^T A x - \lambda(x^T x - 1)$$
$$\frac{\partial L}{\partial x} = 2Ax - \lambda(2x) = 0 => Ax = \lambda x$$

- Hence x must be an eigenvector of A with eigenvalue λ.

- This is a constrained optimization problem that can be solved using Lagrange Multipliers:

$$L = x^T Ax - \lambda(x^T x - 1)$$
$$\frac{\partial L}{\partial x} = 2Ax - \lambda(2x) = 0 => Ax = \lambda x$$

- Hence x must be an eigenvector of A with eigenvalue λ.
- Multiplying by x^T:

$$x^T Ax = \lambda x^T x = \lambda(\text{since } x^T x = 1)$$

- This is a constrained optimization problem that can be solved using Lagrange Multipliers:

$$L = x^T Ax - \lambda(x^T x - 1)$$
$$\frac{\partial L}{\partial x} = 2Ax - \lambda(2x) = 0 => Ax = \lambda x$$

- Hence x must be an eigenvector of A with eigenvalue λ.
- Multiplying by x^T:

$$x^T Ax = \lambda x^T x = \lambda (\text{since } x^T x = 1)$$

- Therefore, the critical points of this constrained problem are the eigenvalues of A.

- This is a constrained optimization problem that can be solved using Lagrange Multipliers:

$$L = x^T Ax - \lambda(x^T x - 1)$$
$$\frac{\partial L}{\partial x} = 2Ax - \lambda(2x) = 0 => Ax = \lambda x$$

- Hence x must be an eigenvector of A with eigenvalue λ.
- Multiplying by x^T:

$$x^T Ax = \lambda x^T x = \lambda (\text{since } x^T x = 1)$$

- Therefore, the critical points of this constrained problem are the eigenvalues of A.
- The maximum value is the largest eigenvalue, while the minimum value is the smallest eigenvalue.

The story so far...

The story so far...

- The eigenvectors corresponding to different eigenvalues are linearly independent.

The story so far...

- The eigenvectors corresponding to different eigenvalues are linearly independent.
- The eigenvectors of a square symmetric matrix are orthogonal.

The story so far...

- The eigenvectors corresponding to different eigenvalues are linearly independent.
- The eigenvectors of a square symmetric matrix are orthogonal.
- The eigenvectors of a square symmetric matrix can thus form a convenient basis.

The story so far...

- The eigenvectors corresponding to different eigenvalues are linearly independent.
- The eigenvectors of a square symmetric matrix are orthogonal.
- The eigenvectors of a square symmetric matrix can thus form a convenient basis.
- We will put all of this to use.